Article:
Wang, Y. orcid.org/0009-0006-4353-9278, Ma, X. orcid.org/0009-0009-8746-656X, Robson, A.J. orcid.org/0000-0002-1449-9477 et al. (2 more authors) (2025) A hybrid machine learning approach to predict and evaluate surface chemistries of films deposited via APPJ. Plasma Processes and Polymers. ISSN 1612-8850

https://doi.org/10.1002/ppap.70035

Plasma Processes and Polymers

PLASMA PROCESSES AND POLYMERS

RESEARCH ARTICLE OPEN ACCESS

A Hybrid Machine Learning Approach to Predict and Evaluate Surface Chemistries of Films Deposited via APPJ

Yong Wang[1] | Xudong Ma[2] | Alexander J. Robson[3] | Robert D. Short[3] | James W. Bradley[1]

[1]Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, UK | [2]School of Computer Science, University of Bristol, Bristol, UK | [3]Department of Chemistry, University of Sheffield, Sheffield, UK

Correspondence: James W. Bradley (jbradley@liverpool.ac.uk)

Received: 30 March 2025 | **Accepted:** 28 April 2025

Funding: The authors thank the Engineering and Physical Sciences Research Council (EPSRC) for supporting this work with grants EP/S005153/1 and EP/S004505/1. The authors are also thankful to the Chinese Scholarship Council for their support.

Keywords: deep learning | films | machine learning | plasma polymerization | TEMPO

ABSTRACT

We developed a hybrid machine learning model, integrating Artificial Neural Network (ANN), Random Forest (RF) and AdaBoost (AB), to predict and evaluate the plasma polymerization process of TEMPO monomer, specifically for Nitric Oxide films. This model is specifically designed to adeptly navigate the intricate landscape of the plasma polymerization process. Through genetic algorithm optimization, we have fine-tuned our hybrid model's algorithm weights, achieving results that closely match experimental data. TEMPO-Helium flow ratio is identified as the most critical parameter for the surface N percentage, with a relative importance of 41%. Frequency has the greatest influence on the N-O percentage, with a relative importance of 30%. The intertwined influence of different polymerization parameters on the film's surface chemistry has been detailed.

1 | Introduction

Nitric oxide (NO) films find diverse and crucial applications in the biomedical field [1, 2]. These films, capable of controlled NO release, contribute to wound healing by promoting angiogenesis and collagen synthesis while also serving as antimicrobial coatings on medical devices to prevent infections [3]. They act as drug delivery systems for localized therapy, aid in cardiovascular health by vasodilation, and have potential in cancer therapy and neuroprotection [4]. NO-releasing films extend their utility to respiratory health, dental applications, bioimaging, tissue engineering, and anti-inflammatory treatments [5–8]. Moreover, they play a role in diabetes management and can be used in various applications to improve overall health outcomes by harnessing the therapeutic properties of nitric oxide while minimizing systemic side effects [9].

In recent years, atmospheric pressure plasma jets (APPJs) have emerged as versatile tools for thin film deposition with diverse applications. These plasma jets are employed to deposit functional coatings, including antibacterial, hydrophobic, and hydrophilic films, enhancing wear resistance, corrosion protection, and surface properties on a wide range of substrates [10–14]. They find utility in photovoltaic and solar cell production, as well as in flexible electronics, improving energy efficiency and enabling flexible device fabrication [15–17]. APPJs are instrumental in modifying polymer surfaces, producing optical and membrane coatings, and enhancing biomedical devices [18, 19]. They also contribute to gas barrier films for packaging, gas sensors, and energy storage applications, demonstrating their broad impact across industries through precise and controlled film deposition processes [17, 20, 21]. However, the creation and

characterization of films through atmospheric pressure plasma polymerization can be a time-consuming process, with the need to carefully adjust and optimize various experimental parameters such as applied voltage, frequency, monomer flow rate, carrier gas flow rate, and sample-nozzle distance, which directly influence the properties of the resulting films. Due to the multifaceted nature of parameter control and the comprehensive analysis required, considerable effort is invested in both film fabrication and characterization.

Machine learning (ML), a branch of artificial intelligence, has revolutionized scientific research by demonstrating exceptional performance, particularly when provided with labeled data [22, 23], which is called supervised ML. In scientific studies, labeled data refers to information that has been meticulously annotated or categorized, allowing ML algorithms to learn patterns and relationships within the data [24]. This invaluable capability has opened new avenues for scientific inquiry and discovery across various domains, including biology, physics, chemistry, and environmental science, enabling researchers to make data-driven predictions, classify complex phenomena, and gain deeper insights into intricate processes that would be challenging to decipher using traditional analytical methods alone [25–28].

In our study, we developed a unique hybrid ML model that integrates deep learning Artificial Neural Network (ANN) with traditional ML methods, namely Random Forest, and AdaBoost. This model aims to correlate plasma polymerization conditions with the chemistries of the film's surface. ANNs are excellent for modeling complex, nonlinear relationships, which is essential for capturing the intricate dependencies between experimental conditions and surface chemistries. Random Forests offer robust and stable predictions with insights into feature importance. Feature importance quantifies the impact of each feature on the predictions made by a machine learning model. AdaBoost's ability to iteratively improve weak learners' performance ensures high accuracy and precision, particularly in challenging datasets with noise and variability. Weak learners such as decision stumps are simple models that perform only marginally better than random guessing, often used in ensemble methods to build a stronger predictive model. By choosing a hybrid model that combines ANN, Random Forest, and AdaBoost, we leverage the unique strengths of each method to create a robust and accurate model for mapping the relationship between experimental conditions and deposited surface chemistries. Additionally, a genetic algorithm (GA) fine-tunes the relative weights of the ANN, Random Forest, and AdaBoost components. Employing this refined hybrid model, we examined plasma polymerization of (2,2,6,6-Tetramethylpiperidin-1-yl)oxyl (TEMPO) via APPJ. TEMPO encompasses a stable nitroxide radical, exhibiting characteristics akin to those of NO, yet devoid of the latter's notably transient existence. This exploration highlighted the impact and significance of various experimental parameters, such as applied voltage, frequency, TEMPO-Helium flow ratio, and nozzle-sample distance, on the film's surface chemistries. Additionally, the model facilitated evaluations of film surface chemistries under new conditions.

2 | Methodology

2.1 | Experimental Schematic

The plasma jet as shown in Figure 1 is a cylindrical dielectric tube of 10 cm in length with an inner and outer diameter of respectively 2 and 3 mm. A copper electrode was located 10 mm from the jet nozzle and powered by high-voltage AC. The power supply consisted of a signal generator (TG 2000), a power amplifier (STA-800), and a homemade transformer. The signals applied to the electrode were monitored using a digital oscilloscope (Tektronix, DPO 4034B) and measured by a Pintek high-voltage probe (HVP-15 HF).

TEMPO (98%, CAS 2564-83-2) was purchased from Fisher Scientific and used without any further purification. The TEMPO is in a solid state at room temperature, so an extra constant heating platform was used to turn the TEMPO into a liquid state before being introduced to the jet. The aluminum samples (1×1 cm, Purity 99.5%, Advent Research Materials Ltd) were placed into a beaker containing isopropanol and cleaned in an ultrasonic bath for 3 min. The samples were placed under the jet nozzle and the deposition time was 20 min. The deposited films were analyzed using X-ray photoelectron spectroscopy (XPS) (Kratos Analytical AXIS Supra). The analysis used a monochromatic source A1 Kα 1486.7 eV, operating at 15 kV, 15 mA, and equipped with an electron flood gun for charge neutralization. Binding energies were calculated regarding the C1s peak at 285 eV.

Under the prescribed operational constraints of the transformer, the initiation of the TEMPO-Helium mixture was efficiently achieved by employing voltages ranging from 7 to 9 kV and frequencies spanning from 5 to 50 kHz for the jet system. The TEMPO-Helium flow ratio, defined as the ratio of the TEMPO-Helium mixture flowing through the side channel to the Helium flowing through the main channel, were adjusted to be within the range of 1%–3% with the flow rate of the main channel set at two standard liters per minute. Alternatively, for monomers with known Antoine coefficients, the monomer concentration can be calculated using the Antoine equation (ESI*). Concurrently, the distances between the nozzle and the sample were adjusted, ranging from 4 to 6 cm. A total of 59 samples were deposited under the specific experimental parameters (ESI*).

2.2 | Description of the Hybrid Model

Figure 2 illustrates the logical framework for prediction, evaluation, and optimization via the hybrid ML model. The integrated development environment used for this study is Pycharm. Four pivotal experimental parameters, including the applied voltage and frequency for discharge, the TEMPO-Helium flow ratio, and the distance between the nozzle and the sample, were chosen as the input features. MinMaxScaler from

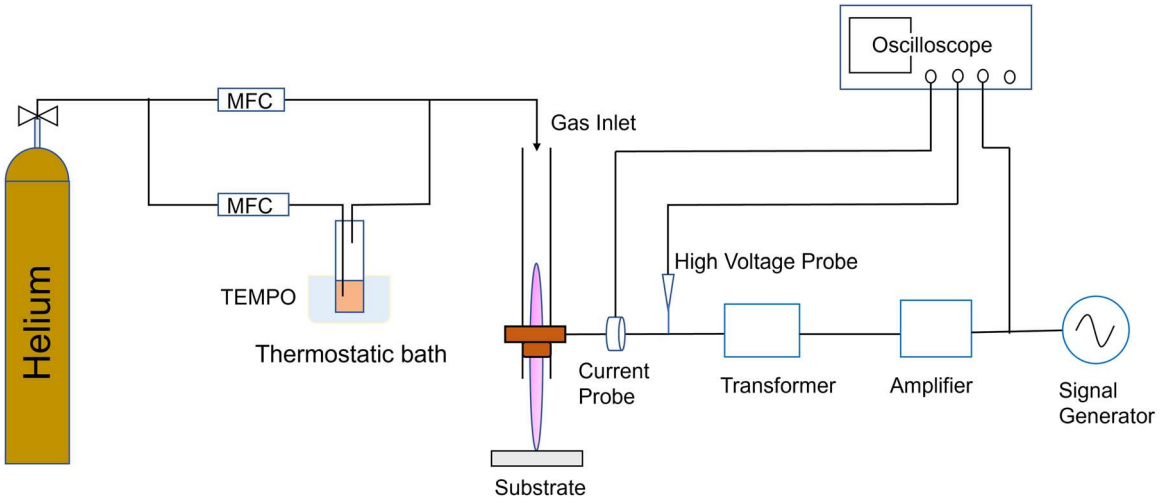

FIGURE 1 | A diagram of the plasma jet system.

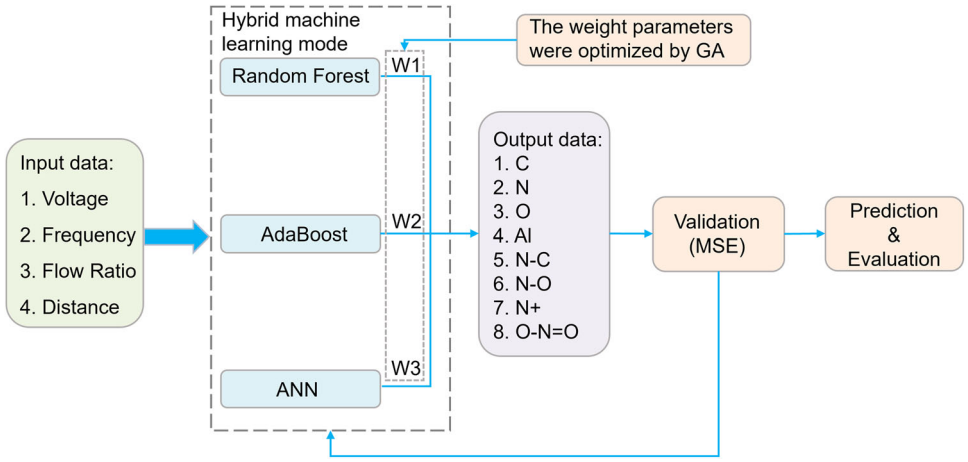

FIGURE 2 | Scheme of the logical structure of the prediction and evaluation using the hybrid ML model.

Scikit-learn was used to ensure that all features from training data and test data have the same scale. Selected surface chemistries, encompassing the atomic percentage of the surface and the chemical composition of Nitrogen (N), act as the labeled output "y."

The surface chemistries of the atmospheric-pressure plasma polymerized films were simulated using a hybrid ML model via a linear combination of ANN, AdaBoost (AB), and Random Forest (RF) algorithms, as shown in Equation (1).

$$P = W_1 \times P_{ANN} + W_2 \times P_{AB} + W_3 \times P_{RF} \quad (1)$$

where W1, W2, and W3 are the relative weights of ANN, AB, and RF algorithms, respectively. All three algorithms are independently trained using the same training data set, comprising 80% of the items randomly chosen from our entire data set. The optimal models derived from these training sessions are preserved as P_{CNN}, P_{AB} and P_{RF}, as illustrated in Equation (1). To determine the optimal weights (W1, W2, and W3) of the hybrid model, a GA is employed [29, 30]. The GA concluded its run after 50 iterations. The GA algorithm was developed using the "Sko" library in Python 3.11.

The mean squared error (MSE) is defined as the mean difference between the experimental ground truth (y) and the predicted results ($\hat{y}$) [31].

$$MSE = \frac{1}{n}\sum_{1}^{n}(y - \hat{y})^2 \quad (2)$$

MSE is chosen in the training algorithms because it leads to simple gradient calculations and optimization landscapes, which are well understood and efficiently solvable using gradient descent and its variants. In this study, the MSE also serves as the metric to assess the performance of the hybrid ML model.

2.2.1 | Artificial Neural Network

In the ANN model, we constructed a densely connected neural network with three hidden layers to map input features to surface chemistries. Figure 3 depicts the architecture of the ANN, which encompasses an input layer, an output layer, and three hidden layers. These hidden layers contain 32, 64, and 64 nodes respectively. The mathematical representation for the output of each hidden layer and output layer is as follows [32]:

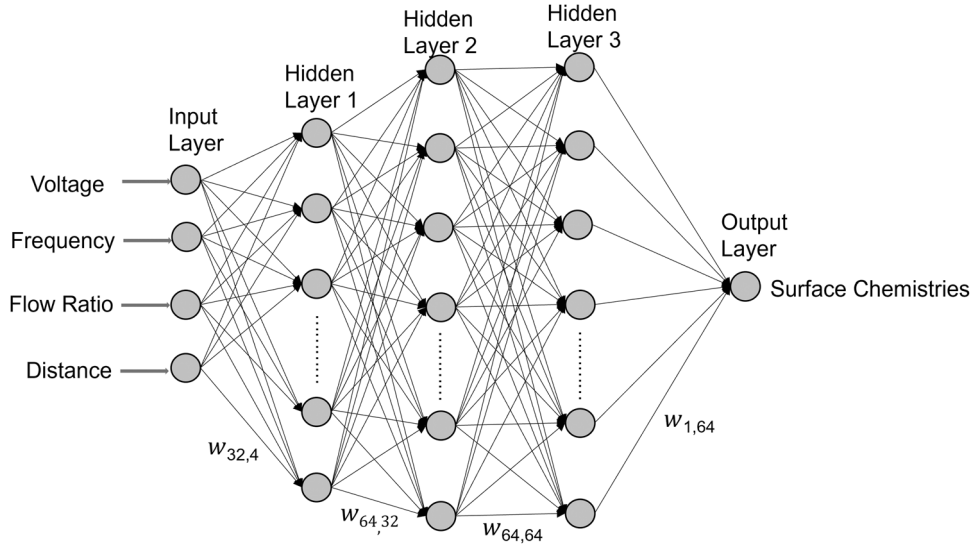

FIGURE 3 | The structure of the artificial neural network.

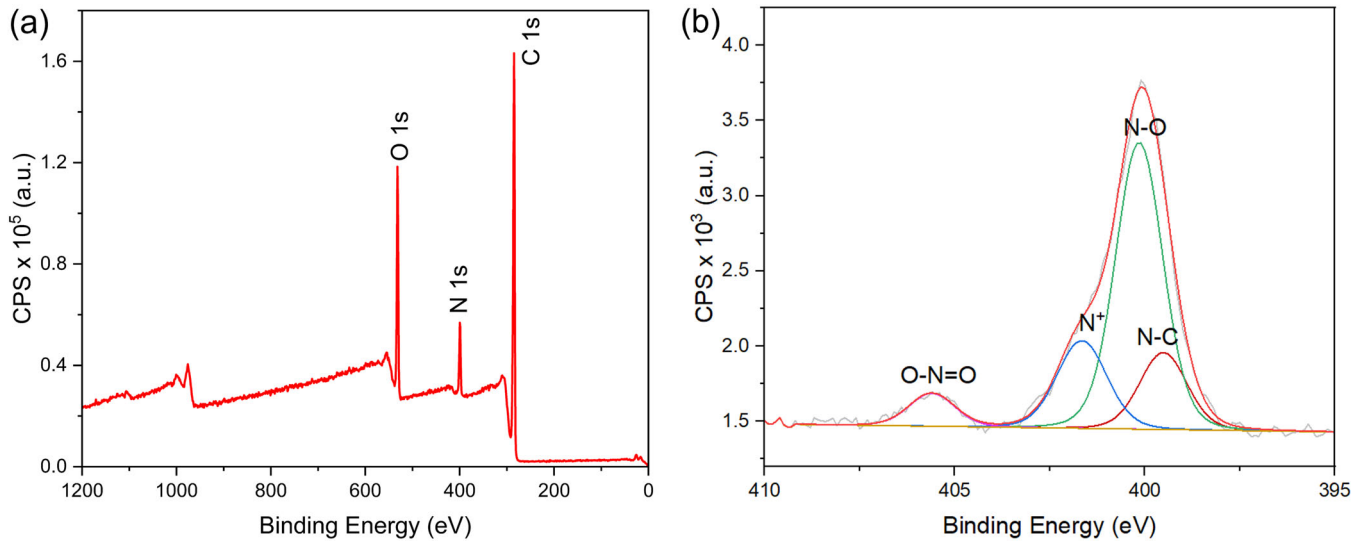

FIGURE 4 | (a) XPS spectrum and (b) high-resolution fittings for N 1 s peak of the sample deposited at 8 kV, 5 kHz, TEMPO-Helium ratio of 2, and nozzle-sample distance of 6 mm.

$$Z'(k) = f(Z(k)) = f\left(\sum_{i=1}^{N}(w_{ij}(k)\cdot x_j(k) + b_i)\right) \quad (3)$$

here, k means the hidden layer number (ranging from 1 to 4, presenting the first, second, and third hidden layer and output layer respectively), and N means there are N nodes in this layer. w_{ij} represents the weight between the i-th neuron in the current layer and the j-th neuron in the previous layer or the j-th input feature if the weight is for the first hidden layer. x_j is the output of the j-th neuron in the previous layer or the j-th input feature if the current layer is the first hidden layer. b_i denotes the bias for the i-th neuron in this layer. Z(k) presents the weighted sums of this layer's input; Z'(k) presents the output of this layer after being activated by the activation function. f represents the activation function of the hidden layer neuron. This introduces nonlinearity into the model and allows the network to learn complex patterns, which is defined as:

$$ReLu(x) = \max(0, x) \quad (4)$$

The utilization of ReLU as the activation function is primarily due to its ability to accelerate the training process and mitigate the vanishing gradient issue, thereby facilitating quicker and more efficient learning of intricate data patterns. An L2 regularization term with a coefficient of 0.03 is applied to all hidden layers to mitigate overfitting. Equation (5) defines the cost function $J(w)$ as the MSE between the predicted outcomes $Z'(4)_i$ and the actual experimental ground truth y_i [24]. To attain the optimized weights for each process, this cost function can be minimized using a gradient algorithm. For the training of neural networks in this context, we've selected a commonly

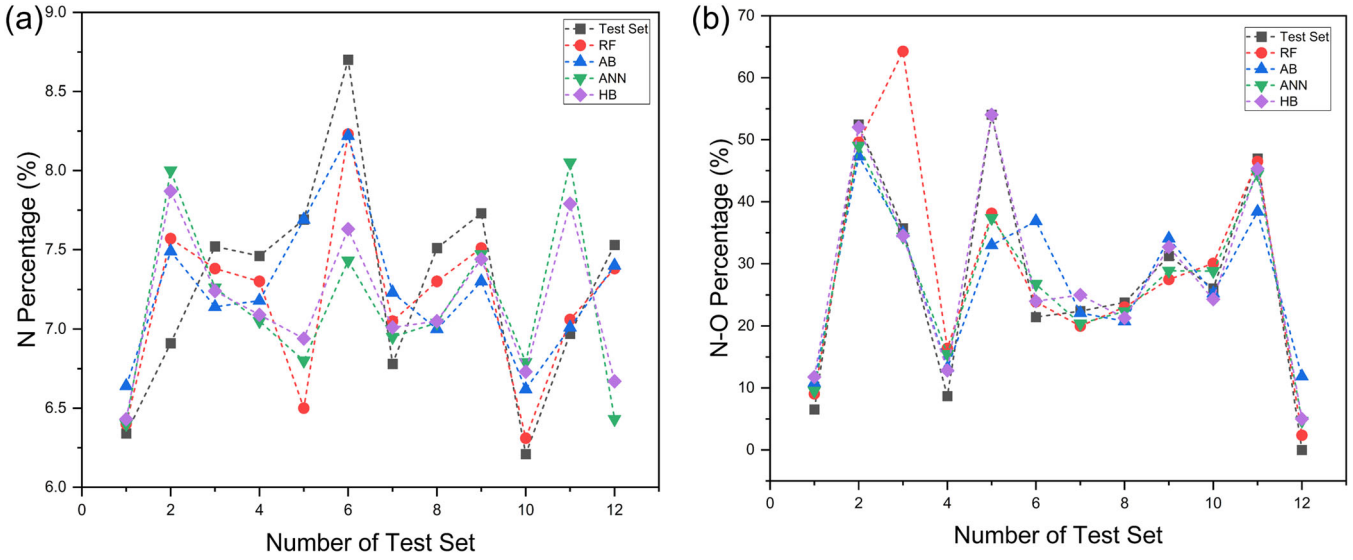

FIGURE 5 | Validation of RF, AB, ANN, and Hybrid (HB) algorithms at test data when (a) N and (b) N-O, as the target y respectively.

used optimizer, Adam. We trained the model for 1000 epochs and saved the best-performing model, characterized by the minimum cost, for subsequent use. The hyperparameters were tuned to ensure minimization of MSE on the test data set. The ANN algorithm was developed using the "Scikit-learn" and "Keras" library in Python 3.11.

$$J(w) = \frac{1}{n}\sum_{1}^{n}(y_i - Z'(4)_i)^2 \quad (5)$$

2.2.2 | Random Forest

The Random Forest algorithm is an ensemble learning method that amalgamates several weak learners to boost the overall model's performance [33]. It merges the concept of bagging with decision trees serving as the foundational models. These decision trees initiate from a root node, traverse through various decision nodes, and conclude with a prediction at the leaf nodes. Bagging involves generating multiple subsets of training samples with replacements, ensuring each observation has an equal likelihood of being selected [24]. Each of these training subsets is then utilized to train an individual decision tree, and the final prediction is derived from averaging the outcomes of all these trees. The feature importance in a Random Forest model is calculated based on the decrease in node Gini impurity (ESI*) that each feature provides across all decision trees in the forest. The Random Forest algorithm was developed using the "Scikit-learn" library in Python 3.11.

2.2.3 | AdaBoost

Boosting is an ensemble technique aimed at transforming weak learners into more potent ones, where "weak" and "strong" pertain to the accuracy of these learners in predicting the target variables [34]. In boosting, every training sample is employed to train an individual decision tree, and data is chosen with replacement, prioritizing over-weighted data. Successive trees learn from their predecessors by adjusting to the residual error. Once these weak learners have been trained, their estimates are aggregated using a weighted average to arrive at the final prediction. Feature importance in the AdaBoost model is calculated based on how often a feature is used across all weak learners and how much it contributes to reducing the error in predicting the target variable. The AdaBoost algorithm was developed using the "Scikit-learn" library in Python 3.11.

3 | Results and Discussions

3.1 | XPS Characterization of the Film

Figure 4a shows the XPS spectrum for N 1 s peak deposited at 8 kV, 5 kHz, TEMPO-Helium ratio of 2, and nozzle-sample distance of 6 mm. The film surface consists of 76.3% C, 16.2% O, and 7.5% N. Figure 4b presents the high-resolution fittings for the N 1 s peak of TEMPO film, where the spectrum is deconvoluted into four distinct components [35]. These components are identified as N-C at an energy level of 399.3 eV, N-O at 400 eV, N^+ at 401.6 eV, and O-N = O at 405.6 eV. The O-N = O groups are formed by the symmetrical quenching of two tempo radicals in the plasma state [36].

3.2 | Model Validation and Prediction

To analyze the experimental data, machine learning models were employed, highlighting the impact and significance of various experimental parameters on the film's surface chemistries. In the course of this study, 80% of the data set was allocated for training three distinct ML models, alongside the application of a GA to ascertain the coefficients W1, W2, and W3. The remaining 20% of the data served as the dedicated test set. As illustrated in Figure 5, the validation process was conducted for RF, AB, ANN, and the Hybrid algorithms (HB) on the test set.

When the target y was represented by the variable N, the mean squared errors (MSE) for RF, AB, and ANN were found to be 0.41, 0.19, and 0.52, respectively. Conversely, when the target y

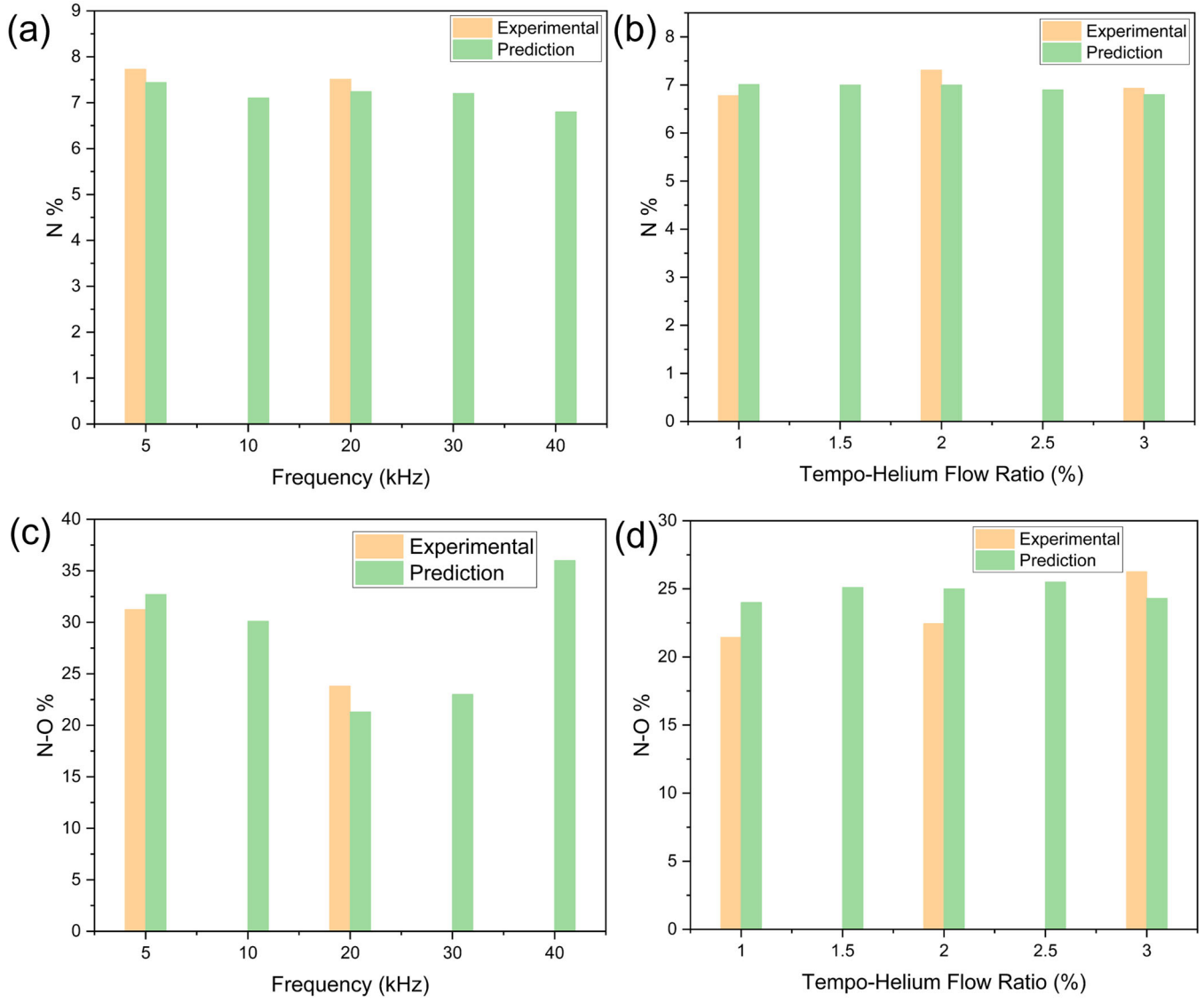

FIGURE 6 | (a) Influence of frequency on both experimental ground truth and model-estimated N% while holding voltage at 8 kV, TEMPO-Helium flow ratio at 2, and nozzle-sample distance at 6 mm. (b) Impact of the TEMPO-Helium flow ratio on experimental ground truth and model-predicted N% with settings fixed at 9 kV for voltage, 20 kHz for frequency, and 6 mm for nozzle-sample distance. (c) Influence of frequency on the observed and model-estimated N-O% while maintaining the voltage at 8 kV, TEMPO-Helium flow ratio at 2, and nozzle-sample distance at 6 mm. (d) Impact of the TEMPO-Helium flow ratio on observed and model-predicted N-O% under conditions of 9 kV voltage, 20 kHz frequency, and 6 mm nozzle-sample distance.

was denoted by N-O, the corresponding MSE values for RF, AB, and ANN amounted to 56.43, 45.79, and 60.67, respectively. The percentage of N-O is generally high, even small relative errors can lead to large absolute squared errors, thus inflating the MSE.

Optimal weights for RF, AB, and ANN were determined to be 0.19, 0.31, and 0.49, respectively, in the case of N as the target y. Conversely, for N-O as the target y, the optimal weights were observed to be 0.52, 0.48, and 0.13 for RF, AB, and ANN, respectively. Furthermore, the MSE of the hybrid algorithm was calculated to be 0.14 when N was considered as the target y, and 30.14 when N-O was utilized as the target y. It is noteworthy that while the MSE for individual algorithms exhibited variation, the hybrid algorithms consistently demonstrated the lowest MSE values, underscoring the robustness of the hybrid approach.

Figure 6a,b provide a comparative analysis between the observed and simulated surface N percentages as frequency and TEMPO-Helium flow ratio increase, respectively. When operating the jet at 9 kV and 20 kHz, with a nozzle-sample distance of 6 mm, the surface N percentage remains consistent despite variations in the TEMPO-Helium flow ratio. This suggests that the TEMPO-Helium ratio flow doesn't significantly influence the surface N % under these conditions. Likewise, when operating the jet at 8 kV, using a TEMPO-Helium flow ratio of 2, and maintaining a nozzle-sample distance of 6 mm, the percentage of N on the films exhibits only minor fluctuations with increasing frequency.

Figure 6c,d depict the relationship between observed and simulated N-O percentages in response to changes in frequency and TEMPO-Helium flow ratio, respectively. As observed in Figure 6c, the N-O percentage declines initially with increasing frequency, hitting its lowest point at 20 kHz. Beyond this, the percentage begins to climb gradually as the frequency continues to rise. On the other hand, as the TEMPO-Helium flow ratio goes up, the N-O percentage exhibits minimal fluctuation, suggesting that the N-O production remains consistent. This indicates that the TEMPO-Helium flow ratio adjustments exert a limited influence on the N-O percentage within this area of the input space (under 9 kV, 20 kHz, nozzle-sample distance of 6 mm).

3.3 | Feature Importance

When depositing films using atmospheric pressure plasma with the TEMPO monomer, the inherent chemical elements of the tempo monomer—carbon, nitrogen, and oxygen—are retained. Under certain experimental conditions, traces of aluminum can also be observed, particularly when the film's thickness is minimal. The primary aim of plasma polymerization using the TEMPO monomer is to ensure a high retention of the N-O functional group on the deposited surfaces. This study primarily focused on high-resolution fitting of nitrogen (N) components, including N-C, N-O, N^+, and O-N=O. Figure 7 delineates the influence of varying operational parameters on the film's surface chemistries, which is obtained from averaging the values obtained by RF and AB. Notably, frequency and TEMPO-Helium flow ratio emerge as the dominant factors, accounting for over 80% of the total relative importance when considering the film's atomic percentage. However, when examining specific functional groups, such as N-C and N-O, the significance of voltage and distance escalates, nearly matching the importance of the aforementioned factors. Thus, to achieve maximum retention of the N-O functional group on the surfaces, meticulous calibration of all experimental parameters is essential.

Here, the feature's importance is identified globally, but its influence on the target y may be insignificant in certain local areas or specific regions of the input space. This nonuniform influence across different regions is likely due to the model's nonlinearity [37].

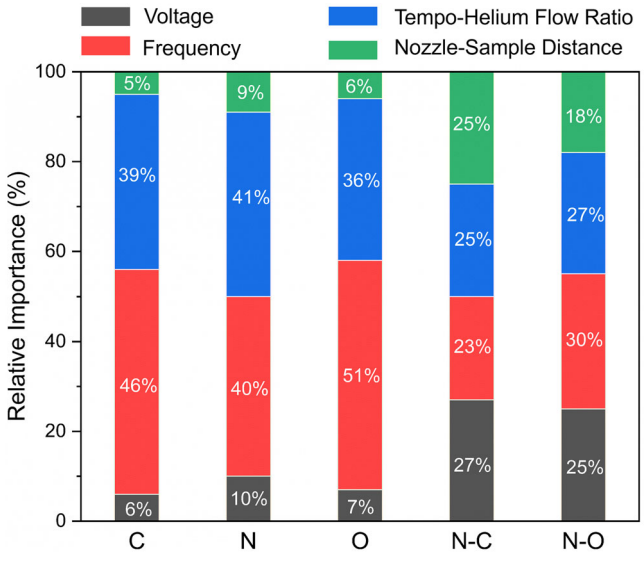

FIGURE 7 | Feature Importance for surface atomic percentage, N-C, and N-O functional groups.

3.4 | Coupling Effect of Input Features

The hybrid ML model has been employed to explore the interplay of plasma deposition parameters on film surface chemistries. Figure 8 visualizes the interaction of frequency and TEMPO-Helium flow ratio on both surface atomic percentages and specific functional groups. Despite using different target variables (y), the sum of the predicted atomic percentages for different elements at the same condition is very close to 100%. This indicates that the model effectively distributes the total atomic composition among the different elements, maintaining a realistic and balanced prediction.

As shown in Figure 8a, at relatively low TEMPO-Helium flow ratios (below 2), the C percentage decreases as the frequency increases. Conversely, at higher TEMPO-Helium flow ratios, the C percentage remains at high levels. Similarly, Figure 8c illustrates that at low TEMPO-Helium flow ratios, the O percentage increases with increasing frequency, whereas at higher flow ratios, the O percentage remains low. For sufficiently thick TEMPO films, the theoretical elemental composition consists of 81.2% C, 9.1% N, and 9.1% O (ESI*) [38]. At low TEMPO-Helium flow ratios, where the introduction of TEMPO monomer is limited, the observed increase in O percentage may result from the fixation of oxygen from the ambient air. As evidenced by the mass spectrum in ESI*, a high concentration of oxygen ions is present in the plasma plume at low flow ratios. In contrast, at higher flow ratios (greater than 2), a high introduction of monomer leads to the elemental composition in the resulting film that more closely aligns with the theoretical values of sufficiently thick TEMPO films. Figure 8b indicates that the flow ratio shows a limited influence on the N percentage at high frequencies, while the N percentage declines as the TEMPO-Helium ratio increases at 5 kHz. In Figure 8d, an increasing frequency correlates with a growing N-C percentage, with the TEMPO-Helium flow ratio showing minimal variance in effect. As illustrated in Figure 8e, the N–O percentage initially decreases with increasing frequency from 5 kHz to 20 kHz, remains at low levels between 20 kHz and 30 kHz, and subsequently increases from 30 kHz to 40 kHz. The frequency significantly influences the ionization process and memory charge within the plasma, leading to the operation of the jet in different modes at varying frequencies [39, 40]. At 20 kHz, the jet operates in the “bullets” mode, as confirmed by the discharge current waveform (ESI*). The observed decrease in the N–O percentage may be attributed to the etching effect induced by the stable and repetitive plasma “bullets” [41]. Lastly, Figure 8f demonstrates that the N+ percentage remains relatively low across all conditions. However, a distinct peak in N+ percentage is observed at 40 kHz with the TEMPO-Helium ratio set to 1.

Figure 9 delves into the interplay between frequency and TEMPO-Helium flow ratio concerning surface atomic percentages and various functional groups with frequency set to 5 kHz and nozzle-sample distance set to 6 mm. As shown in Figure 9a,

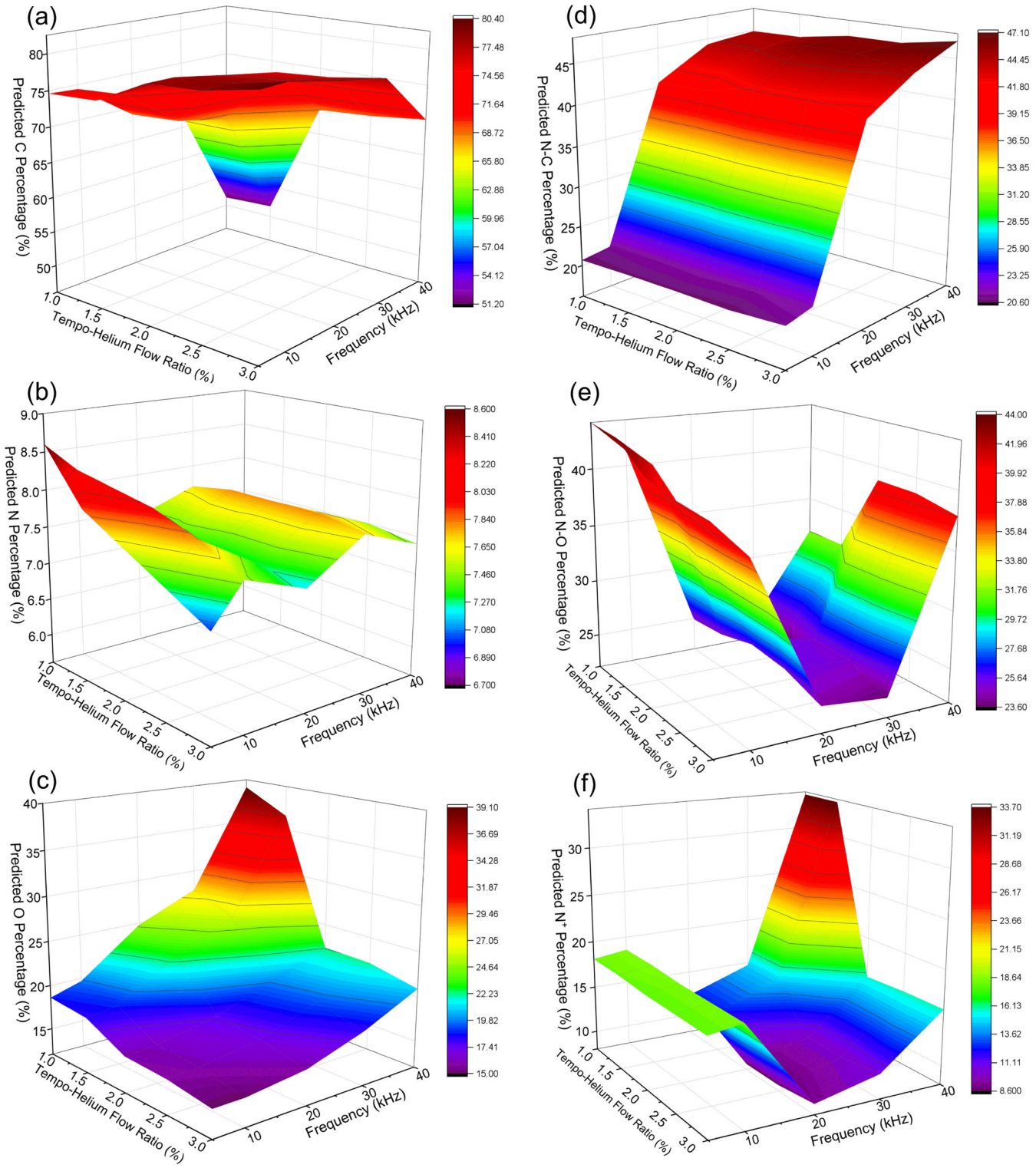

FIGURE 8 | Predictions for (a) C%, (b) N%, (c) O%, (d) N-C%, (e) N-O%, and (f) N + %, based on variations in frequency and TEMPO-Helium flow ratio, while maintaining the voltage and nozzle-sample distance at 9 kV and 6 mm, respectively.

the C percentage increases with the TEMPO-Helium flow ratio, while the applied voltage has a limited influence. Similarly, Figure 9c illustrates that the O percentage decreases as the flow ratio increases, with voltage exerting minimal impact. The high O percentage observed at low flow ratios may also be attributed to the fixation of oxygen from ambient air due to the limited introduction of TEMPO monomer. In Figure 9b, a peak N percentage is observed at 9 kV with the TEMPO-Helium flow ratio set to 1. At 8 kV and 5 kHz, the jet operates in a chaotic mode [40], characterized by the discharge current (ESI*). This mode is marked by weak gas discharge and an inadequate supply of seed electrons, resulting in the inability to maintain a dynamic balance between the applied voltage and the memory voltage essential for stable and periodic plasma. Weak

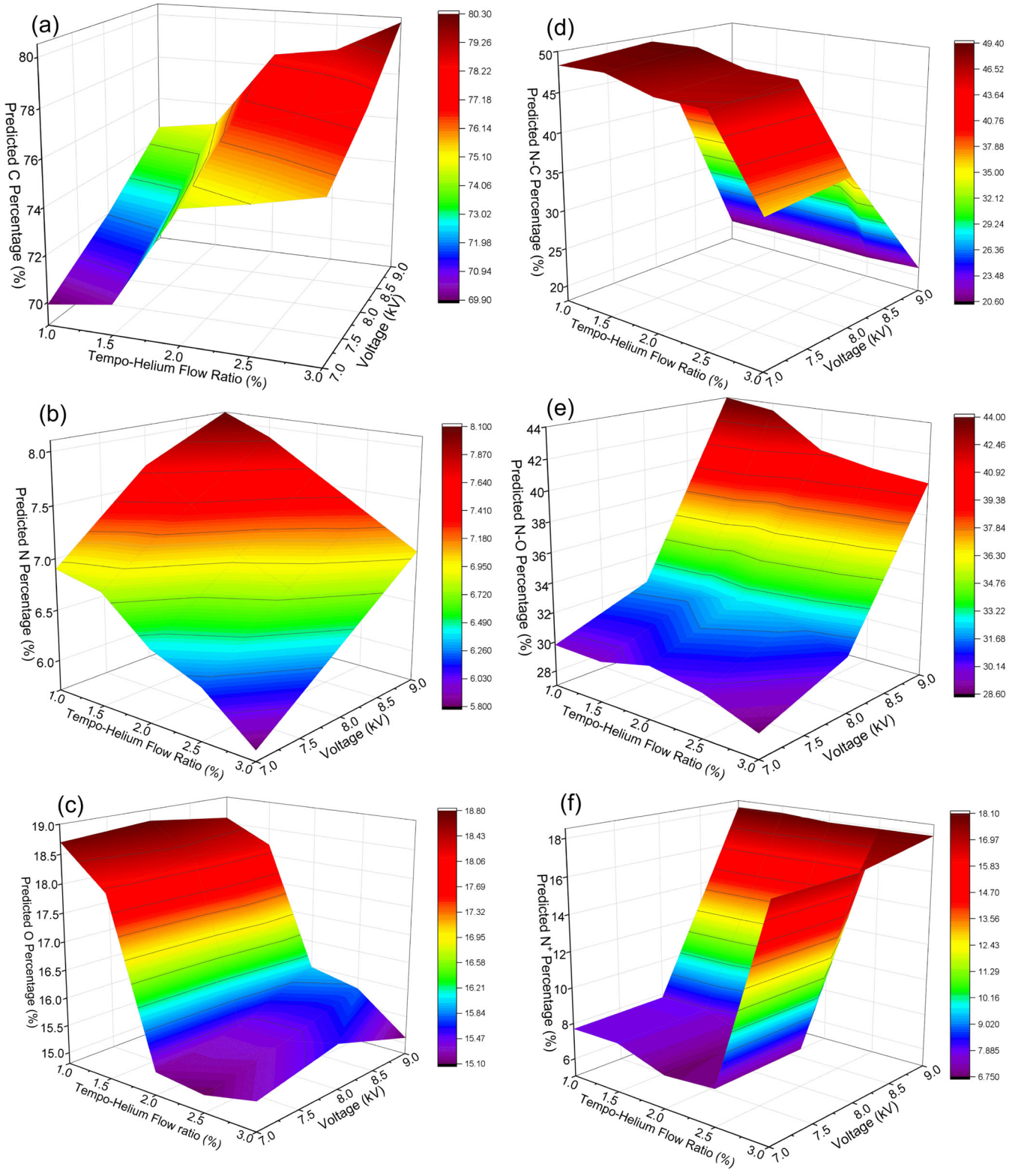

FIGURE 9 | Predicted variations for (a) C%, (b) N%, (c) O%, (d) N-C%, (e) N-O%, and (f) N + %, influenced by changes in voltage and TEMPO-Helium flow ratio, while holding the frequency and nozzle-sample distance constant at 5 kHz and 6 mm, respectively.

discharge at 5 kHz, combined with the fixation of oxygen from the air, possibly results in the above atomic percentage distribution at different voltages and TEMPO-Helium flow ratios. As shown in Figure 9d, the N–C percentage remains high between 7 kV and 8 kV before decreasing, with the TEMPO-Helium flow ratio exerting a limited influence. In contrast, the N–O percentage remains low from 7 kV to 8 kV but increases from 8 kV to 9 kV, with the flow ratio having a marginal impact. Furthermore, Figure 9f demonstrates that the N^+ percentage remains low when the flow ratio is below

2.5 and the voltage is below 8 kV but increases as both the flow ratio and voltage continue to rise.

In summary, the hybrid mode in this study elucidates the relationship between experimental conditions and the surface chemistries of the films. The TEMPO-Helium flow ratio has a significant influence on the C and O percentages on the film surface. At 20 kHz, the jet operates in the "bullets" mode, which is not conducive to the retention of N–O groups on the film surface. Additionally, it allows us to predict surface chemistries under new conditions. However, for the prediction to be effective, the conditions should share the same distribution as the training data set.

4 | Conclusions

A hybrid ML model, melding three distinct algorithms, has been crafted to forecast and assess the impact of various plasma deposition parameters on the surface chemistries of films. Utilizing the Genetic Algorithm (GA), optimal weights for each algorithm were determined. The MSE of the hybrid algorithm was calculated to be 0.14 when N was considered as the target y, and 30.14 when N-O was utilized as the target y. While frequency and TEMPO-Helium flow ratio stand out as pivotal parameters influencing surface atomic percentages, the significance of voltage and nozzle-sample distance becomes pronounced for certain functional groups. A sufficiently thick TEMPO film typically requires a TEMPO-Helium flow ratio greater than 2. Furthermore, higher N-O percentages correlate with lower frequencies and higher voltages. This study underscores the capability of a robust hybrid ML model to predict and evaluate the plasma polymerization process, hinting at its prospective utility across various plasma-driven chemical procedures efficiently and accurately.

Author Contributions

Yong Wang and James W. Bradley contributed to the study design, while the algorithms were developed by Yong Wang and Xudong Ma. XPS analyses were performed by Alexander J. Robson. The manuscript was written by Yong Wang and James W. Bradley, with input from all authors.

Acknowledgments

The authors thank the Engineering and Physical Sciences Research Council (EPSRC) for supporting this study with grants EP/S005153/1 and EP/S004505/1. The authors are also thankful to the Chinese Scholarship Council for their support. There is no conflict of interest.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. T. Yang, A. N. Zelikin, and R. Chandrawati, "Progress and Promise of Nitric Oxide-Releasing Platforms," *Advanced Science* 5, no. 6 (2018): 1701043.

2. A. B. Seabra, G. Z. Justo, and P. S. Haddad, "State of the Art, Challenges and Perspectives in the Design of Nitric Oxide-Releasing Polymeric Nanomaterials for Biomedical Applications," *Biotechnology Advances* 33, no. 6 (2015): 1370–1379.

3. M. Wu, Z. Lu, K. Wu, C. Nam, L. Zhang, and J. Guo, "Recent Advances in the Development of Nitric Oxide-Releasing Biomaterials and Their Application Potentials in Chronic Wound Healing," *Journal of Materials Chemistry B* 9, no. 35 (2021): 7063–7075.

4. G. Jin, Z. Gao, Y. Liu, et al., "Polymeric Nitric Oxide Delivery Nanoplatforms for Treating Cancer, Cardiovascular Diseases, and Infection," *Advanced Healthcare Materials* 10, no. 3 (2021): 2001550.

5. Y. Wang, A. J. Robson, S. Simon, R. D. Short, and J. W. Bradley, "Plasma Polymerization of Isopentyl Nitrite at Atmospheric Pressure: Gas Phase Analysis and Surface Chemistry," *Plasma Processes and Polymers* 21 (2024): e2300162.

6. A. L. Urzedo, M. C. Gonçalves, M. Nascimento, C. B. Lombello, G. Nakazato, and A. B. Seabra, "Cytotoxicity and Antibacterial Activity of Alginate Hydrogel Containing Nitric Oxide Donor and Silver Nanoparticles for Topical Applications," *ACS Biomaterials Science & Engineering* 6, no. 4 (2020): 2117–2134.

7. D. Amilan Jose, N. Sharma, R. Sakla, R. Kaushik, and S. Gadiyaram, "Fluorescent Nanoprobes for the Sensing of Gasotransmitters Hydrogen Sulfide (H2S), Nitric Oxide (NO) and Carbon Monoxide (CO)," *Methods* 168 (2019): 62–75.

8. J. O. Kim, J. K. Noh, R. K. Thapa, et al., "Nitric Oxide-Releasing Chitosan Film for Enhanced Antibacterial and In Vivo Wound-Healing Efficacy," *International Journal of Biological Macromolecules* 79 (2015): 217–225.

9. R. Ahmed, R. Augustine, M. Chaudhry, et al., "Nitric Oxide-Releasing Biomaterials for Promoting Wound Healing in Impaired Diabetic Wounds: State of the Art and Recent Trends," *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie* 149 (2022): 112707.

10. K. G. Doherty, J. S. Oh, P. Unsworth, et al., "Plasma Polymerization Using Helium Atmospheric-Pressure Plasma Jet With Heptylamine Monomer," *Plasma Processes and Polymers* 16, no. 4 (2019): e1800185.

11. M. Mokter Hossain, Q. H. Trinh, M. S. P. Sudhakaran, L. Sultana, and Y. S. Mok, "Improvement of Mechanical Strength of Hydrophobic Coating on Glass Surfaces by an Atmospheric Pressure Plasma Jet," *Surface and Coatings Technology* 357 (2019): 12–22.

12. A. Demaude, C. Poleunis, E. Goormaghtigh, et al., "Atmospheric Pressure Plasma Deposition of Hydrophilic/Phobic Patterns and Thin Film Laminates on Any Surface," *Langmuir* 35, no. 30 (2019): 9677–9683.

13. C. L. Ko, Y. L. Kuo, W. J. Lee, H. J. Sheng, and J. Y. Guo, "The Enhanced Abrasion Resistance of an Anti-Fingerprint Coating on Chrome-Plated Brass Substrate by Integrating Sputtering and Atmospheric Pressure Plasma Jet Technologies," *Applied Surface Science* 448 (2018): 88–94.

14. A. Anagri, A. Baitukha, C. Debiemme-Chouvy, et al., "Nanocomposite Coatings Based on Graphene and Siloxane Polymers Deposited by Atmospheric Pressure Plasma. Application to Corrosion Protection of Steel," *Surface and Coatings Technology* 377 (2019): 124928.

15. J. H. Tsai, I. C. Cheng, C. C. Hsu, and J. Z. Chen, "DC-Pulse Atmospheric-pressure Plasma Jet and Dielectric Barrier Discharge Surface Treatments on Fluorine-Doped Tin Oxide for Perovskite Solar Cell Application," *Journal of Physics D: Applied Physics* 51, no. 2 (2017): 025502.

16. T. Wang, X. Wang, B. Yang, X. Chen, C. Yang, and J. Liu, "Low Temperature Atmospheric Microplasma Jet Array for Uniform Treatment of Polymer Surface for Flexible Electronics," *Journal of Micromechanics and Microengineering* 27, no. 7 (2017): 075005.

17. H. Zhang, Z. Guo, Q. Chen, X. Wang, Z. Wang, and Z. Liu, "Deposition of Silicon Oxide Coatings by Atmospheric Pressure Plasma Jet for Oxygen Diffusion Barrier Applications," *Thin Solid Films* 615 (2016): 63–68.

18. M. Favaro, E. Zanazzi, A. Patelli, et al., "Aluminum Doped Zinc Oxide Coatings at Low Temperature by Atmospheric Pressure Plasma Jet," *Thin Solid Films* 708 (2020): 138118.

19. J. Ibrahim, S. A. Al-Bataineh, A. Michelmore, and J. D. Whittle, "Atmospheric Pressure Dielectric Barrier Discharges for the Deposition of Organic Plasma Polymer Coatings for Biomedical Application," *Plasma Chemistry and Plasma Processing* 41 (2021): 47–83.

20. C. Zhang, J. Wang, and X. Geng, "Tungsten Oxide Coatings Deposited by Plasma Spray Using Powder and Solution Precursor for Detection of Nitrogen Dioxide Gas," *Journal of Alloys and Compounds* 668 (2016): 128–136.

21. J. Z. Chen, C. C. Hsu, C. Wang, et al., "Rapid Atmospheric-Pressure-Plasma-Jet Processed Porous Materials for Energy Harvesting and Storage Devices," *Coatings* 5, no. 1 (2015): 26–38.

22. J. Wei, X. Chu, X. Y. Sun, et al., "MachineLearning in Materials Science," *InfoMat* 1, no. 3 (2019): 338–358.

23. R. Roscher, B. Bohn, M. F. Duarte, and J. Garcke, "Explainable Machine Learning for Scientific Insights and Discoveries," *IEEE Access* 8 (2020): 42200–42216.

24. V. Nasteski, "An Overview of the Supervised Machine Learning Methods," *Horizons* 4 (2017): 51–62.

25. C. Xu and S. A. Jackson, "Machine Learning and Complex Biological Data," *Genome Biology* 20 (2019): 76.

26. L. M. Dery, B. Nachman, F. Rubbo, and A. Schwartzman, "Weakly Supervised Classification in High Energy Physics," *Journal of High Energy Physics* 2017, no. 5 (2017): 145.

27. M. Haghighatlari and J. Hachmann, "Advances of Machine Learning in Molecular Modeling and Simulation," *Current Opinion in Chemical Engineering* 23 (2019): 51–57.

28. D. J. Lary, A. H. Alavi, A. H. Gandomi, and A. L. Walker, "Machine Learning in Geosciences and Remote Sensing," *Geoscience Frontiers* 7, no. 1 (2016): 3–10.

29. A. Lambora, K. Gupta, and K. Chopra (2019). Genetic Algorithm—A Literature Review. In 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon) (pp. 380–384). IEEE.

30. Mirjalili, S. "Evolutionary Algorithms and Neural Networks," *Studies in Computational Intelligence* 780, no. 1 (2019): 43–53.

31. Khan, M., and S. Noor, "Performance Analysis of Regression-Machine Learning Algorithms for Predication of Runoff Time," *Agrotechnology* 8, no. 1 (2019): 1–12.

32. T. Yuan, Y. Xu, Y. Wang, P. Han, and J. Chen, "A Tilt Receiver Correction Method for Visible Light Positioning Using Machine Learning Method," *IEEE Photonics Journal* 10, no. 6 (2018): 1–12.

33. V. Rodriguez-Galiano, M. Sanchez-Castillo, M. Chica-Olmo, and M. Chica-Rivas, "Machine Learning Predictive Models for Mineral Prospectivity: An Evaluation of Neural Networks, Random Forest, Regression Trees and Support Vector Machines," *Ore Geology Reviews* 71 (2015): 804–818.

34. S. González, S. García, J. Del Ser, L. Rokach, and F. Herrera, "A Practical Tutorial on Bagging and Boosting Based Ensembles for Machine Learning: Algorithms, Software Tools, Performance Study, Practical Perspectives and Opportunities," *Information Fusion* 64 (2020): 205–237.

35. M. J. Barnes, A. J. Robson, J. Naderi, R. D. Short, and J. W. Bradley, "Plasma Polymerization of (2, 2, 6, 6-Tetramethylpiperidin-1-yl) Oxyl in a Collisional, Capacitively Coupled Radio Frequency Discharge," *Biointerphases* 15, no. 6 (2020): 061007.

36. T. D. Michl, D. T. T. Tran, K. Böttle, et al., "To be a Radical or Not to be One? The Fate of the Stable Nitroxide Radical TEMPO [(2, 2, 6, 6-Tetramethylpiperidin-1-yl) oxyl] Undergoing Plasma Polymerization Into Thin-Film Coatings," *Biointerphases* 15, no. 3 (2020): 031015.

37. N. B. Kumarakulasinghe, T. Blomberg, J. Liu, A. S. Leao, and P. Papapetrou, *Evaluating Local Interpretable Model-Agnostic Explanations on Clinical Machine Learning Classification Models. In 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS)* (IEEE, 2020), 7–12.

38. T. D. Michl, B. R. Coad, M. Doran, et al., "Nitric Oxide Releasing Plasma Polymer Coating With Bacteriostatic Properties and no Cytotoxic Side Effects," *Chemical Communications* 51, no. 32 (2015): 7058–7060.

39. B. L. Sands, S. K. Huang, J. W. Speltz, M. A. Niekamp, J. B. Schmidt, and B. N. Ganguly, "Dynamic Electric Potential Redistribution and its Influence on the Development of a Dielectric Barrier Plasma Jet," *Plasma Sources Science and Technology* 21, no. 3 (2012): 034009.

40. J. L. Walsh, F. Iza, N. B. Janson, V. J. Law, and M. G. Kong, "Three Distinct Modes in a Cold Atmospheric Pressure Plasma Jet," *Journal of Physics D: Applied Physics* 43, no. 7 (2010): 075201.

41. M. G. Kong, B. N. Ganguly, and R. F. Hicks, "Plasma Jets and Plasma Bullets," *Plasma Sources Science and Technology* 21, no. 3 (2012): 030201.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.