

Article



Impact of Pulse Parameters of a DC Power Generator on the Microstructural and Mechanical Properties of Sputtered AlN Film with In-Situ OES Data Analysis

Wei-Yu Zhou¹, Hsuan-Fan Chen¹, Xue-Li Tseng¹, Hsiao-Han Lo², Peter J. Wang², Ming-Yu Jiang², Yiin-Kuen Fuh^{1,*} and Tomi T. Li¹

- ¹ Department of Mechanical Engineering, National Central University, Taoyuan 320317, Taiwan
- ² Delta Electronics Incorporated, Taipei 100116, Taiwan
- * Correspondence: michaelfuh@gmail.com; Tel.: +886-3-426-7305; Fax: +886-3-425-4501

Abstract: In the present study, the sputtered aluminum nitride (AIN) films were processed in a reactive pulsed DC magnetron system. We applied a total of 15 different design of experiments (DOEs) on DC pulsed parameters (reverse voltage, pulse frequency, and duty cycle) with Box–Behnken experimental method and response surface method (RSM) to establish a mathematical model by experimental data for interpreting the relationship between independent and response variables. For the characterization of AlN films on the crystal quality, microstructure, thickness, and surface roughness, X-ray diffraction (XRD), atomic force microscopy (AFM), and field emission-scanning electron microscopy (FE-SEM) were utilized. AlN films have different microstructures and surface roughness under different pulse parameters. In addition, in-situ optical emission spectroscopy (OES) was employed to monitor the plasma in real-time, and its data were analyzed by principal component analysis (PCA) for dimensionality reduction and data preprocessing. Through the CatBoost modeling and analysis, we predicted results from XRD in full width at half maximum (FWHM) and SEM in grain size. This investigation identified the optimal pulse parameters for producing high-quality AlN films as a reverse voltage of 50 V, a pulse frequency of 250 kHz, and a duty cycle of 80.6061%. Additionally, a predictive CatBoost model for obtaining film FWHM and grain size was successfully trained.

Keywords: Reactive pulsed DC magnetron sputtering; AlN; BoxBehnken method; Pulse parameter; In-situ OES; PCA; RSM; CatBoost modelling

1. Introduction

AlN is a versatile material with high sound acoustic wave velocity, high thermal conductivity, low thermal-expansion coefficient, wide band gap, and low dielectric constant which are significant technical benefits for the semiconductor industry. Reactive sputtering offers a promising method in film deposition at low temperatures with a deposited film of homogenous composition, superior substrate adhesion, and lower contamination [1-4]. Due to the excellent piezoelectric performance of highly C-axis oriented AlN films, they are widely used in the above-mentioned applications, making the production of highly c-axis oriented AlN films with low surface roughness a major goal [5]. In their study, the crystal quality of the films was evaluated using the FWHM of the AlN (002) peak, and pulse parameters corresponding to the lowest FWHM value were determined through experiments. A study investigated the effect of process parameters on residual stress in AlN films, which is crucial for their mechanical stability. Residual stress correlates with film thickness and crystallite orientation. In-situ OES data was analyzed using PCA to predict the stress state, and RSM was applied to determine the optimum conditions for minimizing residual stress [6]. Another study focused on developing a cost-effective process to produce solar coatings for solar receiver tubes in linear focusing concentrated solar power (CSP) technology. The reactive magnetron sputtering technology in transition



Citation: Zhou, W.-Y.; Chen, H.-F.; Tseng, X.-L.; Lo, H.-H.; Wang, P.J.; Jiang, M.-Y.; Fuh, Y.-K.; Li, T.T. Impact of Pulse Parameters of a DC Power Generator on the Microstructural and Mechanical Properties of Sputtered AIN Film with In-Situ OES Data Analysis. *Materials* **2023**, *16*, 3015. https://doi.org/10.3390/ ma16083015

Academic Editor: Viktor Shapovalov

Received: 6 March 2023 Revised: 7 April 2023 Accepted: 8 April 2023 Published: 11 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). mode was applied to deposit all ceramic constituents of solar coatings, which significantly improved deposition rates. The developed process allowed for a 44% reduction in energy consumption, and the produced coating exhibited excellent thermal stability [7]. Their work optimized process parameters for high Young's modulus [002] oriented aluminum nitride (AlN) thin films by pulsed DC reactive sputtering, achieving Young's modulus up to 335 GPa and a surface roughness of 1.2 nm. The optimal pulse parameters were a frequency of 100 kHz and a duty cycle of 80% [1]. Our previous study has investigated the effect of N2 gas flow and power on the minimum film residual stress of AlN thin films deposited by pulsed DC reactive sputtering on Si (100). Large-scale in-situ OES, XRD, SEM, TEM, FTIR, and alpha-step profiler data were analyzed to determine the film stress states. They proposed a residual stress (VRS) value calculated by a standard deviation in the first principal component direction (PC1-DEV) method. The study suggests a methodology using PCA to determine residual stress characterization as an in-situ plasma monitoring tool for the AlN thin film deposition process [2].

Plasma properties are studied to better understand plasma vacuum processing and improve process control and repeatability. Typically, duty cycles between 50 and 90 percent and pulse frequencies between 10 and 350 kHz are used throughout the deposition process [8]. Therefore, understanding the basics of the creation and behavior of pulsed DC magnetron discharges is required to master the new benefits of pulsed DC sputtering. Furthermore, to establish the connections between film properties from film measurements to pulse parameters, it is essential to explain relationships between them and film properties such as film structure, composition, and characteristics, which are particularly interesting for practical applications.

Pulsed DC magnetron sputtering is a widely-used technology in the industry today, and it is especially effective for fabricating materials with excellent electrical insulation properties. It can obtain thin films with high density and good crystallinity at low temperatures [9–12]. In the current experiment, AlN films were processed by a reactive DC magnetron sputtering with an asymmetric bipolar-pulsed DC power source at various pulse settings in reverse voltage, pulse frequency, and duty cycle. The asymmetrical bipolar-pulsed DC power supply can adjust its pulse parameters according to requirements to improve efficiency. Compared with radiofrequency (RF) power supply, it can have a better deposition rate and prevents an electrically insulating layer from being deposited on the cathode surface by adding a reverse voltage, which can reduce charge buildup and arcing in affecting film quality [13]. A typical waveform and pulse feature was shown in Figure 1. This study used XRD, AFM, and FE-SEM to determine the microstructure of films. RSM, in conjunction with a mathematical Box–Behnken method [6,14], was employed to setup a DOE model for how the power supply DC pulse factors impact the microstructural properties of films.



Figure 1. A typical pulsed DC sputtering pattern in an asymmetric bipolar voltage waveform.

In addition, this study trained thin film quality prediction models through PCA dimensionality reduction combined with the CatBoost model. The CatBoost classifier

algorithm was adopted for training transmission line faults data in the multi-dataset from an electrical power system owing to its training speed, accuracy, and ability. The trained model implements an excellent accuracy of 99.54% [15]. The OES data have been widely built-in virtual metrology (VM) system. Multi-features in OES data are reduced dimensionality using principal component analysis (PCA) [16]. There is research that presents a non-invasive method for ensuring plasma parameters inside a radio-frequency ion thruster (RIT) by OES data incorporating PCA [17]. Data preprocessing with a machine learning or deep learning model is already common in artificial intelligence architecture. The PCA preprocessed the OES data, and the results were handed over to the multilayer perceptron (MLP) model for training [18]. The initialization of the variational autoencoder (VAE) used in MLP had the best result to predict plasma density of etch tool by OES data [19]. A real-time machine VM based on artificial intelligence has been developed. It can use OES data to supervise plasma nitridation processes in predicting the accuracy of electron temperature at 90% and electron density at 97% [20]. Another study presented generative adversarial networks (GANs) of deep learning to solve fault detection with imbalanced data in the plasma etching process. It verified the benefits of GANs solving imbalance OES data problems [21]. This study aims to identify the optimal pulse parameters for producing high-quality AlN films. Additionally, a predictive CatBoost model from PCA was successfully trained by collecting OES data for obtaining film FWHM and grain size.

2. Methodology

2.1. Experimental Setup and PVD Process

Figure 2 shows a high vacuum chamber and equipment used in the experiment. In this study, P-type Si (100) substrates (4-inch) were used to deposit 500-600 nm thick of AlN films in a reactive DC magnetron sputtering system with an asymmetric bipolarpulsed DC power generator (HPP-1K5A01KAT, Delta Incorporated, Taiwan) under a base pressure of 8×10^{-6} torr at room temperature. A pure aluminum target (5N) with 101.6 mm diameter is used and the distance from the target to the substrate is 80 mm. The deposition time is 30 min at various sputtering pulse parameters. The interaction variables with the factor levels and PVD process setup for a total of 15 experiments are listed in Table 1. Moreover, we use optical emission spectroscopy (OES) detector (model: SE2020-025-FUVN, OtO Photonics, Taiwan@ resolution 1.2-6.0 nm FWHM) in the process to record the plasma spectral intensity at wavelength 250-850 nm with a sampling rate of 1/3 Hz and integration time to 50 ms. The OES data played a crucial role in the modeling process. OES measurements provided a real-time, in-situ monitoring of the plasma conditions during the deposition process, enabling us to collect valuable information on the chemical composition, excitation, ionization states, and other plasma characteristics. In addition to the role, OES data serves as a valuable training dataset for the thin film state classification model. By employing the spectral information obtained from OES measurements, the model can identify and distinguish between different thin film states based on their unique spectral signatures. This allows for the development of a robust machine learning algorithm, which can effectively classify thin films based on their properties and deposition conditions.

Each Si substrate was cleaned with acetone, isopropyl alcohol, and deionized water in an ultrasonic oscillator for 3 min to remove grease and particles. Then, substrates were pickled with 2% hydrofluoric acid (HF) for one minute to remove surface native oxide. Additionally, prior to the deposition, Ar+ ion bombardment was applied to clean the Al target surface to remove the oxide layer, contamination layer, and any adsorbates [22].



Figure 2. The schematic diagram of the sputter chamber.

| Table 1. Th | ne interaction | variables of | the fact | or level | s and | sputtering | conditions | by experimer | ntal design |
|-------------|----------------|--------------|----------|----------|-------|------------|------------|--------------|-------------|
|-------------|----------------|--------------|----------|----------|-------|------------|------------|--------------|-------------|

| Maninulated Variables | Units | Sions | Levels | | | |
|-----------------------|-------|----------------|----------|------------|-----------|--|
| Manipulated variables | | Signs | Low (-1) | Medium (0) | High (+1) | |
| Reverse Voltage | V | X ₁ | 50 | 70 | 90 | |
| Pulse Frequency | kHz | X ₂ | 100 | 175 | 250 | |
| Duty Cycle | % | X ₃ | 80 | 85 | 90 | |
| Response variable | | | | | | |
| FWHM | 0 | Y | | | | |
| Sputtering parameter | Units | | Va | lues | | |
| Pressure | mTorr | | | 2 | | |
| Power | W | | 12 | 200 | | |
| Deposition time | min | 30 | | | | |
| Distance | cm | | | 8 | | |
| Substrate temperature | °C | | R | .T. | | |
| Total gas flow rate | sccm | | e | 50 | | |
| N ₂ | sccm | | 6 | 50 | | |

2.2. Measurements

X-ray diffraction (XRD) measurements were performed by a model of D8 Advance (Bruker US). $\theta/2\theta$ scans were recorded on a monochromator-less diffractometer delivering Cu-K α (λ = 1.5418 Å) radiation which was operated at 40 kV and 40 mA in 34°–40°. FWHM data were calculated from AlN (002) peak measured from the center of AlN film.

The surface morphologies of AlN films were characterized by an atomic force microscope (AFM, Nanoview 1000, FSM, China) equipped with an n-type contact cantilever silicon tip in radius size of 7 nm. Use the tapping mode to measure the surface roughness of thin films to avoid surface scratches on test pieces and to obtain higher quality images.

The characterization of film microstructure and thickness was carried out by Field emission scanning electron microscopy (FE-SEM, JSM-7000F, Japan). The cross-section of AlN films was taken from the observed AlN (002) columnar nanocrystalline, confirming its quality [23] and measurements of the film's thickness at \times 20,000 magnification. Furthermore, we take the film surface morphology images at \times 80,000 magnification to calculate the grain size. Figure 3 shows the results of the three techniques utilized in detecting the microstructure of the thin films of sample S8.



Figure 3. Typical thin film quality measurements by XRD, SEM and AFM for sample S8.

2.3. Box–Behnken Design Coupled with RSM on Experiments

Box–Behnken design coupled with RSM was employed to setup a design of experiment (DOE) for optimizing the quality of AlN films. Box–Behnken method is more efficient and requires fewer experiments to analyze the results [24,25]. In addition, through two and three-dimensional analysis, the influence of pulse parameters on films can be more intuitively found. In this experiment, the independent variables were based on three pulsed DC generator parameters in reverse voltage (X1), pulse frequency (X2) and duty cycle (X3). RSM layouts were obtained from statistical software (Minitab 19, Scientific Formosa, Inc.) that can generate surface plots from experimental data to describe trend under specific pulse parameters to model and analyze problems in which responses are influenced by three input variables.

2.4. Dimensionality Reduction and GBDT-Based Algorithm

In this study, we used the OES datasets for machine learning model training; Figure 4 shows 15 sets of OES data affected by pulse parameters with 9966 datasets. Each dataset consists of 1900 features. When the number of dimensions exceeds three, it becomes increasingly difficult to visualize the relationships between variables. To relieve high-dimension issues mentioned, we use principal component analysis (PCA) dimensionality reduction techniques to preprocess the OES datasets. The PCA is a linear transformation that seeks to find the eigenvectors (principal components) and eigenvalues (variances of the principal components) of the covariance matrix which is calculated using the original dataset. The first principal component corresponds to the eigenvector with the largest eigenvalue, and each subsequent principal component corresponds to the eigenvector with the next largest eigenvalue. These principal components are orthogonal to each other. It means that they are uncorrelated. The following is a general equation for PCA in Equation (1). Where variable *N* is the number of all data in the dataset (assuming sample *x*), v_1 and v_2 are two features in sample *x*, and v_{i1} and v_{i2} represent the *i*-th sample of v_1 and v_2 in the sample *x*, respectively.

$$Cov(v_1, v_2) = \sum_{i=1}^{N} \frac{(v_{i1} - v_1)(v_{i2} - v_2)}{N - 1}$$
(1)

We used CatBoost machine learning model for training after PCA data preprocessing. CatBoost model is a gradient boosting decision tree-based (GBDT) machine learning algorithm developed by Yandex in 2017. This algorithm model uses an ensemble method for the improvement of the accuracy of the model by combining the predictions of multiple decision trees. It can prevent overfitting by controlling the growth of decision trees and has good generalization ability. Compared to other GBDT models, the CatBoost model has better explainability, allowing direct observation of the impact of each decision tree. It also supports an adaptive loss function, which can adapt to different task requirements, such as binary classification, multi-classification, regression, etc. The equation for the final prediction can be represented by Equation (2). Where F(x) is the final prediction, W_j is the weight assigned

to the *j*-th decision tree, $h_j(x)$ is the prediction made by the *j*-th decision tree for the dataset (assuming sample *x*), and the variable *M* is the number of decision tree splits.

$$F(x) = \sum_{i}^{M} W_{j} \times h_{j}(x)$$
⁽²⁾





3. Results and Discussion

AlN films were fabricated using a reactive DC magnetron sputtering with a total of 15 different DOEs combined with three DC pulse parameters in reverse voltage (50–90 V), pulse frequency (100–250 kHz), and duty cycle (80–90%). Pulse parameters have a direct effect on the state of the plasma. Therefore, we optimized the power parameters to obtain better quality of AlN films. Table 2 presents the 15-run experimental designs from Box–Behnken method.

| Run | | Factors | |
|-----|----|---------|----------------|
| | X1 | X2 | X ₃ |
| S1 | 0 | -1 | +1 |
| S2 | +1 | +1 | 0 |
| S3 | -1 | +1 | 0 |
| S4 | -1 | -1 | 0 |
| S5 | +1 | -1 | 0 |
| S6 | 0 | 0 | 0 |
| S7 | 0 | +1 | +1 |
| S8 | 0 | 0 | 0 |
| S9 | -1 | 0 | +1 |
| S10 | +1 | 0 | +1 |
| S11 | 0 | -1 | -1 |
| S12 | +1 | 0 | -1 |
| S13 | 0 | 0 | 0 |
| S14 | 0 | +1 | -1 |
| S15 | -1 | 0 | -1 |

Table 2. 15-run experimental designs from Box–Behnken method.

3.1. Characterization

Figure 5 shows XRD ($\theta/2\theta$ scan) spectra of AlN films (S1–S15) deposited at various sputtering pulse parameters. We observed prominent peaks corresponding to the AlN crystallite phases, which are critical in assessing the quality and crystallinity of the films.

These peaks include the (100) and (002) planes, commonly reported in the literature for hexagonal wurtzite AlN films. It can provide insights into the film's overall purity and the effectiveness of the sputtering process. The microstructures of AlN films were analyzed extensively by XRD scanning in the range of 34° – 40° . The impact of three DC pulsed parameters on film FWHM was studied by varying the design of pulse parameters. It can be clearly observed that each group has an obvious AlN 002 peak at about 36° . The film FWHM was obtained by calculations from a Lorentz curve fit in the experimental data between 34° and 38° . Among them, FWHM in fifteen sets of experiments is between 0.1° – 0.506° . We observed that experiment S3 has the minimum FWHM value, indicating that the AlN film has better crystallinity under this parameter. In an asymmetric bipolar pulsed-DC reactive sputtering and to prevent the target from poisoning due to the arc causing fine dust particles deposition. The asymmetric bipolar pulsed DC is also dependent on pulse frequency. To prevent the build-up potential (voltage) from crossing the parasitic capacitors, the pulse frequency has to maintain enough beyond its breakdown voltage [26–28].



Figure 5. XRD ($\theta/2\theta$ scan) spectra of AlN films (S1–S15) deposited at various DC pulsed parameters of 50–90 V reverse voltage, 100–250 kHz pulse frequency, 80–90% duty cycle as well as N₂ concentration of 100% and flow rate of 60 sccm.

AlN is an excellent piezoelectric material for electroacoustic devices for example, bulk acoustic wave (BAW) and surface acoustic wave (SAW) devices. However, uneven surfaces or microstructure, such as cracks and voids, will cause transmission loss. Furthermore, if the roughness of the surface is greater than the traveling wavelength, the acoustic waves cannot pass through the material surface, which will affect the SAW speed [5,29]. Therefore, the surface roughness is extremely important to the device performance; we use AFM (tapping mode) and FE-SEM to measure the surface roughness, the microscopic crystal morphology, and the cross-sectional structure of AlN films. The results are shown in Table 3. The surface roughness data are recorded in the center of the film using the tapping mode and the measurement range is 5 μ m \times 1.2 μ m. This mode will bring the probe closer to the sample and increase the amplitude. The amplitude change is used to obtain high-resolution images when contacting the sample. The advantage is that the contact force between the probe and the sample can be reduced, thereby preventing the film surface from scratching. AFM images of AlN films were shown in Figure 6. The RMS surface roughness values on 15 sets of AlN films were 1.97 nm-4.46 nm. According to Figure 7, experiments S1, S7, S10, S11, and S14, which display roughness above the average, all have a reverse voltage greater than or equal to 70 V in common.

| Sample | 002 Peak FWHM (°) | Surface Roughness (nm) | Deposition Rate (nm/min) | Grain Size (nm) |
|--------|----------------------|---------------------------|-----------------------------|-----------------|
| S1 | 0.29 | 3.04 | 20.33 | 32.9 |
| S2 | 0.23 | 2.53 | 21.57 | 52.0 |
| S3 | 0.10 | 1.97 | 20.03 | 43.3 |
| S4 | 0.31 | 2.11 | 19.23 | 37.7 |
| S5 | 0.27 | 2.53 | 19.67 | 39.7 |
| S6 | 0.34 | 2.74 | 20.87 | 48.7 |
| S7 | 0.23 | 4.17 | 20.63 | 49.3 |
| S8 | 0.32 | 2.68 | 19.00 | 48.1 |
| S9 | 0.51 | 2.22 | 20.43 | 42.8 |
| S10 | 0.27 | 3.06 | 20.00 | 39.7 |
| S11 | 0.28 | 3.07 | 19.23 | 45.3 |
| S12 | 0.47 | 2.65 | 18.67 | 45.3 |
| S13 | 0.32 | 2.76 | 18.67 | 48.7 |
| S14 | 0.38 | 4.46 | 20.20 | 46.9 |
| S15 | 0.19 | 2.13 | 18.00 | 46.9 |

Table 3. Results of characterization for 15 experimental sets.



Figure 6. AFM images of AlN films (S1-S15 in RMS surface roughness) as deposited on Si (100) substrates.

Figure 8 shows the cross-sectional micrographs of AlN films deposited on silicon (100) substrates taken by FE-SEM. The 30 min process exhibited a film thickness between 540–647 nm and a deposition rate of 18.0–21.6 nm/min. It was observed that when the duty cycle was set at a small value (S11, S12, and S15), resulting in a shorter sputtering period (ton), the deposition rate was significantly lower. The duty cycle is proportional to the sputtering period in a complete current waveform. The value represented by the duty cycle is presented in Equation (3). In addition, the increase in pulse frequency will also increase plasma activity. The top-views of AlN films are depicted in Figure 9. Thin film morphology confirms the XRD data discussed above; the (002) oriented films showed pebble-like structures with average grain sizes ranging between 30.9–52.0 nm, indicating lower surface mobility of adatoms under pure nitrogen conditions [30] and a clear columnar structure is also shown in the cross-sectional micrographs in Figure 8.



Figure 7. The surface roughness with root-mean-square (RMS) after process completion for the sputtered AlN films in the various pulse parameters as shown in Figure 6 (S1–S15).

$$DutyCycle = \frac{(t_{on})}{(t_{on}) + (t_{rev})} \times 100\%$$
(3)



Figure 8. The cross-sectional micrographs of FE-SEM on AlN thin films (S1-S15 in film thickness).

 Size
 Size

Figure 9. The top-views of SEM images on AlN films (S1–S15 in average grain size) as deposited on Si (100) substrates.

3.2. Statistical Analysis and Optimization

In this experiment, the Box–Behnken factorial outline with three factors was selected to analyze the three pulse parameters. A linear response equation was generated from the measured and calculated FWHM values. Multiple regression analyses were performed on the experimental data and response variables. We established a second-order polynomial equation, as shown in Equation (4), to explain the relationship between pulse parameters and film FWHM. Equation (4) can predict the response of each factor at three levels (-1, 0, +1), the high levels were labeled as +1, and the low levels were labeled as -1. The reverse voltage (X_1) and duty cycle (X_3) have a more significant linear impact. The coefficients are 0.1064 and -0.238, respectively. In the interaction between pulse parameters, the reverse voltage and duty cycle also have the most significant impact, and its value is -0.001302. According to the above results, the duty cycle (X_3) has the greatest influence, followed by reverse voltage (X_1) and pulse frequency (X_2).

$$Y(FWHM) = 5.75 + 0.1064X_1 + 0.01129X_2 - 0.238X_3 -0.000014X_2 * X_2 + 0.00204X_3 * X_3 +0.000029X_1 * X_2 - 0.001302X_1 * X_3 - 0.000103X_2 * X_3$$
(4)

All responses from the results of variance (ANOVA) analysis by a quadratic polynomial model with a corresponding valid model obtained from Minitab 19 software are depicted in Table 4. In statistical hypothesis testing, the significance level is typically set at 0.05 or below, meaning that results with a *p*-value greater than 0.05 are generally not considered statistically significant. In general, the smaller the *p*-value is, the more significant the results are, with a *p*-value of less than 0.05 indicating strong evidence against the null hypothesis [31,32]. It can be seen from the variance analysis of the regression model that the quadratic model has a certain significance, the *p*-value is 0.04, and the R-squared value is 85.91%, suggesting that this model has strong explanatory power. Only 14.09% of the total variation remains unexplained, indicating it is a good fit model between the prediction and the data. In addition, the smaller *p*-value of the squared and interaction terms of the independent variables indicates a more significant impact of these terms in the model. However, even the R-square (adj) value is 67.13%, and a lower value may suggest a simpler model with fewer variables. The model can still provide valuable insights and predictive power for the response variable.

| Summary of Model | | | | | |
|-------------------------------|-----------------|----------------------|--------------------------|---------|-----------------|
| R-squared 85.91% | R-squ 6 | 1ared (adj) 7.13% | R-squared (pred) N.D. | | |
| ANOVA | | | | | |
| Source | DF ^a | Adj SS ^b | Adj MS ^c | F-Value | <i>p</i> -Value |
| * Model | 8 | 0.124731 | 0.015591 | 4.57 | 0.040 |
| Linear | 3 | 0.007574 | 0.002525 | 0.74 | 0.565 |
| X1 | 1 | 0.002016 | 0.002016 | 0.59 | 0.471 |
| X ₂ | 1 | 0.005512 | 0.005515 | 1.62 | 0.251 |
| X3 | 1 | 0.000045 | 0.000045 | 0.01 | 0.912 |
| * Square | 2 | 0.035624 | 0.017812 | 5.23 | 0.049 |
| $X_2 X_2$ | 1 | 0.023623 | 0.023623 | 6.93 | 0.039 |
| X_3X_3 | 1 | 0.009661 | 0.009661 | 2.83 | 0.143 |
| * 2-Way Interaction | 3 | 0.081533 | 0.027178 | 7.97 | 0.016 |
| X_1X_2 | 1 | 0.007744 | 0.007744 | 2.27 | 0.182 |
| $X_1 X_3$ | 1 | 0.067860 | 0.067860 | 19.91 | 0.004 |
| X ₂ X ₃ | 1 | 0.005929 | 0.005929 | 1.74 | 0.235 |
| Error | 6 | 0.020452 | 0.003409 | | |
| Lack of fit | 4 | 0.020168 | 0.005042 | 35.42 | 0.028 |
| Pure Error | 2 | 0.000285 | 0.000142 | | |
| Total | 14 | 0.145183 | | | |

Table 4. A summary from response surface regression analysis.

Footnotes for ^a DF: degree of freedom; ^b Adj SS: adjusted sum of squares; ^c Adj MS: adjusted mean square; * the importance of correlation parameters.

Furthermore, to better visualize and evaluate the effects of pulse parameters on optimizing the FWHM, we converted the above-mentioned second-order polynomial equation (Equation (4)) into 2D and 3D response surface plots. These plots provide an intuitive representation of the relationships among the three main variables (reverse voltage, pulse frequency, and duty cycle), shown in Figure 10.

In constructing these plots, one parameter was fixed at a medium level, while the FWHM was analyzed as a function of the other two interaction factors. Specifically, Figure 10a illustrates the response of FWHM when the reverse voltage interacts with the pulse frequency, with the duty cycle held constant at a medium level (85%). The plot shows that better FWHM values are achieved when the reverse voltage is at a low level, and the pulse frequency is at a high level. Similarly, Figure 10b displays the response of the reverse voltage interacting with the duty cycle while the pulse frequency is maintained at a medium level (175 kHz). Lastly, Figure 10c presents the response of the pulse frequency interacting with the reverse voltage fixed at a medium level (70 V).

By examining these response surface plots, we can gain a deeper understanding of the complex interactions among the pulse parameters, which aids in optimizing the FWHM based on the second-order polynomial equation (Equation (4)).

In summary, the response of the interaction of multiple variables in the above RSM diagram is integrated, and a set of optimized parameters is obtained, presented in Table 5. The best results obtained in this research are reverse voltage 50 V, pulse frequency 250 kHz, and duty cycle 80.6061%. The estimated FWHM value is 0.11°. Due to the much higher mobility of electrons compared to ions, the positive voltage is usually set at 10–20% of the negative working voltage to achieve complete discharge. However, it must be kept low enough to prevent the occurrence of back sputtering and reduce the risk of arc and target poisoning. Deposition at higher pulse frequencies enhances plasma activity and concentration, resulting in better film uniformity. Additionally, a higher duty cycle could result in an increase in the sputtering period, which may lead to a better deposition rate. However, it could also cause a breakdown voltage exceeding the parasitic capacitors on target, resulting in arcing and subsequently deteriorating the quality of the deposited film [28,33].



Figure 10. 2D views (factor vs. level) and 3D views. (**a**) 3D in X_1*X_2 : frequency interacting with reverse voltage, X_3 remains a constant; (**b**) 3D in X_1*X_3 : duty cycle interacting with reverse voltage, X_2 remains a constant; and (**c**) 3D in X_2*X_3 : duty cycle interacting with frequency, X_3 remains a constant. Here, factors are reverse voltage (X_1), pulse frequency (X_2), and duty cycle (X_3).

Table 5. Multiple response prediction.

| Factor X ₁ Reverse Voltage | Factor X1Factor X2Reverse VoltagePulse Frequency | | Response FWHM | |
|--|--|----------|------------------|--|
| 50 V | 250 kHz | 80.6061% | 0.110969° | |

3.3. Data Preprocessing and Model Training

We trained two CatBoost models to classify FWHM and grain, respectively. For the FWHM model, if the measured data exceeds the average value of 0.34 degrees, we label it as "Wide", and the opposite is labeled as "Narrow". The schematics of the labeled data are shown in Figure 11a. For the grain size model, if the measured data exceed the average value of 45.1 nm, we label it as "Large", and the opposite is labeled as "Small". The schematic of the labeled data is shown in Figure 12a.



Figure 11. (a) The schematics of the labeled data of the FWHM model. (b) The first vs. second principal components (PC0 and PC1). (c) The third vs. fourth principal components (PC2 and PC3). (d) The fifth vs. sixth principal components (PC4 and PC5).



Figure 12. (**a**) The schematic of the labeled data of the grain size model. (**b**) The first vs. second principal components (PC0 and PC1). (**c**) The third vs. fourth principal components (PC2 and PC3). (**d**) The fifth vs. sixth principal components (PC4 and PC5).

For data preprocessing, we use PCA to reduce the dimensionality of OES data from 1900 features to 6 features (PC0 to PC5). After PCA dimensionality reduction, the data is transformed into a new set of principal components (PCs) that capture the maximum variance

in the original dataset. The cumulative variance is the sum of variances of the principal components (eigenvalue of the covariance matrix illustrated in Section 2.4). The means of the cumulative variances increase as the number of principal components increases. Typically, we hope to select the smallest number of principal components that capture a large percentage of the total variance, to reduce the dimensionality of the data and simultaneously preserve the information as much as possible. The cumulative variances from PC0 to PC5 are 73.59%, 79.12%, 82.51%, 84.73%, 89.12%, and 91.14%, respectively. The cumulative variances mean the total variation in a dataset explained by each principal component. From the calculation results using Python software (scikit-learn toolkit), we can interpret 91.14% of the features in the dataset through six principal components. The dimensionality data can improve the calculating time and eliminate outliers and visualization effectively after dimensionality reduction. From the visualized results, the same labeled data are well clustered. The well-clustered data refer to a dataset where the data points are grouped closely together after transformed space. When the data is well-clustered, it becomes easier for machine learning models to learn the relationships between the features and the target variable. As a result, machine learning models trained on well-clustered data tend to have better training performance, including higher accuracy and faster convergence during training. Figures 11b,c and 12b,c, show the results of dimensionality reduction. The blue and orange dots signify two distinct types of data. Despite the proximity of the clusters, the results of the dimensionality reduction reveal a clear and discernible clustering pattern. This noticeable separation between the data groups suggests that the accuracy of the prediction model is likely to be enhanced. Consequently, the clustering visualized in the figure serves as a valuable indicator of the model's potential performance.

The machine learning model training process typically involves two important steps: data split and hyperparameter tuning. The data split involves dividing the available dataset into training, validation, and testing sets. The training data are used to train the model, while validation data are used to evaluate the model's performance during training and select the best hyperparameters. Finally, the test data are used to evaluate the final performance of the trained model after hyperparameter tuning. Hyperparameter tuning involves selecting the best value of hyperparameters that maximizes the model's performance on validation data. Once the best set of hyperparameters is selected, the model can be trained again using the entire training data, and the final performance can be evaluated on the test data. Following these two steps can ensure that the machine learning model is properly trained and optimized for the given dataset.

For CatBoost model training, the datasets have been split into the training dataset 70%, the validation dataset 15%, and the testing dataset 15%. This distribution to splitting dataset into training, validation, and test subsets allows for a sufficiently large training set (70%) to develop and refine the prediction model while also providing separate validation (15%) and test (15%) sets to assess its performance and generalization capabilities. The validation set is useful for tuning model parameters and preventing overfitting, while the test set serves as a final evaluation to ensure that the model performs well on new or unseen data. This method of dataset partitioning is commonly used in machine learning applications and is considered good practice for robust model evaluation. We mainly tuned the parameters of iteration and learning rate. We fixed the parameters of max_depth = 3 and 12_leaf_reg = 4, and both the FWHM and grain size models are set to be the same iteration (=1000) and learning rate (=0.01). The max_depth and l2_leaf_reg are the hyperparameters in the CatBoost. The parameters of the max_depth can control the maximum depth of each decision tree in the ensemble model (CatBoost in this study). The l2_leaf_reg can control the L2 regularization to the leaf weights of each decision tree in the ensemble. The best value for max_depth and l2_leaf_reg depends on the scale and complexity of the dataset. Then the trained model for FWHM and grain size received the performance metrics as shown in Table 6. The FWHM model had a recall of 89.52%, precision of 94.51%, F1-score of 91.95%, and accuracy of 90.67%. The grain size model had a recall of 86.40%, precision of 79.61%, F1-score of 82.87%, and accuracy of 82.07%. Four basic terms that refer to four possible outcomes of binary classification in machine learning are the values of *TP* (true

positive), *TN* (true negative), *FP* (false positive), and *FN* (false negative). TP and TN are correct prediction results, and *FP* and *FN* are incorrect. The values of *TP*, *FP*, *FN*, and *TN* as accuracy, precision, recall, and F1-score, respectively, can be adopted to evaluate the performance of a prediction model. Details are illustrated in Equations (5)–(7).

$$Recall = \frac{TP}{TP + FN}$$
(5)

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

$$F1-score = \frac{2 \times Recall \times Precision}{Recall + Precision}$$
(7)

Table 6. Basic terms and critical indicators of two trained models.

| Goal | Fundamental Terms | | | | Performance Metrics | | | |
|---------------|-------------------|--------|--------|--------|---------------------|---------------|--------------|--------------|
| | TP [%] | FP [%] | FN [%] | TN [%] | Recall [%] | Precision [%] | F1-Score [%] | Accuracy [%] |
| FWHM | 53.32 | 3.09 | 6.23 | 37.35 | 89.52 | 94.51 | 91.95 | 90.67 |
| Grain Size | 43.36 | 11.10 | 6.82 | 38.70 | 86.40 | 79.61.51 | 82.87 | 82.07 |

4. Conclusions

We use Box–Behnken experimental design to reduce the number of experiments and collected data to develop a reliable model, which enables us to make more appropriate adjustments to parameters according to device requirements, thus saving a lot of time and effort in improving its quality. We obtained the optimized conditions of the pulse parameters of the best FWHM as reverse voltage: 50 V, pulse frequency: 250 kHz, and duty cycle: 80.6061%. In addition, using AFM and SEM to analyze the film's microstructure, the relationship among surface roughness, deposition rate and pulse parameters was obtained. We compared relevant literature and found that the FWHM ranged from $0.28^{\circ}-0.36^{\circ}$ and decreased with a decrease in duty cycle (20–70%). However, the pulse frequency used in the literature was 100 kHz, which may not be sufficient to achieve the desired plasma activity and concentration [34]. Compared with other studies using asymmetric bipolar pulse deposition of AlN thin films, our results also showed better performance [35]. This study also successfully trained two models that can predict film FWHM and grain size. Machine learning algorithms can analyze amounts of data and identify OES data features that are difficult for humans to find. Accurately predicting a semiconductor deposition process can reduce the number of failed runs and improve overall efficiency. It can make more accurate predictions about the deposition process, reducing the time and cost required for testing. Moreover, machine learning algorithms can provide insights into the deposition process that may not be apparent through traditional analysis methods. This can help engineers make more informed decisions about the process and improve the overall quality of the semiconductor products produced. In the future, we will develop more expertise in plasma physics, data analysis, and machine learning, as well as access to the necessary equipment, and data (parameters) to develop a complete plasma information-based virtual metrology (PI-VM) system. This can provide real-time feedback to engineers, enabling them to optimize the semiconductor processing and improve yields. We know that uncertainty factors can impact the prediction model, specifically the blank signals and noise measured by the OES during the setup period. Suppose these noise components are inadvertently included in the model training process without proper attention or filtering. In that case, the model's prediction accuracy may be compromised, resulting in errors and reduced reliability of the predictions. The novelty of this study is that it verifies the application of the PCA method to rapidly reduce the dimensionality of OES data and validates the feasibility of using reduced OES data for predicting film quality (FWHM and grain size). Since the prediction model belongs to a classification model that classifies two classes (above or below the average value), it cannot be used to predict the actual value of FWHM and grain size. A predictive model of actual values is planned for the future.

5. Future Work

One of the primary goals for our future studies is to investigate the density of AlN films, as it is a crucial parameter for semiconductor applications and is directly related to the sputtering parameters. We will explore opportunities for collaboration with other laboratories or institutions to access this technology. By utilizing X-ray reflectivity (XRR) equipment, and measurements, we can estimate the density of the AlN films and examine its relationship with the sputtering conditions. Furthermore, we will seek to refine our prediction model by incorporating additional parameters, such as film stress, adhesion, and thermal conductivity, to achieve a more comprehensive understanding of the film properties. Additionally, we plan to compare the results obtained using our optimized parameters with previously published studies to validate the effectiveness of our approach. Through these efforts, we aim to enhance the optimization of sputtering parameters for AlN film deposition in semiconductor applications and contribute to the broader body of knowledge in this field.

In the upcoming experiments, we will model and simulate the problem of target erosion. First, develop a computer model to simulate the sputtering process and predict the effect of target erosion on film properties. The model incorporates parameters such as target material, deposition conditions, and erosion rate Incorporate parameters such as target material, deposition conditions, and erosion rate. Then, use the model to simulate the sputtering process under varying degrees of target erosion and predict the corresponding film properties. The software expected to be used for computer modeling and simulation of sputtering processes is COMSOL Multiphysics.

Author Contributions: W.-Y.Z., X.-L.T. and H.-F.C. carried out methodology, software and validation with support from H.-F.C. W.-Y.Z. took the lead in writing the manuscript with support from T.T.L., Y.-K.F. and H.-F.C. for review and editing; Y.-K.F. and T.T.L. conceived the ideas and followed investigations; H.-H.L., P.J.W. and M.-Y.J. helped perform the analysis with constructive discussions. All authors provided critical feedback and helped shape the research, analysis, and manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially funded by Delta Electronics, Inc. under contract No. 108K92.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in article.

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

References

- 1. Talukder, A.A.; Baule, N.; Steinhorst, M.; Shrestha, M.; Fan, Q.H.; Schuelke, T. Pulsed direct-current reactive sputtering of high Young's modulus [002] oriented aluminum nitride thin films. *Thin Solid Films* **2022**, 751, 139239. [CrossRef]
- Lu, T.Y.; Yang, Y.P.; Lo, H.H.; Wang, P.J.; Lai, W.; Fuh, Y.K.; Li, T.T. Minimizing film residual stress with in situ OES big data using principal component analysis of deposited AlN films by pulsed DC reactive sputtering. *Int. J. Adv. Manuf.* 2021, 114, 1975–1990. [CrossRef]
- Panda, P.; Ramaseshan, R.; Ravi, N.; Mangamma, G.; Jose, F.; Dash, S.; Suematsu, H. Reduction of residual stress in AlN thin films synthesized by magnetron sputtering technique. *Mater. Chem. Phys.* 2017, 200, 78–84. [CrossRef]
- Lu, Y.; Reusch, M.; Kurz, N.; Ding, A.; Christoph, T.; Kirste, L.; Žukauskaitė, A. Surface morphology and microstructure of pulsed DC magnetron sputtered piezoelectric AlN and AlScN thin films. *Phys. Status Solidi A* 2018, 215, 1700559. [CrossRef]
- 5. Sanz-Hervás, A.; Clement, M.; Iborra, E.; Vergara, L.; Olivares, J.; Sangrador, J. Degradation of the piezoelectric response of sputtered c-axis AlN thin films with traces of non-(0002) X-ray diffraction peaks. *Appl. Phys. Lett.* **2006**, *88*, 161915. [CrossRef]
- Lo, H.H.; Chen, W.L.; Wang, P.J.; Lai, W.; Fuh, Y.K.; Li, T.T. Residual stress classification of pulsed DC reactive sputtered aluminum nitride film via large-scale data analysis of optical emission spectroscopy. *Int. J. Adv. Manuf. Technol.* 2022, 119, 7449–7462. [CrossRef]

- D'Angelo, A.; Diletto, C.; Esposito, S.; Graditi, G.; Guglielmo, A.; Lanchi, M.; Rossi, G. Spectrally selective solar coating based on W-AlN cermet fabricated by reactive sputtering processes at high deposition rate. In *AIP Conference Proceedings*; AIP Publishing LLC: Melville, NY, USA, 2022; Volume 2445, No. 1; p. 020002.
- Belkind, A.; Freilich, A.; Lopez, J.; Zhao, Z.; Zhu, W.; Becker, K. Characterization of pulsed dc magnetron sputtering plasmas. New J. Phys. 2005, 7, 90. [CrossRef]
- Vlček, J.; Pajdarová, A.D.; Musil, J. Pulsed dc magnetron discharges and their utilization in plasma surface engineering. *Contrib. Plasma Phys.* 2004, 44, 426–436. [CrossRef]
- Musil, J.; Leština, J.; Vlček, J.; Tölg, T. Pulsed dc magnetron discharge for high-rate sputtering of thin films. J. Vac. Sci. Technol. A 2001, 19, 420–424. [CrossRef]
- 11. Lin, J.; Moore, J.J.; Sproul, W.D.; Mishra, B.; Wu, Z.; Wang, J. The structure and properties of chromium nitride coatings deposited using dc, pulsed dc and modulated pulse power magnetron sputtering. *Surf. Coat. Technol.* **2010**, 204, 2230–2239. [CrossRef]
- 12. Yeh, T.S.; Wu, J.M.; Hu, L.J. The properties of TiN thin films deposited by pulsed direct current magnetron sputtering. *Thin Solid Films* **2008**, *516*, 7294–7298. [CrossRef]
- Safi, I. Recent aspects concerning DC reactive magnetron sputtering of thin films: A review. Surf. Coat. Technol. 2000, 127, 203–218. [CrossRef]
- 14. Adamczyk, J.; Horny, N.; Tricoteaux, A.; Jouan, P.Y.; Zadam, M. On the use of response surface methodology to predict and interpret the preferred c-axis orientation of sputtered AlN thin films. *Appl. Surf. Sci.* **2008**, 254, 1744–1750. [CrossRef]
- 15. Ogar, V.N.; Hussain, S.; Gamage, K.A. Transmission line fault classification of multi-dataset using catboost classifier. *Signals* **2022**, 3, 468–482. [CrossRef]
- 16. Chien, K.C.; Chang, C.H.; Djurdjanovic, D. Virtual metrology modeling of reactive ion etching based on statistics-based and dynamics-inspired spectral features. J. Vac. Sci. Technol. B Nanotechnol. Microelectron. 2021, 39, 064003. [CrossRef]
- 17. Nauschütt, B.T.; Chen, L.; Holste, K.; Klar, P.J. Non-invasive assessment of plasma parameters inside an ion thruster combining optical emission spectroscopy and principal component analysis. *EPJ Tech. Instrum.* **2021**, *8*, 13. [CrossRef]
- 18. Yang, Y.P.; Lu, T.Y.; Lo, H.H.; Chen, W.L.; Wang, P.J.; Lai, W.; Fuh, Y.K.; Li, T.T. Machine learning assisted classification of aluminum nitride thin film stress via in-situ optical emission spectroscopy data. *Materials* **2021**, *14*, 4445. [CrossRef]
- 19. Kwon, O.; Lee, N.; Kim, K. Improvement of virtual diagnostics performance for plasma density in semiconductor etch equipment using variational auto-encoder. *IEEE Trans. Semicond. Manuf.* 2022, *35*, 256–265. [CrossRef]
- Park, J.H.; Cho, J.H.; Yoon, J.S.; Song, J.H. Machine learning prediction of electron density and temperature from optical emission spectroscopy in nitrogen plasma. *Coatings* 2021, 11, 1221. [CrossRef]
- Choi, J.E.; Seol, D.H.; Kim, C.Y.; Hong, S.J. Generative adversarial network-based fault detection in semiconductor equipment with class-imbalanced data. Sensors 2023, 23, 1889. [CrossRef]
- 22. Chodun, R.; Nowakowska-Langier, K.; Zdunek, K. Methods of optimization of reactive sputtering conditions of Al target during AlN films deposition. *Mater. Sci.* 2015, 33, 894–901. [CrossRef]
- 23. Xu, X.H.; Wu, H.S.; Zhang, C.J.; Jin, Z.H. Morphological properties of AlN piezoelectric thin films deposited by DC reactive magnetron sputtering. *Thin Solid Films* **2001**, *388*, 62–67. [CrossRef]
- Khatti, T.; Naderi-Manesh, H.; Kalantar, S.M. Application of ANN and RSM techniques for modeling electrospinning process of polycaprolactone. *Neural Comput. Appl.* 2019, 31, 239–248. [CrossRef]
- Ewulonu, C.M.; Chukwuneke, J.L.; Nwuzor, I.C.; Achebe, C.H. Fabrication of cellulose nanofiber/polypyrrole/polyvinylpyrrolidone aerogels with box-Behnken design for optimal electrical conductivity. *Carbohydr. Polym.* 2020, 235, 116028. [CrossRef]
- Sellers, J. Asymmetric bipolar pulsed DC: The enabling technology for reactive PVD. Surf. Coat. Technol. 1998, 98, 1245–1250. [CrossRef]
- Schiller, S.; Goedicke, K.; Reschke, J.; Kirchhoff, V.; Schneider, S.; Milde, F. Pulsed magnetron sputter technology. Surf. Coat. Technol. 1993, 61, 331–337. [CrossRef]
- Cherng, J.S.; Chang, D.S. Effects of pulse parameters on the pulsed-DC reactive sputtering of AlN thin films. *Vacuum* 2009, 84, 653–656. [CrossRef]
- 29. Liu, H.Y.; Tang, G.S.; Zeng, F.; Pan, F. Influence of sputtering parameters on structures and residual stress of AlN films deposited by DC reactive magnetron sputtering at room temperature. *J. Cryst. Growth.* **2013**, *363*, 80–85. [CrossRef]
- Cheng, H.; Sun, Y.; Zhang, J.X.; Zhang, Y.B.; Yuan, S.; Hing, P. AlN films deposited under various nitrogen concentrations by RF reactive sputtering. J. Cryst. Growth. 2003, 254, 46–54. [CrossRef]
- 31. Baligidad, S.M.; Chandrasekhar, U.; Elangovan, K.; Shankar, S. RSM optimization of parameters influencing mechanical properties in selective inhibition sintering. *Mater. Today Proc.* **2018**, *5*, 4903–4910. [CrossRef]
- 32. Abid, S.; Messadi, R.; Hassine, T.; Ben Daly, H.; Soulestin, J.; Lacrampe, M.F. Optimization of mechanical properties of printed acrylonitrile butadiene styrene using RSM design. *Int. J. Adv. Manuf. Technol.* **2019**, *100*, 1363–1372. [CrossRef]
- 33. Arnell, R.D.; Kelly, P.J. Recent advances in magnetron sputtering. Surf. Coat. Technol. 1999, 112, 170–176. [CrossRef]

- 34. Takenaka, K.; Satake, Y.; Uchida, G.; Setsuhara, Y. Low-temperature formation of c-axis-oriented aluminum nitride thin films by plasma-assisted reactive pulsed-DC magnetron sputtering. *Jpn. J. Appl. Phys.* **2017**, *57*, 01AD06. [CrossRef]
- 35. Iqbal, A.; Mohd-Yasin, F. Reactive sputtering of aluminum nitride (002) thin films for piezoelectric applications: A review. *Sensors* **2018**, *18*, 1797. [CrossRef]

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