

Virtual metrology modeling of reactive ion etching based on statistics-based and dynamics-inspired spectral features

Cite as: J. Vac. Sci. Technol. B **39**, 064003 (2021); <https://doi.org/10.1116/6.0001277>

Submitted: 09 July 2021 • Accepted: 17 September 2021 • Published Online: 07 October 2021

 Kun-Chieh Chien,  Chih-Hao Chang and Dragan Djurdjanovic

COLLECTIONS

Paper published as part of the special topic on [64th International Conference on Electron, Ion, And Photon Beam Technology and Nanofabrication, EIPBN 2021](#)



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Particle Accelerators: Their Triumphant History and Uncertain Future](#)

The Physics Teacher **59**, 521 (2021); <https://doi.org/10.1119/10.0006457>

[Perspectives for next generation lithium-ion battery cathode materials](#)

APL Materials **9**, 109201 (2021); <https://doi.org/10.1063/5.0051092>

[Windsurfing analysis breaks down propulsion mechanics of sailboards](#)

Scilight **2021**, 411105 (2021); <https://doi.org/10.1063/10.0006710>



Advance your science and
career as a member of

AVS

LEARN MORE



Virtual metrology modeling of reactive ion etching based on statistics-based and dynamics-inspired spectral features

Cite as: J. Vac. Sci. Technol. B 39, 064003 (2021); doi: [10.1116/6.0001277](https://doi.org/10.1116/6.0001277)

Submitted: 9 July 2021 · Accepted: 17 September 2021 ·

Published Online: 7 October 2021



Kun-Chieh Chien, Chih-Hao Chang,^{a)} and Dragan Djurdjanovic^{a)}

AFFILIATIONS

Walker Department of Mechanical Engineering, The University of Texas at Austin, Austin, Texas 78712

Note: This paper is part of the Special Collection: 64th International Conference on Electron, Ion, And Photon Beam Technology and Nanofabrication, EIPBN 2021.

^{a)}Authors to whom correspondence should be addressed: chichang@utexas.edu and dragand@me.utexas.edu

ABSTRACT

Due to increasing demand on the fabrication yield and throughput in micro/nanoscale manufacturing, virtual metrology (VM) has emerged as an effective data-based approach for real-time process monitoring. In this work, a novel automated methodology, without the need for domain knowledge and experience, for extracting useful features from raw optical emission spectroscopy (OES) data is presented. Newly proposed OES features are combined with other types of data, which include tool settings, sensor readings, physical measurements, non-numerical data, and process control parameters. Using partial least squares and support vector regression, VM models for predicting the critical dimension after reactive ion etching are built. The results from the VM model indicate that the coefficient of determination of up to 0.65 and the root mean square Error of 0.08 can be achieved. Compared to the traditional features obtained by the current solution in industry, the performances of VM models via the proposed methodology can enhance the coefficient of determination by 62.5% and reduce the root mean square error by 23.1%.

Published under an exclusive license by the AVS. <https://doi.org/10.1116/6.0001277>

I. INTRODUCTION

Virtual metrology (VM) refers to methods that employ available manufacturing process-related data and relevant sensor readings to predict the properties of the resulting product. In micro/nanoscale manufacturing, applying VM to the process brings several advantages due to the increasing demand on the fabrication yield and throughput. Advanced process control can be enhanced by replacing traditional lot-to-lot control with wafer-to-wafer control.¹ In addition, VM can inform and improve the sampling selection of the physical measurement. The current approach relies on physical metrology of randomly sampled products and is generally not time and cost-efficient since nondefective wafers are measured, while defective ones are easily missed. With the assistance of VM, only wafers with undesirable predicted qualities can be sent for further inspection. However, plasma processes, such as reactive ion etching (RIE), involve complex physical and chemical processes, which inherently complicate VM and make it difficult to apply. Furthermore, due to fluctuation in process conditions and

tool aging, root causes or key variables for VM models are hard to define.

In the RIE process, the moment that the target layer to be etched is completely removed is called the endpoint. To control the quality and fidelity of etching structures, detecting an endpoint in the process is critical. Optical emission spectroscopy (OES), depicting spectral emissions of the etch plasma at different wavelengths, is widely used for endpoint detection in the RIE process because the characteristic emission can be used to determine the chemical compound that is being etched. The concept of endpoint detection by using OES data is shown in Fig. 1. The endpoint detection selects an OES signal at a specific wavelength to be monitored, which corresponds to the critical chemical species in the recipe. If the process ends before the endpoint, etching of the target material is not completed. On the contrary, etching past the endpoint results in etching beyond the target layer and also deteriorates the outgoing etch quality. Endpoint detection methodologies and other applications of OES endpoint detection are summarized by

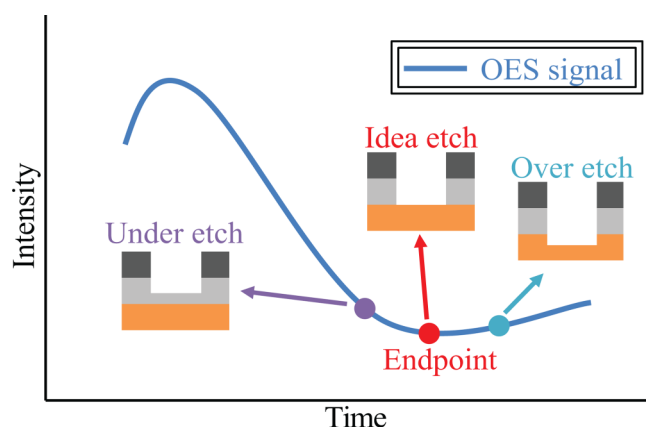


FIG. 1. OES signal and endpoint detection. Before the endpoint, the target material is not removed completely. After the endpoint, etching is beyond the target layer.

Herman.² In more recent work, several studies^{3–8} presented the use of principal component analysis (PCA) or partial least square (PLS) to eliminate the correlated relationship among OES data in order to detect the endpoint of plasma etching. Jeon *et al.*⁹ further used PCA and hidden Markov model, with hidden Markovian states representing nominal etching and over-etching. Other machine learning (ML) techniques, such as neural networks and K-means cluster analysis, have also been used to determine the endpoint in several studies.^{10–14}

In addition to endpoint detection, OES data have also been used in VM modeling using different methods and aiming to predict various postetch quality characteristics. Hong *et al.*¹⁵ applied PCA and autoencoder neural networks on OES data to model the etch rate, selectivity, uniformity, and anisotropy. Hirai and Kano¹⁶ used locally weighted PLS to predict the etching conversion differential. Ragnoli *et al.*¹⁷ compared the performance of principal component regression (PCR), PLS, forward selection component analysis, and forward selection regression on the prediction of etch rate. Etch bias was predicted by using various statistical techniques in Zeng's study.¹⁸ A fused lasso algorithm was used to build a reliable VM model.¹⁹ Park *et al.*²⁰ used PCR to build a VM for etch rate. Maggipinto *et al.*^{21,22} used a convolutional autoencoder approach and a computer vision-inspired deep learning architecture to predict the etch rate. Plasma information (PI) parameters based on plasma physics were used to develop PI-VM models to predict the etching depth and profile and to enhance the performance of predictive control.^{23–26}

OES can be seen as a dense set of time series of wavelength-specific spectral intensities emitted from the plasma during the etching process. Given such an abundance of data, extraction of useful information from those time series becomes a highly challenging and important task. On the other hand, the sample size is another concern for building VM models. Not every wafer is measured in an industrial manufacturing line, which reduces the amount of available data used to build a VM model. Also, since the etching mechanism and optical signal characteristics for different

etching tools and recipes are very dissimilar, creating a global VM model is challenging and can have limited performance. Therefore, multiple VM models created based on different conditions are necessary and which will further reduce the data size. Therefore, a method to extract features without the limitation of the size of data is needed. The current industrial solution or a traditional method for extraction OES signals is to analyze some fixed portions of several predefined OES key signals based on etch mechanisms of the process and user experience. However, the etch mechanisms for complex processes can be difficult to completely capture, and not all end-users have similar levels of domain knowledge. This can lead to some important information not being retained in the feature extraction process.

As a result, based on the above considerations, a methodology proposed by Ul Haq and Djurdjanovic²⁷ for extracting statistics-based and dynamics-inspired signatures from whole time series is the best fit for this work based on several reasons. First, it can fully depict the underlying dynamics of the system or process since it analyzes the whole length of signals. Second, since the proposed method is not based on AI/ML, it is not constrained by the small sample size. In addition, features extracted by this method are highly explainable, compared to features extracted based on deep learning techniques such as convolutional autoencoder. Those highly explainable features can provide further information for process enhancement.

In this article, we propose an automated method using PCA and statistics-based and dynamics-inspired features to realize VM in the RIE process without the need for domain knowledge and experience. As described in Sec. II, the dimensionality of OES data is reduced using PCA to form integration bands, after which statistics-based and dynamics-inspired spectral features are extracted from those integration bands and single key wavelengths using methods from Ul Haq and Djurdjanovic.²⁷ The procedure proposed in Sec. II is illustrated in Fig. 2. The results in Sec. III show that PCA can successfully identify key wavelengths of sapphire etching, and the proposed method performs better compared with the VM model built on traditional features extracted based on domain knowledge and user experience.

II. AUTOMATED FEATURE EXTRACTION FROM OES DATA

The OES signals contain information about how spectral emissions from the plasma and record how the presented chemical species evolve over time during the RIE process. However, the amount of OES data is overwhelming for VM modeling, with signals often being noisy and fluctuate with chamber conditions, which further complicates the analysis. Hence, dimensionality reduction is necessary for extracting useful information from the OES signals. The traditional approach to extract features is based on the users' domain knowledge and experience. In this process, one or several wavelengths are selected as key wavelengths that represented critical species and elements based on the understanding of the recipe and the etching process. Certain types of desired features within several fixed time intervals in those key wavelengths are further identified according to the characteristics of key OES signals. Let us take CF₄ RIE of silicon as an example. In that

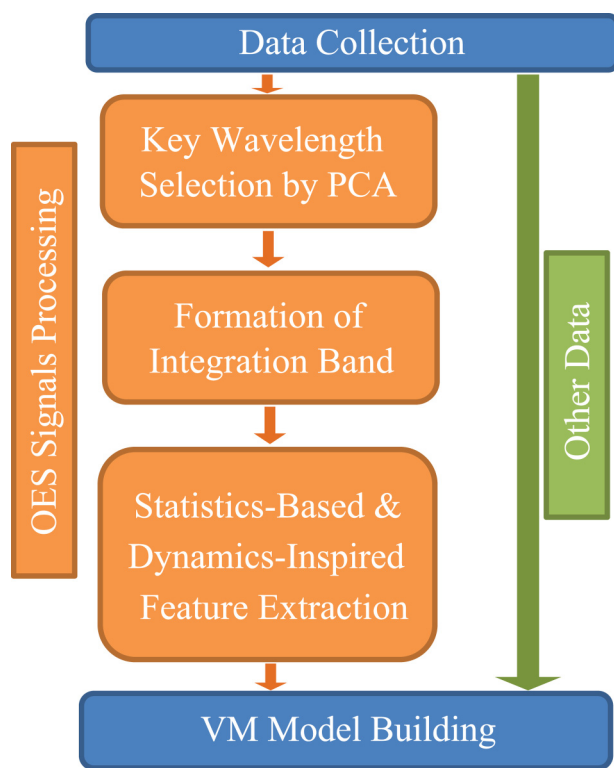


FIG. 2. Framework of VM modeling proposed in this article. OES signals are processed via key wavelength selection by PCA, formation of integration band, and statistics-based and dynamics-inspired feature extraction before inputting into the VM model. Other data are directly inputted into the VM model.

process, the wavelength of 440 nm is usually identified as the key wavelength for the endpoint detection since it corresponds to the volatile species SiF_3 . A schematic illustration of the traditional approach to the extraction of useful features from time-domain evolutions of OES intensities for a single wavelength is illustrated in Fig. 3. Here the mean, maximum, and minimum in an interval from 5 to 7 s are extracted because there are abrupt changes of signal levels at those time instances. Between 10 and 13 s, only the mean within this period is chosen since the signal is relatively steady. After 15 s, the OES signal decreases gradually, so the slope within this interval is extracted.

Unfortunately, this traditional method cannot fully retain the information in the OES data. The inherent drifting of OES signals makes extracting features within several fixed intervals impractical. Key wavelengths also become more difficult to identify if multiple gases or etch materials are used in the RIE process. Furthermore, important information embedded across multiple wavelengths is ignored if only one or a few wavelengths are selected. Namely, OES emissions of chemical species do not occur only at a few individual wavelengths but across a continuum of wavelengths, and this entire spectrum could potentially contain useful information. Therefore, including multiple wavelengths rather than taking only one single

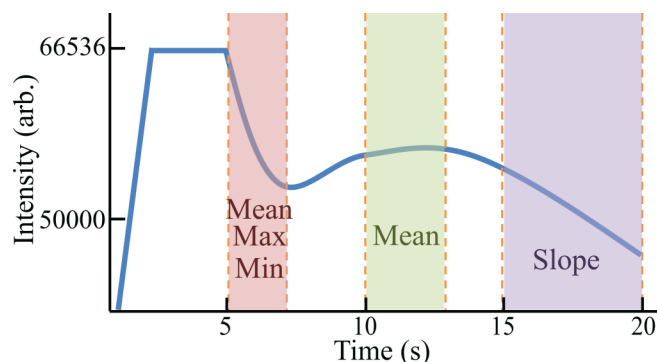


FIG. 3. Schematic of the traditional method to extract features of a key OES signal. The mean, maximum, and minimum from 5 to 7 s are extracted because of abrupt changes of signal levels. The mean between 10 and 13 s is extracted because the signal is relatively steady. The slope is extracted after 15 s because of a gradual decrease in the signal.

wavelength can enhance the robustness and reliability of VM models to reduce the effects of noise and the signal drift. Consequently, a better method to extract information in OES data is needed.

To that end, in this article, we propose an automated approach for processing OES data. Initially, the dimensionality of OES readings is reduced using PCA-based formation of integration bands of wavelengths. Then, statistics-based and dynamics-inspired features are extracted from selected key wavelengths and integration bands to construct the VM models.

A. Key wavelength selection by PCA

PCA is a multivariate data analysis method commonly used for dimensionality reduction through the projection of correlated data onto a smaller set of uncorrelated data streams, while preserving as much variation from the original dataset as possible.^{28,29} The methodology is to apply PCA to OES data to select the key wavelengths. The covariance of a given OES data X , a $m \times n$ matrix of m timestamps and n wavelengths, is scaled to have column-wise zero mean and then factorized by eigen-decomposition. This operation leads to the OES data X being the product of the score matrix T and the loading matrix W with the additional residual matrix E ,

$$X = TW^T + E, \quad (1)$$

where

$$T = [t_1, t_2, \dots, t_l], \quad (2)$$

$$W = [w_1, w_2, \dots, w_l], \quad (3)$$

and l is the number of retained principal components which is normally much less than n . The term TW^T represents the PCA modeled component, and residual matrix E is the unmodeled component. Since both the score matrix T and loading matrix W are

orthogonal, Eq. (1) can be rewritten as follows:

$$t_k = X \cdot w_k, \text{ where } k \leq l. \quad (4)$$

Since the score vector t_k is the projection of the original OES data matrix X onto the k th principal component, Eq. (4) implies that loading vector w_k shows the contribution of each wavelength to form score vector t_k . The key wavelengths can then be selected by setting the threshold for loadings.³ First, the sum of squares of all the loading values in each principal component at each wavelength is calculated by

$$r_i = \sum_{k=1}^l w_{ik}^2, \quad (5)$$

where r_i represents the overall contribution of the i th wavelength and w_{ik} is the loading at the i th wavelength on the k th principal component. Then, if r_i is larger than a preselected threshold r , the i th wavelength is selected as a key wavelength. In this work, the average overall contribution R calculated as

$$R = \frac{\sum_{i=1}^n r_i}{n} \quad (6)$$

is used as the threshold for key wavelength selection.

Since the PCA deposition for different wafers has slightly different sets of key wavelengths, another criterion is imposed for wavelength selection of the batch, which is that a wavelength needs to be selected as a key wavelength in over 80% of wafers.

B. Formation of integration bands of OES wavelengths

Once the key wavelengths have been identified by PCA, another challenge is the treatment of those contiguous key wavelengths. An integration band, i.e., integrating a range of OES spectra, is a common practice in industry to monitor the endpoint in the RIE process, since qualitative information may occur in neighboring wavelengths due to signal drifting in the spectrum. In this article, two approaches to decide the width of the integration band, namely, how many neighboring key wavelengths should be integrated to form an integration band, are discussed.

One method is based on summing all contiguous key wavelengths together to form an integration band. For example, OES is recorded only at wavelengths that are integers of nanometers in lengths, and wavelengths of 301, 302, 303, 304, 305, 306, and 400 nm are selected as the key wavelengths. Then, the OES intensities between 301 and 306 nm are summed together to form an integration band. On the other hand, since 400 nm does not have any neighboring key wavelengths, no integration band is formed. The other method considered in this article further divides contiguous key wavelength sets into smaller subsets based on peaks and valleys of their loading values. In the previous example, let 0.2, 0.4, 0.3, 0.5, 0.6, and 0.2 be respective loading values for wavelengths between 301 and 306 nm. In that case, the peaks of loading values happen to be at 302 and 305 nm, while the valley of loading values is at 303 nm. Based on that, this set of continuous key wavelengths is divided into two subsets, one between 301 and 303 nm and the other between 304 and 306 nm. Thus, instead of just one

integration band between wavelengths of 301 and 306 nm, two integration bands would be formed if the second method is used—one integration band would be formed by summing OES intensities between 301 and 303 nm, while the second integration band would be formed by summing OES intensities between 304 and 306 nm.

C. Statistics-based and dynamics-inspired feature extraction

Dynamic and statistical characteristics of the temporal evolution of OES intensities for various wavelengths contain important information about the RIE process. Following Ul Haq and Djurdjanvic,²⁷ one effective approach to analyze those time series of spectral intensities is to parse those data streams into a sequence of steady state and transient sections and extract dynamics-inspired features³⁰ from transient signal portions, while steady-state portions of those data streams can be characterized via a set of statistics-based signatures.

To autonomously parse the OES signals into steady state and transient regions, a noise threshold Δ_T is needed. To do so, several steps are presented, including filtering the signal by a low-pass filter, determining the gradient of the filtered signal, and calculating the maximum standard deviation σ of data points within the regions that gradient below a predefined threshold. Then, the noise threshold Δ_T can be obtained by

$$\Delta_T = c\sigma,$$

where

$$c \in \text{constant}. \quad (7)$$

A moving window with a predefined length M and a width $2\Delta_T$ is used to parse the OES signal by sliding it along the signal until more than 90% of the points are within the window. Once this condition is satisfied, it indicates that a steady state has been reached and the initial point of the window becomes the start point of this steady-state region. To determine the end of the steady state, the window is extended until less than 90% of the points are within the window. The window function then returns to its original size to find the next steady-state region. Segments of time series of wavelength-specific OES intensities that reside in-between thus-identified windows of steady-state behavior can be seen as transient segments. A representative segmentation result for a 440 nm OES signal is shown in Fig. 4. Once the signals are parsed into steady states and transients, a set of statistics-based signatures, such as mean, standard deviation, and kurtosis, can be extracted from steady-state segments, while dynamics-inspired features, such as rise time, overshoot, settling time, and area under the transient curve, can be extracted from the transient sections. More details on the features extracted from parsed OES data streams will be discussed in Sec. III.

After extracting features out of time series of wavelength-specific OES intensities, feature alignment and feature selection are needed in order to build a VM model. Following Ul Haq and Djurdjanvic,²⁷ the feature alignment process is accomplished by appropriately inserting “placeholders” if the length of a feature set is smaller than the longest feature set. Feature selection is accomplished in two stages. In the first stage, a threshold is set for the correlation coefficient between the feature and the system response,

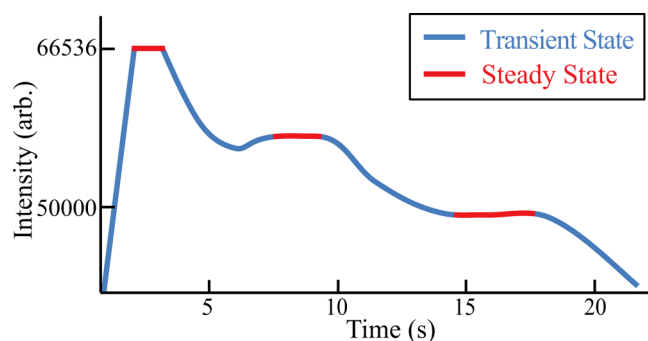


FIG. 4. Transient and steady states of a 440 nm OES signal.

which in this case is the critical dimension of the nanostructures at the end of the RIE process. In the second stage, recursive feature elimination with cross-validation is applied in order to further decrease the number of features employed in the VM modeling.

III. RESULTS AND DISCUSSION

To evaluate the novel proposed method, we present the analysis result of OES data from two RIE studies. The first study aims to demonstrate the key wavelength selection using PCA and is conducted on sapphire etching using an inductively coupled plasma reactive ion etching (ICP-RIE) tool. The second study involves VM modeling of critical dimensions of silicon nanostructures during the RIE pattern transfer process. Advantages of the use of newly available OES features extracted methods described in Sec. II are evaluated through comparisons with VM results obtained using traditional OES features.

A. Determination of key OES wavelengths using PCA

In this section, an ICP-RIE process for getting a higher aspect ratio of sapphire nanostructures is investigated.³¹ In this process, silicon nitride is used as the mask material and chlorine (Cl_2) is used as the active gas for etching the sapphire substrate. The processing parameters were pressure of 8 mTorr, the Cl_2 flow rate of 30 SCCM, ICP power of 1500 W, and radio frequency (RF) power of 200 W. The OES data are acquired using a fiber spectrophotometer (Model USB4000, Ocean Optics, Inc.) at the viewing port of the plasma chamber.

The original broadband OES data at 1, 3, and 5 min are plotted in Fig. 5(a). Please note that it is difficult to observe differences through direct comparison of raw OES data. However, after applying PCA to the OES readings, the spectral peaks can be clearly identified by plotting the square of each loading in the first principal component, as shown in Fig. 5(b). Further confirmation of chemical species corresponding to those peaks is evident in the fact that those peaks indeed correspond to critical elements and species during the sapphire etching process, namely, Si (252 and 288 nm), AlCl (261 nm), SiCl (281 and 283 nm), Al (396 nm), Cl (254, 726, 741, and 755 nm), and N_2 (775 nm). This study demonstrates that PCA can capture where major variations are in the

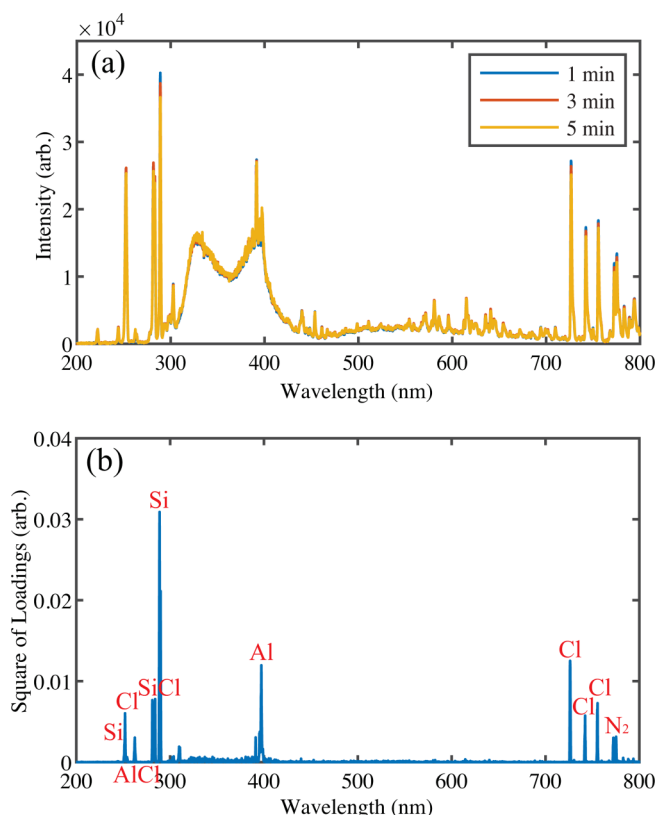


FIG. 5. Sapphire etching using Cl_2 plasma. (a) Raw OES data at 1, 3, and 5 min. (b) Square of loadings after PCA and critical chemical species.

original OES data, even though some parts of the broadband signals may be contaminated by noise.

B. VM modeling of critical dimensions in RIE

The key wavelengths identified by PCA can serve as inputs for VM models that estimate outgoing product quality in etching processes. To demonstrate the proposed method, in this section, we build VM models for predicting critical dimensions of silicon nanostructures obtained through a pattern transfer process that consists of four RIE steps. In this experiment, silicon wafers are coated with a photoresist layer, an antireflective coating (ARC) layer, and a nitride mask layer, as shown in Fig. 6. The first three etch steps consist of the RIE of the ARC, the nitride mask, and the silicon substrate, respectively. The last etching step is so-called ashing, which is the plasma etching of the remaining residuals above the silicon substrate. The first and the second RIE steps are performed in the fixed etch-time mode, while the third and fourth steps are conducted in the endpoint-detection mode. Key wavelength selection, integration band formation, and statistics-based and dynamics-inspired feature extraction are performed, respectively, for all four different RIE steps.

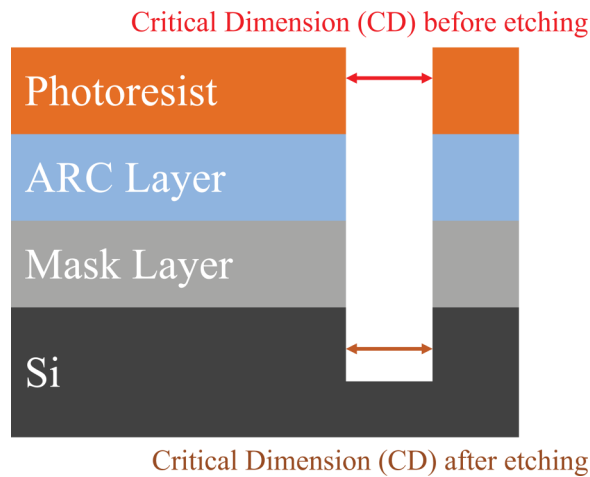


FIG. 6. Silicon wafer with three layers, namely, a photoresist layer, an antireflective coating layer, and a nitride mask layer, for the RIE process in this study.

Two critical dimension measurements, defined as the line-width of the pattern, are taken before and after the RIE process using critical dimension scanning electron microscopes (CD-SEM). The critical dimension before the RIE process is measured at the photoresist layer, and the critical dimension after the RIE process is measured at the silicon substrate. The same measurement conditions and edge detection procedures are used by the CD-SEM to measure the critical dimension automatically and consistently on the wafer. However, critical dimension measurements are taken on multiple CD-SEMs using the same recipe for different wafers, and this will lead to a measurement bias because of inherent differences between individual machines. Therefore, the identifier of each specific CD-SEM is considered as a non-numerical feature in our VM model.

The data analyzed in this section correspond to one recipe and were collected from one chamber of one RIE tool within one single maintenance cycle in 2 months. As a result, a total of 151 wafers with metrology measurements are available. The collected

data can be divided into three categories, namely, OES data, system data related to the plasma etcher, and data related to previous processes. OES data were collected with the spectral resolution of 2048 wavelengths between 200 and 900 nm, and the sampling rate is 10 Hz. The plasma etcher-related system data include tool settings, sensor readings (e.g., RF hours and chamber temperature), non-numerical data (e.g., process sequence number), the measurement data (critical dimension after etching and CD-SEM name), and process control parameters. Data related to previous processes contain the measurement data (critical dimension before etching and CD-SEM name) and non-numerical data (e.g., names of tools of previous processes). Except for OES data, which will be further processed to extract features based on the proposed method and traditional approach, the other two types of data are directly used to build VM models.

To evaluate the novel proposed technique in this work, the result of the VM models based on statistics-based and dynamics-inspired features is compared with VM models based on traditional features. Table I lists the newly proposed features in this work and the traditional feature set. In the transient states, eight types of features in the raw data and the gradient of the raw data are extracted. In the steady states, ten types of features in the raw data and the gradient of the raw data are extracted. All newly proposed features are automatically extracted from OES data as described previously. In contrast, traditional features are typically obtained by identifying key wavelengths and relevant feature types based on user experience and domain knowledge. Those features were identified through consultations with our industrial collaborators.

Different VM modeling methods are also considered. Specifically, in this article, we compared prediction performances of the PLSs-based VM and support vector regression (SVR)-based VM approaches since both those techniques are frequently used in industry today. Given that PLS and SVR methods are not compatible with missing input values, missing values are replaced by the mean of the available values of the corresponding features. One-hot encoding³² is used for processing non-numerical data, while standard 5-fold cross-validation³³ is applied to avoid overfitting.

Two metrics are used for evaluating the performance, namely, the coefficient of determination (R^2) and the root mean square error (RMSE). R^2 is the proportion of the variance in the

TABLE I. List of newly proposed features and traditional features.

Statistics-based and dynamics-inspired features				
Steady state		Transient		
Raw data	The gradient of the raw data	Raw data	The gradient of the raw data	Traditional features
Median	Median	Amplitude	Maximum	Mean, maximum, and minimum of several specific regions in six key wavelengths (predefined by the user)
Duration	Standard deviation	Duration	Minimum	
Standard deviation	deviation	Rise time		
Pulse counts	Minimum	Pre overshooting		
Amplitude	Maximum	Post overshooting		
	Range	Area under curve		

TABLE II. Results for VM modeling of critical dimension after etching (the third step).

	PLS#1 (newly proposed OES features + system data)	PLS#2 (traditional OES features + system data)	PLS#3 (without OES)
R ²	0.35	0.13	−0.0132
RMSE	0.113	0.131	0.141

dependent variable (e.g., critical dimension after etching) that is able to be predicted from the independent variables (e.g., identified features), which indicates better performance when it is close to one. The RMSE is the standard deviation of the prediction errors, which, obviously, indicate better performance when it is close to zero. Below are the mathematical expressions of these metrics,

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \frac{1}{n} \sum_{i=1}^n y_i)^2}{\sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (9)$$

where n is the wafer number, y_i is the measurement of the critical dimension after etching for i th wafer, and \hat{y}_i is the predicted value of the VM model for that same wafer.

For the results of key wavelength selection by PCA, only loadings in the first principal component are used because the explained variance ratio of the first principal components happened to be larger than 0.95. In essence, this indicates that the major variation in the original OES data can be represented using only using the first principal component, while the remaining components seem to be dominated by noise and environmental variations.

Since the critical dimension of the structure is measured at the silicon layer, the corresponding third etching step is considered as a key step for the whole RIE process based on consultations with our industrial collaborators. Therefore, PLS VM models are built based on the OES data at the third step and other system data. The results shown in Table II demonstrate that the R² of the PLS#1 model, which is built by the proposed features and other system data, is 0.35. PLS#1 outperforms the PLS#2 model, which is built by traditional features and other system data, by 170%. It can also

TABLE III. Performance comparison between two methods for forming integration bands.

	Method 1 (continuous)	Method 2 (peaks and valleys)
R ²	0.57	0.65
RMSE	0.088	0.08

be observed that including the proposed features in the VM model also reduces the RMSE by around 14%. PLS#3 represents the result where OES data are not used and performs the worst with R² close to 0, since it lacks important information for the VM modeling. As expected, without OES data, the VM models cannot predict the critical dimension after etching.

We also examined the performance of the two methods regarding the selection of the integration band width for PLS-based VM models of the entire four-step RIE process, as presented in Table III. The results show that by further dividing the set of continuous key wavelengths into smaller subsets based on the peaks and valleys of their loading values, R² can be enhanced from 0.57 to 0.65, while the RMSE reduces from 0.088 to 0.08. This is probably caused by the fact that the width of the integration band has informational value in the sense that if the width is too wide, the important information becomes hidden since more unnecessary or uncorrelated information and noise get included.

Finally, the PLS-based and SVR-based VM models for the entire four-step RIE process are presented in Table IV. The results show that regardless of whether one uses PLS- or SVR-based VM, considering all steps of the RIE process leads to dramatic improvements of VM performance, compared to what is obtained with building VM models when only the third RIE step is involved. This indicates that although the critical dimension is measured after etching into the silicon substrate, the process conditions of other layers also have a significant impact on the critical dimension. Furthermore, the results indicate that even though they are not as accurate as the VM models built by the newly proposed features, VM models based on traditional features can also be improved when SVR is applied in comparison to PLS-based VM. The cause of this phenomenon is likely the capability of SVR models to model nonlinearities in the data, which does not exist when PLS modeling is used. Last but not least, combination of statistics-based and dynamics-inspired features always performs better than the traditional feature set. This implies that the statistics-based and dynamics-inspired feature set possesses novel and more useful information for predicting the critical dimension after etching.

TABLE IV. Results for VM modeling of critical dimension after etching (entire steps).

Feature Set	PLS			SVR		
	Newly proposed	Traditional	Without OES	Newly proposed	Traditional	Without OES
R ²	0.65	0.4	−0.015	0.62	0.49	0.03
RMSE	0.08	0.104	0.136	0.08	0.1	0.132

IV. CONCLUSIONS AND FUTURE WORK

This article presents an approach for automatic extraction of informative features out of OES data for VM modeling of critical dimensions after RIE processes. The proposed method is based on key wavelength selection using PCA, formation of integration bands based on two different techniques to decide the width of the integration band, and extraction of statistics-based and dynamics-inspired features to build the VM model. The results demonstrate that PLS-based and SVR-based VM models have been successfully constructed for the prediction of critical dimension after etching the photoresist pattern into silicon. Including the entire four-step RIE process and using peaks and valleys in loading values to determine the width of the integration band can improve the performances of VM models. In addition, compared to VM models based on traditional features, the results indicate that our proposed method improves the R^2 from 0.4 to 0.65, by as much as 62.5%, and reduces the RMSE from 0.104 to 0.08, an improvement of 23.1%. These bring strong support to the statistics-based and dynamics-inspired feature set that can successfully represent the underlying dynamics of OES signals and retain more useful information for VM modeling, which does not exist in the traditional feature set. Also, both studies show that the proposed method is capable to deal with different devices, processes, and recipes, which means the VM model based on newly proposed features is robust that can be applied to different RIE processes independent of the etch chemistry.

Future research that could augment this work should include further enhancement of VM performance through exploration and testing of the proposed feature extraction and VM modeling approaches across different maintenance cycles is necessary. However, in the long run, some occasionally segments might add to the original data, and those might affect the results of feature alignment. To address this challenge, a similarity analysis, such as dynamic time warping,³⁴ can be introduced when considering data across different maintenance cycles, which will be explored as a part of future work. Since the proposed VM approach is highly explainable, it can potentially provide further information for the process control-based analysis of contributions each of those features makes to the final VM output, which could be accomplished using, e.g., variable importance in projection scores.³⁵ This information could be coupled with VM-based estimates of outgoing product quality to enable the automatic wafer-to-wafer process control, which would be much more responsive than the currently available lot-to-lot process control. Finally, large-scale implementation, such as different tools, devices, and recipes, of the newly proposed methods remains a very valuable objective.

ACKNOWLEDGMENTS

This work was supported in-part by the National Science Foundation (NSF) under Grant Nos. CMMI#1552424 and EEC-1160494. Work was partly done at the Texas Nanofabrication Facility supported by NSF Grant No. NNCI-2025227. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

DATA AVAILABILITY

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

REFERENCES

- ¹P. Chen, S. Wu, J. Lin, F. Ko, H. Lo, J. Wang, C. H. Yu, and M. S. Liang, *ISSM 2005 IEEE International Symposium on Semiconductor Manufacturing, 2005 San Jose, CA, 13-15 September 2005* (IEEE, San Jose, CA, 2005), pp. 155–157.
- ²I. P. Herman, *Annu. Rev. Phys. Chem.* **54**, 277 (2003).
- ³H. H. Yue, S. J. Qin, J. Wiseman, and A. Toprac, *J. Vac. Sci. Technol. A* **19**, 66 (2001).
- ⁴D. A. White, B. E. Goodlin, A. E. Gower, D. S. Boning, H. Chen, H. H. Sawin, and T. J. Dalton, *IEEE Trans. Semicond. Manuf.* **13**, 193 (2000).
- ⁵B. E. Goodlin, "Multivariate endpoint detection of plasma etching processes," Ph.D. thesis (Massachusetts Institute of Technology, 2002).
- ⁶R. Chen, H. Huang, C. J. Spanos, and M. Gatto, *J. Vac. Sci. Technol. A* **14**, 1901 (1996).
- ⁷H.-T. Noh, D.-I. Kim, and S.-S. Han, *2015 China Semiconductor Technology International Conference, Shanghai, China, 15-16 March 2015* (IEEE, New York, 2015), p. 1.
- ⁸K. Han, E. S. Yoon, J. Lee, H. Chae, K. H. Han, and K. J. Park, *Ind. Eng. Chem. Res.* **47**, 3907 (2008).
- ⁹S.-I. Jeon, S.-G. Kim, Y.-S. Han, S.-H. Shin, and S.-S. Han, *ECS Trans.* **44**, 1087 (2012).
- ¹⁰E. A. Rietman, R. C. Frye, E. R. Lory, and T. R. Harry, *J. Vac. Sci. Technol. B* **11**, 1314 (1993).
- ¹¹R. L. Allen, R. Moore, and M. Whelan, *J. Vac. Sci. Technol. B* **14**, 498 (1996).
- ¹²S. Limanond, J. Si, and Y.-L. Tseng, *J. Vac. Sci. Technol. B* **16**, 2707 (1998).
- ¹³H. Jang, H. Lee, H. Lee, C.-K. Kim, and H. Chae, *IEEE Trans. Semicond. Manuf.* **30**, 17 (2017).
- ¹⁴B. Kim, S. Im, and G. Yoo, *Electronics* **10**, 49 (2021).
- ¹⁵S. J. Hong, G. S. May, and D.-C. Park, *IEEE Trans. Semicond. Manuf.* **16**, 598 (2003).
- ¹⁶T. Hirai and M. Kano, *IEEE Trans. Semicond. Manuf.* **28**, 137 (2015).
- ¹⁷E. Ragnoli, S. McLoone, S. Lynn, J. Ringwood, and N. Macgearailt, *2009 IEEE/SEMI Advanced Semiconductor Manufacturing Conference, Berlin, Germany, 10-12 May 2009* (IEEE, New York, 2009), pp. 106–111.
- ¹⁸D. Zeng and C. J. Spanos, *IEEE Trans. Semicond. Manuf.* **22**, 419 (2009).
- ¹⁹C. Park and S. B. Kim, *J. Process Control* **42**, 51 (2016).
- ²⁰S. Park, S. Jeong, Y. Jang, S. Ryu, H.-J. Roh, and G.-H. Kim, *IEEE Trans. Semicond. Manuf.* **28**, 241 (2015).
- ²¹M. Maggipinto, C. Masiero, A. Beghi, and G. A. Susto, *Proc. Manuf.* **17**, 126 (2018).
- ²²M. Maggipinto, M. Terzi, C. Masiero, A. Beghi, and G. A. Susto, *IEEE Trans. Semicond. Manuf.* **31**, 376 (2018).
- ²³Y. Jang, H.-J. Roh, S. Park, S. Jeong, S. Ryu, J.-W. Kwon, N.-K. Kim, and G.-H. Kim, *Curr. Appl. Phys.* **19**, 1068 (2019).
- ²⁴S. Park et al., *Plasma Phys. Control. Fusion* **61**, 014032 (2019).
- ²⁵S. Park et al., *Phys. Plasmas* **27**, 083507 (2020).
- ²⁶J.-W. Kwon, S. Ryu, J. Park, H. Lee, Y. Jang, S. Park, and G.-H. Kim, *Materials* **14**, 3005 (2021).
- ²⁷A. A. Ul Haq and D. Djurdjanovic, *J. Ind. Inf. Integr.* **13**, 22 (2019).
- ²⁸I. T. Jolliffe and J. Cadima, *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **374**, 20150202 (2016).
- ²⁹J. E. Jackson, *A User's Guide to Principal Components* (John Wiley & Sons, Hoboken, NJ, 2005).

³⁰IEEE Std 181TM—2011, *IEEE Standard for Transitions, Pulses, and Related Waveforms* IEEE Instrumentation and Measurement Society (IEEE, New York, 2011).

³¹Y.-A. Chen, I.-T. Chen, and C.-H. Chang, *J. Vac. Sci. Technol. B* **37**, 061606 (2019).

³²K. Potdar, T. Pardawala, and C. Pai, *Int. J. Comput. Appl.* **175**, 7 (2017).

³³M. Stone, *J. R. Stat. Soc. Ser. B Methodol.* **36**, 111 (1974).

³⁴M. Müller, *Information Retrieval for Music and Motion* (Springer, Berlin, 2007), p. 69.

³⁵S. Wold, E. Johansson, and M. Cocchi, *3D QSAR in Drug Design: Theory Methods and Applications* (KLUWER ESCOM Science, Germania, 1993), p. 523.