Industrial Engineering
& Management Systems
Vol 14, No 4, December 2015, pp.392-400
ISSN 1598-7248 | EISSN 2234-6473 |
http://dx.doi.org/10.7232/iems.2015.14.4.392
© 2015 KIIE

Quantitative Analysis for Plasma Etch Modeling Using Optical Emission Spectroscopy: Prediction of Plasma Etch Responses

Young-Seon Jeong*
Department of Industrial Engineering, Chonnam National University, Gwangju, Korea

Sangheum Hwang
Lunit Inc., Seoul, Korea

Young-Don Ko
Vent-up Inc., Washington, D.C., USA

(Received: September 9, 2015 / Revised: December 16, 2015 / Accepted: December 18, 2015)

ABSTRACT

Monitoring of plasma etch processes for fault detection is one of the hallmark procedures in semiconductor manufacturing. Optical emission spectroscopy (OES) has been considered as a gold standard for modeling plasma etching processes for on-line diagnosis and monitoring. However, statistical quantitative methods for processing the OES data are still lacking. There is an urgent need for a statistical quantitative method to deal with high-dimensional OES data for improving the quality of etched wafers. Therefore, we propose a robust relevance vector machine (RRVM) for regression with statistical quantitative features for modeling etch rate and uniformity in plasma etch processes by using OES data. For effectively dealing with the OES data complexity, we identify seven statistical features for extraction from raw OES data by reducing the data dimensionality. The experimental results demonstrate that the proposed approach is more suitable for high-accuracy monitoring of plasma etch responses obtained from OES.

Keywords: Plasma Process Modeling, Statistical Quantitative Feature, Semiconductor Manufacturing, Robust Relevance Vector Machine, Optical Emission Spectroscopy

* Corresponding Author, E-mail: young.jeong@jnu.ac.kr

1. INTRODUCTION

Several processing steps have been developed in semiconductor manufacturing industry for increasing the manufacturing productivity and improving the performance; examples include deposition, etching, ashing, polishing, annealing, and patterning (Hastie *et al.*, 2001; Kim *et al.*, 2005; Kolari, 2008; Sugawara, 1998). Among these steps, etching plays an important role in device fabrication, because it can be used for controlling the size and reducing the dimensionality of an integrated circuit system. In this step, a plasma-aided process is potentially addressed for modifying the surfaces of the materials and for solving a critical problem of device dimensionality; during this process low pressure, non-equilibrium, and partially ionized gas discharge plasmas are effectively characterized (Hitchon, 1999; Ko *et al.*, 2010). However, physical properties of either the device or the wafer may not be accurately controlled owing to irregular factors or unexpected events in the system during the process operation. This problem critically affects the device's performance characteristics, such as the

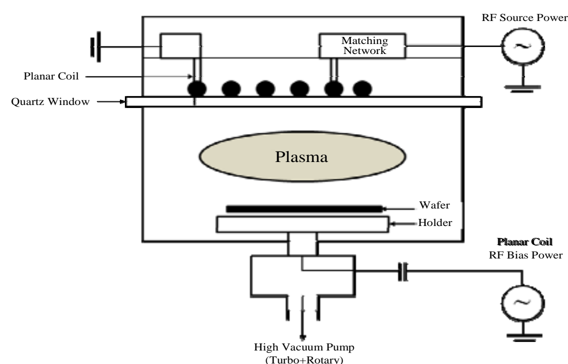


Figure 1. Schematic of the plasma etching system.

power consumption, the extent of leakage current, device reliability, device efficiency, and physical properties. Therefore, the processes of chip and/or device manufacturing should be more carefully monitored and precisely predicted in timely manner. This concern becomes a challenging task in state-of-the-art advanced process controlling. A schematic of the plasma etching system is shown in Figure 1.

For diagnostics of plasma processes, relevant examples are provided by the modeling of X-ray photoelectron spectroscopy data (Kim and Kim, 2007; Kim and Park, 2006) and the prediction of etch properties by learning optical emission spectroscopy (OES) (Hong and May, 2005; Hong *et al*., 2003; Kim and Kwon, 2008; Ko *et al*., 2010). Instead of dealing with full sets of OES data, reduced sets that were generated by using wavelets and principal component analysis (PCA) were used for modeling (Bayissa *et al*., 2008; Kim and Kim, 2007; Kim and Kwon, 2008; Kim and Park, 2006). Plasma etch models were devised by using statistical regression techniques in conjunction with principal component regression (PCR) and neural network methods (Hong and May, 2005; Hong *et al*., 2003; Kim *et al*., 2005; Kim and Kim, 2005). In addition, Zhang *et al*. (2006) proposed a continuous wavelet transform methodology for improving the axial resolution of acoustic micro-imaging, in the context of advanced microelectronic packaging. Ko *et al*. (2008) developed a kernel-based regression model by using wavelet-compressed OES data. They used a wavelet transform with vertical energy thresholding (VET) shrinkage procedures, which were used for reducing the dimensionality of multiple functional data. However, statistical quantitative methods for processing high-dimensional OES data are still lacking.

To address this issue, in this paper, we propose a novel procedure combining a robust supervised learning algorithm with statistical quantitative features extracted from original OES data for modeling plasma processes. To effectively address the high dimensionality of the OES data, seven statistical features are identified here as carrying significant input information, and robust relevance vector machine (RRVM) is employed for monitoring of plasma etch processes. The proposed approach in this paper is likely to make important contributions to: (1) extracting informative statistical features from raw OES data and (2) building a robust regression model for accurately predicting the etch responses during plasma etch processes. To our best knowledge, there are little literatures about the use of statistical quantitative features to model plasma responses in semiconductor manufacturing.

The rest of this paper is organized as follows. In Section 2, we describe the experimental details for obtaining the OES datasets. In Section 3, we present the proposed statistical quantitative approach combined with a RRVM for monitoring of etch processes. Computational results and conclusions are given in Sections 4 and 5, respectively.

2. EXPERIMENTAL DETAILS

Plasma etching was performed by using a magnetically enhanced chemical vapor deposition system. A schematic of plasma equipment was presented as well as the fabrication of the test patterns. Therefore, they are briefly described here. The test patterns were fabricated on (100)-Si substrates. Following a deionized water rinse, a 900-nm-thick oxide film was deposited by using a plasma-enhanced chemical vapor deposition system. A 1.02-nm-thick photoresist film was then spin-coated, followed by photoresist patterning. During the etching, OES was used for collecting in-situ radical data for wavelengths ranging from 227.6 to 791.6 nm, by using a sample interval of 0.3 nm. Examples of OES data and the corresponding etch responses are shown in Figure 2. Face-centered box Wilson experimental design was used in this study, and the scenarios are shown in Table 1.

A 2^{4-1} fractional factorial experiment was performed, and the corresponding eight experiments are shown in

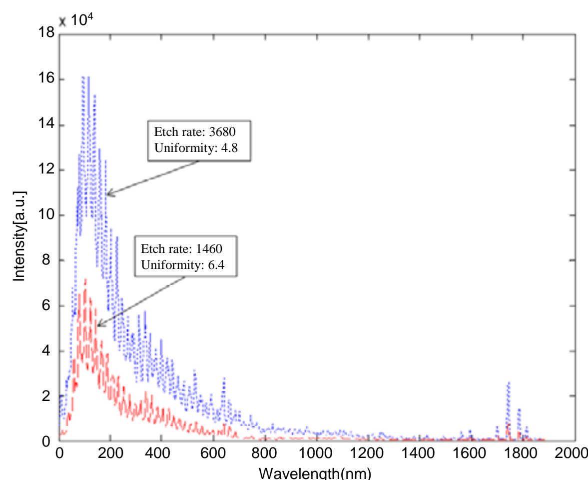


Figure 2. Example of OES data and the corresponding etch responses

Table 1. 2^{4-1} fractional factorial design

Run	CHF_3	CF_4	RF power	Pressure
1	20	10	300	50
2	20	10	800	200
3	20	40	300	200
4	20	40	800	50
5	80	10	300	200
6	80	10	800	50
7	80	40	300	50
8	80	40	800	200
9*	50	25	550	125

* indicates the central point in the training data set.

Table 1. A training data set was combined with one central design point. Face-centered points were generated and used as a testing data set. The data from one of the nine experiments has not been collected owing to an inappropriate parameters setting. The etch rate was calculated by dividing the etched film thickness by the etch time of 60s. To assess the non-uniformity, etch rates were measured at five points on the wafer. Four points were equidistant from the central point. The metric for measuring the non-uniformity was defined as:

$$Nonuniformity = \frac{R_{\max} - R_{\min}}{R_{avg}} \times 100\,(\%) \tag{1}$$

where R_{max} and R_{min} indicate the maximal and minimal etch rates, respectively. R_{avg} is the average etch rate, which is obtained simply by averaging over five values of etch rates.

3. PROPOSED QUANTITATIVE STATISTICAL MODEL

The quantitative statistical approach is based on characterizing the statistical features of OES data and applying robust relevance vector machine (RRVM) for predicting the etch responses from which the OES data were drawn.

3.1 Statistical Features

In this subsection, a novel approach for predicting plasma etching responses based on the statistical features of OES data is proposed. Denoting by $y_j \in \mathbf{Y} = \{y_1, y_2, \cdots, y_n\}$ the raw OES data vector of length n, the following statistical features can be investigated:

1. Curve length (CL): The approximate length of a raw OES data vector.

$$S_{CL} = \sum_{j=1}^{n-1} \left| y_{j+1} - y_j \right| \tag{2}$$

2. Peaks: The number of local maxima.

$$S_{Peak} = \frac{1}{2}\sum_{j=1}^{n-2} \max\left\{0,\ \operatorname{sgn}\left[y_{j+2} - y_{j+1}\right] - \operatorname{sgn}\left[y_{j+1} - y_j\right]\right\} \tag{3}$$

3. Root mean square (RMS): Describes the magnitude of the OES data fluctuation.

$$S_{RMS} = \sqrt{\frac{1}{n}\sum_{j=1}^{n} y_j^2} \tag{4}$$

4. Average nonlinear energy (ANE): Represents the extent of nonlinearity in the raw signal.

$$S_{ANE} = \frac{1}{n-2}\sum_{j=1}^{n-1}\left(y_j^2 - y_{j-1}\, y_{j+1}\right) \tag{5}$$

5. Standard deviation (SD): Measures the extent of variation or dispersion from the average.

$$S_{SD} = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n-1}\left(y_j - \bar{Y}\right)^2} \tag{6}$$

where $\bar{Y} = \frac{1}{n}\sum_{j=1}^{n} y_j$ is the overall mean of the raw signal.

6. Kurtosis: The fourth moment of a distribution, capturing the size of the distribution's tails.

$$S_K = \frac{1}{n}\frac{\sum_{j=1}^{n}\left(y_j - \bar{Y}\right)^4}{\left(S_{SD}\right)^4} \tag{7}$$

7. Skewness: Characterizes the extent of asymmetry of a distribution around its mean.

$$S_{Ske} = \frac{1}{n}\frac{\sum_{j=1}^{n}\left(y_j - \bar{Y}\right)^3}{\left(S_{SD}\right)^3} \tag{8}$$

The selected statistical features have been explored to be useful to capture the underlying nature of specific signals in diverse applications (Guo *et al.*, 2005; Picard *et al.*, 2001). In addition, the advantage of using these statistical features is relatively small computational burden associated with data transformation from raw OES data, making these features advantageous for developing online monitoring procedures.

3.2 Robust Relevance Vector Machine for Regression

Next, we combine the proposed statistical features

into a robust relevance vector machine (RRVM) for predicting the etch responses, such as the etch rate and non-uniformity.

Relevance vector machine for regression (RVR) is a Bayesian sparse kernel technique which shares many of the characteristics of the SVR (Bishop, 2006). The structure of the RVR is described as follows:

$$y_i = f(\mathbf{x}; \boldsymbol{\beta}) + \varepsilon_i \tag{9}$$

where ε_i's are independent samples from $p(\varepsilon_i \mid \gamma) \sim N(0, \gamma^{-1})$. By using the given the training data, the regression model $f(\mathbf{x}; \boldsymbol{\beta})$ should be inferred to have a good generalization capability. The regression model can be expressed as

$$f(\mathbf{x}; \boldsymbol{\beta}) = \sum_{i=1}^{n} \beta_i \phi_i(\mathbf{x}) + \beta_0 \tag{10}$$

where the vector β_i is the weight parameter of the model, β_0 is a bias, and $\boldsymbol{\phi}(\mathbf{x}) = (1, \phi_1(\mathbf{x}), \ldots, \phi_N(\mathbf{x}))^T$ is the basis function vector.

The limitation of a conventional RVR is the sensitivity of outliers. In other words, if the training data set is contaminated by some outlying observations, RVR could give unreliable and inaccurate results. To deal with this problem, Hwang *et al*. (2014) recently proposed a robust relevance vector machine for regression (RRVM) to reduce the effect of outliers by using a weight strategy and utilize a variational inference method to estimate the posterior distribution over model parameters. To overcome an outlier problem, they used the noise ε_i's have heteroscedastic variance $p(\varepsilon_i \mid w_i, \gamma) \sim N(0, (w_i\gamma)^{-1})$. To obtain a robust model coefficients $\boldsymbol{\beta}$, by assuming independently distributed data, a weighting strategy can be employed to the likelihood function as

$$p(\mathbf{y} \mid \boldsymbol{\beta}, \mathbf{w}, \gamma) = (2\pi)^{-N/2} \prod_{i=1}^{N} (w_i\gamma)^{1/2} \exp\left[-\frac{w_i\gamma}{2}\left\{y_i - \boldsymbol{\beta}^T\boldsymbol{\phi}(\mathbf{x}_i)\right\}^2\right] \tag{11}$$

where $\mathbf{y} = (y_1, \ldots, y_N)^T$ and $\mathbf{w} = (w_1, \ldots, w_N)^T$. We approximate the posterior distribution over model coefficients $\boldsymbol{\beta}$ using a variational inference method under the following prior distributions over the model parameters:

$$p(\boldsymbol{\beta} \mid \boldsymbol{\alpha}) = N(\boldsymbol{\beta} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) \tag{12}$$

$$p(\boldsymbol{\alpha} \mid a, b) = \prod_{i=0}^{N} Gamma\left(\alpha_i \mid \tilde{a}, \tilde{b}_i\right) \tag{13}$$

$$p(\mathbf{w} \mid c, d) = \prod_{i=1}^{N} Gamma\left(w_i \mid \tilde{c}, \tilde{d}_i\right) \tag{14}$$

where

$$\boldsymbol{\Sigma} = \left(\boldsymbol{\Phi}^T\mathbf{W}\boldsymbol{\Phi} + \mathbf{A}\right)^{-1} \tag{15}$$

$$\boldsymbol{\mu} = \boldsymbol{\Sigma}\boldsymbol{\Phi}^T\mathbf{W}\mathbf{y} \tag{16}$$

$$\tilde{a} = a + \frac{1}{2}, \quad \tilde{b}_i = b + \frac{\mathrm{E}\left(\beta_i^2\right)}{2} \tag{17}$$

$$\tilde{c} = c + \frac{1}{2}, \quad \tilde{d}_i = d + \frac{\gamma}{2}\mathrm{E}\left\{\left(y_i - \boldsymbol{\beta}^T\boldsymbol{\phi}(\mathbf{x}_i)\right)^2\right\} \tag{18}$$

where

$$\gamma^{-1} = \frac{\sum_{i=1}^{N}\mathrm{E}(w_i)\{(y_i - \mathrm{E}(\boldsymbol{\beta})^T\boldsymbol{\phi}(\mathbf{x}_i))^2 + \boldsymbol{\phi}(\mathbf{x}_i)^T\boldsymbol{\Sigma}\boldsymbol{\phi}(\mathbf{x}_i)\}}{N - tr(\mathbf{I}_{N+1} - \mathbf{A}\boldsymbol{\Sigma})} \tag{19}$$

Here, $\boldsymbol{\Phi}$ represents the matrix whose row vectors are $\boldsymbol{\phi}(\mathbf{x}_1)^T, \ldots, \boldsymbol{\phi}(\mathbf{x}_N)^T$, and $\mathbf{W}$ and A denote a diagonal matrix with $\mathrm{E}(w_i)\gamma$ and $\mathrm{E}(\alpha_i)$ as the *i*-th diagonal element, respectively. To train RRVM by using training dataset, we randomly select initial values for *a*, *b*, *c*, *d* and γ. In this paper, we set to $a = b = 10^{-5}$ and $c = d = \gamma = 1$. Next, by using Eq. (15)-Eq. (18), update the hyperparameters $\boldsymbol{\Sigma}$, $\boldsymbol{\mu}$, $\tilde{a}$, $\tilde{b}_i$, $\tilde{c}$ and $\tilde{d}_i$ and re-estimate γ in Eq. (19). The posterior distributions in Eq. (12)-Eq. (14) can be characterized with the updated hyperparameter values. Finally, compute the lower bound value $\mathcal{L}[Q(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{w})]$, which is defined in Eq. (20) and repeat those works until the change of $\mathcal{L}[Q(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{w})]$ is smaller than a predefined threshold.

$$\begin{aligned}\mathcal{L}[Q(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{w})] = {} & \mathrm{E}[\log p(\mathbf{y} \mid \boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{w}, \gamma)] + \mathrm{E}[\log p(\boldsymbol{\beta} \mid \boldsymbol{\alpha})] \\ & + \mathrm{E}[\log p(\boldsymbol{\alpha} \mid a, b)] + \mathrm{E}[\log p(\mathbf{w} \mid c, d)] \\ & - \mathrm{E}[\log Q_{\boldsymbol{\beta}}(\boldsymbol{\beta})] - \mathrm{E}[\log Q_{\boldsymbol{\alpha}}(\boldsymbol{\alpha})] - \mathrm{E}[\log Q_{\mathbf{w}}(\mathbf{w})]\end{aligned} \tag{20}$$

Given a new input $\mathbf{x}$, the predictive distribution over *y* can be evaluated using the following approximated distribution:

$$p(y \mid \mathbf{x}, \mathbf{y}, \gamma) = N(y \mid \boldsymbol{\mu}^T\boldsymbol{\phi}(\mathbf{x}), \{\mathrm{E}(w)\gamma\}^{-1} + \boldsymbol{\phi}(\mathbf{x})^T\boldsymbol{\Sigma}\boldsymbol{\phi}(\mathbf{x})) \tag{21}$$

where *w* denotes a weight corresponding to a new input $\mathbf{x}$(see the Hwang *et al*. (2014) for the detail derivatives).

4. EXPERIMENTAL RESULTS

The plasma etching process responses were examined by using a variety of modeling techniques: multiple linear regression (MLR), multilayer perceptron neural network (NNet), RVR, and RRVM.

MLR is widely employed prediction model in diverse applications, which assumes a linear relation between independent variables and dependent ones. MLR is simple and computationally efficient, but it has poor prediction results when the relation between variables is nonlinear (Hastie *et al*., 2001).

The architecture of NNet is composed as follows: seven neurons in the input layer, single hidden layer

with ten neurons and one output neuron. Tangent sigmoid function and linear transfer function are used for activation functions in the hidden and output layer. The prediction models were created using all fifteen OES data from plasma etch process, and for the limited samples, we used five-fold cross-validations. In addition, Gaussian kernels were used for the RVR and RRVM. A Gaussian kernel is defined as follows:

$$K(x_1, x_2) = \exp\left(\frac{\|x_1 - x_2\|}{2\sigma^2}\right) \quad (22)$$

where σ is the width parameter, which is optimized by the validation dataset. The parameter σ was varied and the values of $\sigma = 2^{-10}, 2^{-9}, \text{L}, 2^{9}, 2^{10}$ were obtained. This kernel parameter was optimized in terms of the prediction accuracy by using the validation dataset. By using the selected parameters, the prediction accuracy for the testing dataset was calculated.

To compare between the models, two performance measures are computed as follows: Root mean squared error (RMSE) and normalized root mean squared error (NRMSE).

$$\text{RMSE} = \sqrt{\frac{1}{n_t}\sum_{i=1}^{n_t}(y_i - \hat{y}_i)^2} \quad (23)$$

$$\text{NRMSE} = \frac{1}{m_t}\left(\sqrt{\frac{1}{n_t}\sum_{i=1}^{n_t}(y_i - \hat{y}_i)^2}\right) \quad (24)$$

where n_t is the number of samples, m_t is the mean of the observed value of samples, y_i is the i-th observed value, and $\hat{y}_i$ is the i-th predicted value. The smaller value of these performance measures represents the better predictive accuracy of the model.

Table 2 shows the comparative results of testing data in terms of RMSE and NRMSE in case of the etch rate response. As seen in Table 2, the proposed RRVM with statistical features shows better results than existing models in terms of average performance of five-cross validation data. Specifically, the RRVM yielded the average NRMSE of 0.300 while the MLR yielded the average error of 1.125. On the other hand, Table 3 shows the prediction errors of each model, addressing the non-uniformity. In Table 3, the MLR model exhibits large NRMSEs, 1.220 on average while RRVM has average NRMSE of 0.224. Therefore, the proposed RRVM regression model with statistical features demonstrates much better performance than the existing approaches.

In addition, the results of comparison with regression models with original features are shown in Table 4 and Table 5. In is shown in those tables that the accuracy of regression models with original features is worse than that of the regression models with the selected features because the number of original features is much bigger than the number of observations and contains much noises in original features.

In addition, Figure 3 and Figure 4 illustrates each prediction modeling results for the etch rate and uniformity, respectively. As evidenced by the smaller perfor-

Table 2. Summary of the accuracy results for etch rate with statistical features

Prediction models	Performance measures	CV dataset					Average
		1	2	3	4	5	
MLR	RMSE	893	2311	3481	1175	5352	2642
	NRMSE	0.312	1.051	1.336	0.445	2.482	1.125
NNet	RMSE	1075	1245	1395	1420	658	1159
	NRMSE	0.375	0.566	0.535	0.538	0.305	0.464
RVM	RMSE	865	1369	1721	401	409	953
	NRMSE	0.302	0.622	0.661	0.152	0.190	0.385
RRVM	RMSE	872	974	662	276	853	727
	NRMSE	0.304	0.443	0.254	0.105	0.396	0.300

Table 3. Summary of the accuracy results for non-uniformity with statistical features

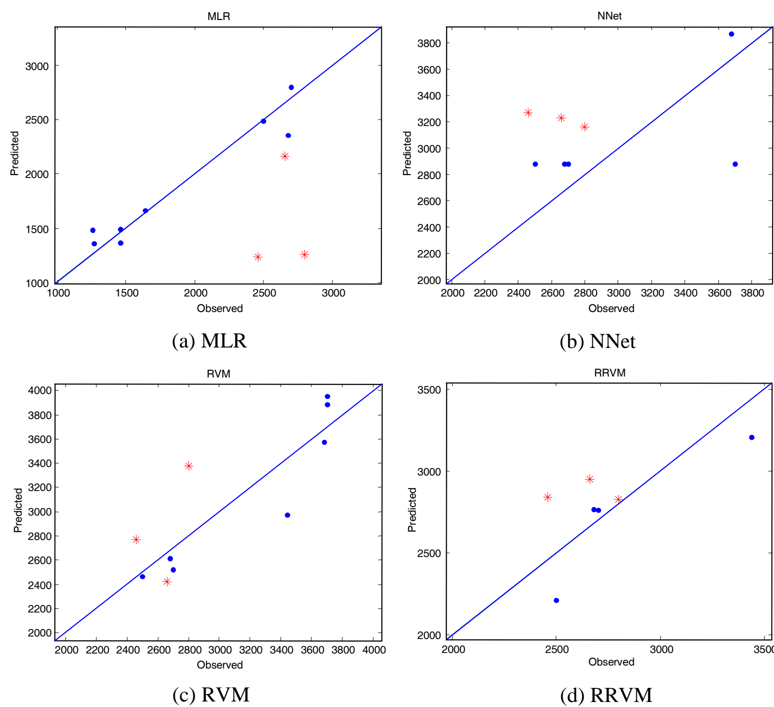
Prediction models	Performance measures	CV dataset					Average
		1	2	3	4	5	
MLR	RMSE	2.366	3.677	9.230	2.471	10.256	5.600
	NRMSE	0.399	0.919	1.978	0.608	2.200	1.220
NNet	RMSE	2.237	2.453	1.419	1.192	3.037	2.068
	NRMSE	0.377	0.613	0.304	0.293	0.651	0.448
RVM	RMSE	2.275	5.368	4.750	1.077	1.185	2.931
	NRMSE	0.383	1.342	1.018	0.265	0.254	0.652
RRVM	RMSE	1.983	1.269	0.660	0.618	0.811	1.068
	NRMSE	0.334	0.317	0.141	0.152	0.174	0.224

Table 4. Summary of the accuracy results for etch rate with original features

Prediction models	Performance measures	CV dataset					Average
		1	2	3	4	5	
MLR	RMSE	1048	2947	3974	2184	6847	3400
	NRMSE	0.366	1.340	1.525	0.827	3.175	1.717
NNet	RMSE	1097	1473	1408	1573	776	1265
	NRMSE	0.383	0.670	0.540	0.596	0.360	0.541
RVM	RMSE	903	1428	1726	573	521	1030
	NRMSE	0.315	0.649	0.662	0.217	0.242	0.442
RRVM	RMSE	912	1047	836	375	903	815
	NRMSE	0.318	0.476	0.321	0.142	0.419	0.339

Table 5. Summary of the accuracy results for non-uniformity with original features

Prediction models	Performance measures	CV dataset					Average
		1	2	3	4	5	
MLR	RMSE	3.738	4.037	9.974	3.937	12.387	6.815
	NRMSE	0.630	1.009	2.137	0.968	2.654	1.480
NNet	RMSE	2.573	2.980	2.473	1.976	4.583	2.917
	NRMSE	0.434	0.745	0.530	0.486	0.982	0.635
RVM	RMSE	2.475	5.837	3.274	1.384	1.387	2.876
	NRMSE	0.417	1.459	0.702	0.340	0.297	0.643
RRVM	RMSE	2.284	2.047	0.937	1.184	1.385	1.567
	NRMSE	0.385	0.512	0.201	0.291	0.297	0.337

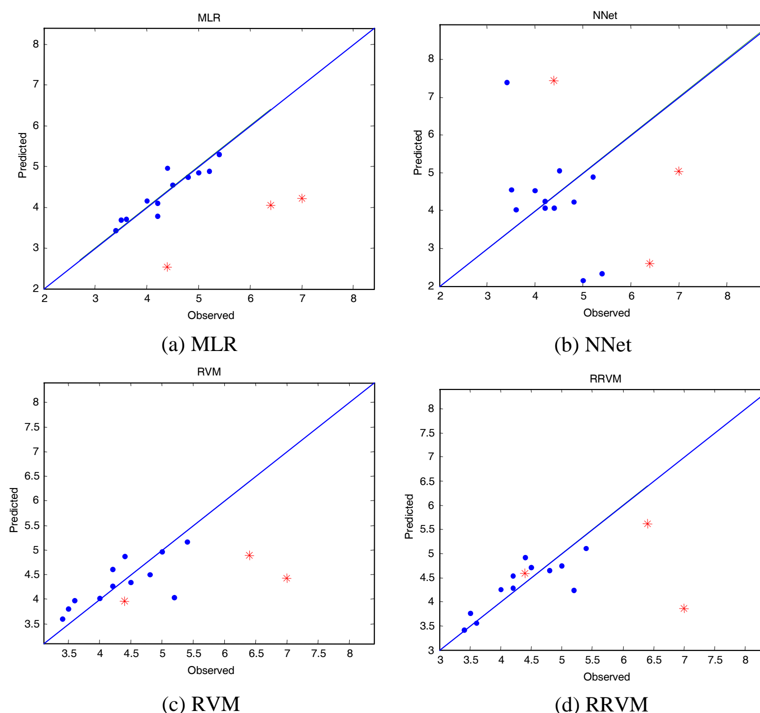


(a) MLR (b) NNet

(c) RVM (d) RRVM

Figure 3. Prediction modeling results for the etch rate.

mance measure, as shown in those figures, RRVM model presents the best prediction results for the testing data. Even though the prediction performance for the training data is similar to each model, RVM and RRVM model demonstrates better performance for the testing data compared with that of MLR and NNet mode.



(a) MLR (b) NNet

(c) RVM (d) RRVM

Figure 4. Prediction modeling results for the nonuniformity.

To explore the superiority of statistical features, we compared the prediction performance with popular existing dimensional reduction methods including linear and non-linear methods such as principal component analysis (PCA), partial least squares (PLS), kernel principal component analysis (KPCA), and kernel partial least squares (KPLS) (Rosipal and Trejo, 2001). For fair comparison, we used a RRVM model with seven principal components (PCs) and seven latent variables (LVs) because seven PCs in PCA and seven LVs in PLS can explain 96.1% variance and 97.5% variance of OES data, respectively. For KPCA and KPLS method, Gaussian kernel function is used. Figure 5 shows that even though the performance of non-linear methods is better than that of linear methods, the regression accuracy based on statistical features shows the best performance among other existing methods.

In summary, the experimental results validate our approach based on the proposed statistical features for effective characterization of OES data for the prediction of etch rate and non-uniformity.

5. CONCLUSIONS

In this paper we proposed a statistical quantitative method for modeling etch responses, such as the etch rate and uniformity, by using optical emission spectroscopy (OES) data in a plasma etch process. We used the experimental results for validating the effectiveness of the proposed RRVM with statistical features in dealing with OES data, which exhibit several characteristics, such as nonlinearity and high dimensionality. This suggests that the proposed procedure can be utilized for improving the etch process yield and the manufacturability of the overall process in the field of semiconductor engineering.

Additional applications of our statistical features-based approach that will be investigated in the future work include analysis of spectral data by taking into consideration the changes attributed to the atomic species, physical reactions, and process conditions in semiconductor manufacturing. Monitoring the optical emission of various emission lines generated by reaction products or reagents is an additional interesting topic for future research.

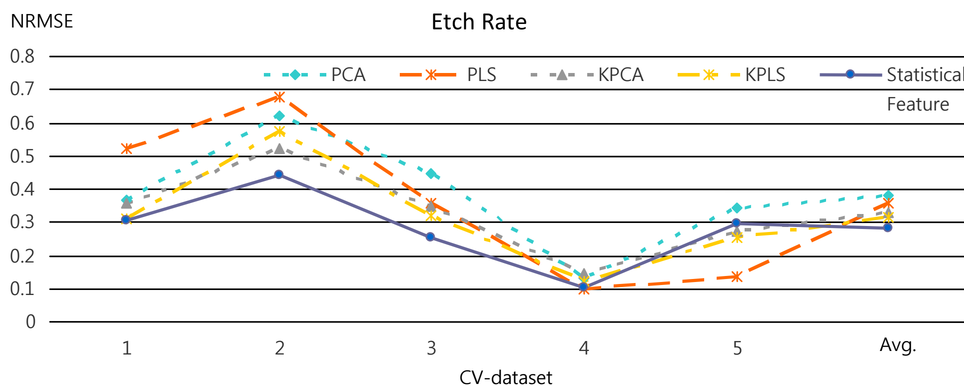
ACKNOWLEDGEMENTS

This study was financially supported by Chonnam National University (Grant no. 2014-0542, 2015-1829).

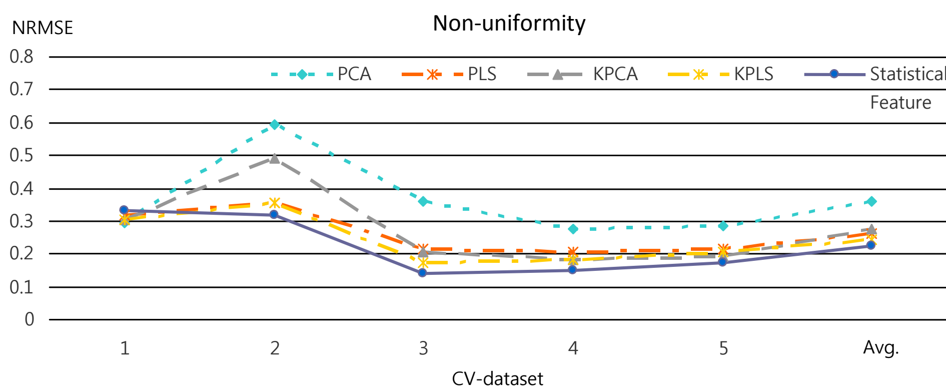
REFERENCES

Bayissa, W. L., Haritos, N., and Thelandersson, S. (2008), Vibration-based structural damage identification using wavelet transformation, *Mechanical Systems and Signal Processing*, **22**, 1194-1215.

Bishop, C. M. (2006), *Pattern recognition and machine learning*, Springer.



(a) Testing NRMSE of etch rate



(b) Testing NRMSE of non-uniformity

Figure 5. Summary of the NRMSEs for dimension reduction methods.

Guo, H., Jack, L. B.m and Nandi, A. K. (2005), Feature generation using genetic programming with application to fault classification, *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, **35**(1), 89-99.

Hastie, T., Tibshirani, R., and Friedman, J. (2001), *The elements of statistical learning*, New York: Springer.

Hitchon, W. N. G. (1999), *Plasma processing for semiconductor fabrication*, Cambridge University Press.

Hong, S. J. and May, G. S. (2005), Neural-network-based sensor fusion of optical emission and mass spectroscopy data for real-time fault detection in reactive ion etching, *IEEE Transactions on Industrial Electronics*, **52**(4), 1063-1072.

Hong, S. J., May, G. S., Park, D.-C. (2003), Neural network modeling of reactive ion etching using optical emission spectroscopy data, *IEEE Transactions on Semiconductor Manufacturing*, **16**(4), 598-608.

Hwang, S., Jeong, M. K., and Yum, B. J. (2014), Robust relevance vector machine with variational inference for improving virtual metrology accuracy, *IEEE Transactions on Semiconductor Manufacturing*, **27** (1), 83-94.

Kim, B., Bae, J. K., and Hong, W.-S. (2005), Plasma control using neural network and optical emission spectroscopy, *Journal of Vacuum Science and Technology A*, **23**(2), 355-358.

Kim, B. and Kim, S. (2005), Diagnosis of plasma proces-sing equipment using neural network recognition of wavelet-filtered impedance matching, *Microelectronic Engineering*, **82**, 44-52.

Kim, B. and Kim, W. S. (2007), Wavelet monitoring of spatial surface roughness for plasma diagnosis, *Microelectronic Engineering*, **84**, 2810-2816.

Kim, B., Kim, J., Lee, S. H., Park, J., and Lee, B. T. (2005), Plasma etching of silicon oxynitride in a low pressure C_2F_6 plasma, *Journal of Korean Physics Society*, **47**, 712-715.

Kim, B. and Kim, W. (2007), Partial X-ray photoelectron spectroscopy to constructing neural network model of plasma etching surface, *Microelectronic Engineering*, **84**, 584-589.

Kim, B. and Kwon, M. (2008), Optimization of PCA-applied in-situ spectroscopy data using neural network and genetic algorithm, *Applied Spectroscopy*, **62**(1), 73-77.

Kim, B. and Park, M. (2006), Prediction of surface roughness using X-ray photoelectron spectroscopy and

neural networks, *Applied Spectroscopy*, **60**(10), 1192-1197.

Ko, Y.-D., Jeong, Y. S., Jeong, M. K., Garcia-Diaz, A., and Kim, B. (2010), Functional kernel-based modeling of wavelet compressed optical emission spectral data: Prediction of plasma etch process, *IEEE Sensors Journal*, **10**(3), 746-754.

Kolari, K. (2008), Deep plasma etching of glass with a silicon shadow mask, *Sensors and Actuators A: Physical*, **141**, 677-684.

Picard, R. W., Vyzas, E., and Healey, J. (2001), Toward machine emotional intelligence: Analysis of affective physiological state, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**(10), 1175-1191.

Rosipal, R. and Trejo, L. J. (2001), Kernel partial least squares regression in reproducing kernel Hilbert space, *Journal of Machine Learning Research*, **2**, 97-123.

Sugawara, M. (1998), *Plasma etching fundamentals and applications*, Oxford University Press, Oxford.

Zhang, G.-M., Harvey, D. M., and Braden, D. R. (2006), Resolution improvement of acoustic micro-imaging by continuous wavelet transform for semiconductor inspection, *Microelectronics Reliability*, **46**, 811-821.