8th International IFAC Symposium on
Dynamics and Control of Process Systems
Preprints Vol.1, June 6-8, 2007, Cancún, Mexico

DEVELOPMENT OF ENDPOINT DETECTION ALGORITHM IN THE MULTI-STEP PLASMA ETCHING PROCESS

Kyounghoon Han[1], Kun Joo Park[2], Heeyeop Chae[3*], Chonghun Han[1] and En Sup Yoon[1]

[1]School of Chemical & Biological Engineering, Seoul National University, Seoul, South Korea
[2]Quantum Plasma Services Co. Ltd., Suwon, South Korea
[3]Department of Chemical Engineering, Sungkyunkwan University, Suwon, South Korea

Abstract: An endpoint detection algorithm based on principal component regression is developed for the multi-step plasma etching process with the whole optical emission spectra data. Because many endpoint detection techniques use a few manually selected wavelengths, noise render them ineffective and it is hard to select important wavelengths. Furthermore, the smaller the open area changes, the more difficult this single wavelength method detection the endpoint. In this paper, the principal component regression between two wafers was used for the real-time endpoint detection In case study, we applied our multiple models to the multi-step plasma etching process, which consisted of continuous polysilicon etching after the bottom anti-reflective coating etching. So we could obtain the simple and clear information for the more effect endpoint detection, which can be used for the improved process monitoring afterwards. *Copyright © 2007 IFAC*

Keywords: Multi-Step Plasma Etching, Optical Emission Spectroscopy, Principal Component Regression, Endpoint Detection

1. INTRODUCTION

In semiconductor processing, plasma etching is typically employed to define the micro- and nano-scale patterns on a silicon wafer. When the target layer is cleared, it is critical to stop the plasma etching to avoid excessive over-etching and this event called the endpoint detection (EPD). Typically the uniformity of film thickness should be maintained within 5% or so and it is inevitable to over-etch to ensure that all the lines, contacts and vias are completed. However, excessive over-etching may remove the film underneath the target layer, and too much over-etching can cause device failures and subsequent yield reduction. Therefore it is critical to determine the endpoint without damaging of the underlayer.
The most widely used method for end point detection is to monitor the optical emission trace of reactive species in plasmas using optical emission spectrometer (OES).

* Corresponding author. Tel.:+82 31 290 7342;
fax:+82 31 290 7272; E-mail:hchae@skku.edu

By measuring optical emission signal intensities at specific wavelengths, one can identify the neutral particles and ions present in the plasma. Most endpoint detection methods using OES focus on identifying a single wavelength corresponding to a chemical species that shows a pronounced transition at the endpoint.
Biolsi *et al*. [1999] demonstrated an advanced endpoint system for small open-area etching by applying threshold signal processing with single wavelength signal. This single wavelength method cannot avoid the noise problem or time delay associated with filtering. Furthermore selection of appropriate wavelengths requires significant experience of process engineers. So these methods usually reliably work only for the large open area wafers (typically larger than 10%). White *et al.* (2000) proposed T^2 and Q statistics for the endpoint detection of low open-area wafers using PCA in conjunction with T^2 detection and recursive mean update. They improved signal sensitivity but their model cannot include the drift of the process. To overcome this limitation, recursive mean and covariance update are needed for real-time adjustment. Yue and co-workers (2001) extracted a reliable endpoint signal using the principal component analysis (PCA). In this algorithm, loading vectors are used for the wavelengths selection, and principal component (PC) values are monitored for the EPD. They suggested sphere criterion method using the loading vectors for the wavelengths

selection but this method also didn't consider the abnormal process condition.
The more OES information could be captured, the more processing power (time, memory, etc.) should be necessary. Furthermore multiple EPD procedure needs for the multi-stage wafer etching as common as real process. In this paper, we present the algorithm for the EPD prediction in the real time multi-step batch process with the whole optical emission spectra using the principal component regression (PCR) and concept of production ratio as shown followings.

2. PRINCIPAL COMPONENT REGRESSION

Principal Component Analysis (PCA) is a favorite tool of chemometricians for data compression and information extraction (Jackson, 1991; Wise and Kowalski, 1995a; Wise et. al 1996; Wold, et. al. 1987a). PCA finds combinations of variables or factors that describe major trends in a data set. Mathematically, PCA relies on an eigenvector decomposition of the covariance or correlation matrix **X** corresponds to samples while columns correspond to variables. For given data matrix **X** with **m** rows and **n** columns the covariance matrix **X** is defined as

$$\mathrm{cov}(X) = \frac{X^T X}{m-1} \qquad (1)$$

This assumes that the columns of **X** have been "mean centered", i.e. adjusted to have a zero mean by subtracting off the mean of each column. Equation 1 gives the correlation matrix of **X**. (Data should be adjusted to zero mean an unit variance by dividing each column by its standard deviation) PCA decomposes the data matrix **X** as the sum of outer product of vectors $\mathbf{t}_i$ and $\mathbf{p}_i$ plus a residual matrix **E** :

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}^T_1 + \mathbf{t}_2\mathbf{p}^T_2 + \ldots + \mathbf{t}_k\mathbf{p}^T_k + \mathbf{E} \qquad (2)$$

Here k must be less than or equal to the smaller dimension of **X,** i.e. $k \le \min\{m, n\}$. The $\mathbf{t}_i$ vectors are known as ***scores*** and contain information on how the samples relate to each other. The $\mathbf{p}_i$ vectors are eigenvectors of the covariance matrix, i.e. for each $\mathbf{p}_i$

$$\mathrm{cov}(\mathbf{x})\mathbf{p}_i = \lambda_i\mathbf{p}_i \qquad (3)$$

where λ_i is the eigenvalue associated with the eigenvector $\mathbf{p}_i$. In PCA the $\mathbf{p}_i$ are known as ***loadings*** and contain information on how variables relate to each other. The $\mathbf{t}_i$ form an orthogonal set($\mathbf{t}_i^T\mathbf{t}_j = 0$ for $i \neq j$), while the **pi** are orthonormal ($\mathbf{p}_i^T\mathbf{p}_j = 0$ for $i \neq j$, $\mathbf{p}_i^T\mathbf{p}_j = 1$ for i=j). Note that for X and any $\mathbf{t}_i$, $\mathbf{p}_i$ pair.

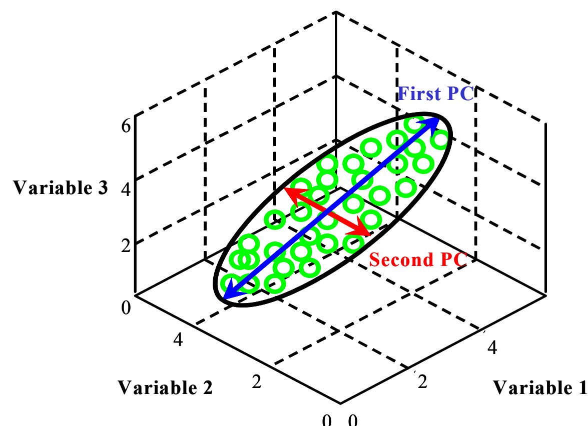


Fig. 1. Principal component model of three dimensional data set lying primarily in a single plane. (Barry M. Wise, *et al.*, 1996).

$$\mathbf{X}\mathbf{p}_i = \mathbf{t}_i \qquad (4)$$

This is because the score vector $\mathbf{t}_i$ is the linear combination of the original **X** data defined by $\mathbf{p}_i$. The $\mathbf{t}_i$, $\mathbf{p}_i$ pairs are arranged in descending ordered according to the associated λ_i. The λ_i are a measure of the amount of variance described by the $\mathbf{t}_i$, $\mathbf{p}_i$ pair. In this research, we can think of variance as information. Because the **ti**, **pi** pairs are in descending order of λi, the first pair capture the largest amount of information of any pair in the decomposition. In fact, it can be shown that the $\mathbf{t}_1$, $\mathbf{p}_1$ pair captured the greatest amount of variation in the data that it is possible to capture the greatest possible variance remaining at that step.
The concept of principal components is shown graphically in Figure 1. The figure shows a three dimensional data set where the data lie primarily in a plane, thus the data is well described by a two principal component (PC) model. The first eigenvector or PC aligns with the greatest variation in the data while the second PC aligns with the greatest amount of variation that is orthogonal to the first PC. Generally it is found that the data can be adequately described using far fewer principal components than original variables.
Principal Component Regression (PCR) is one way to deal with the problem of ill-conditioned matrices. Instead of regressing the system properties (*e.g.* concentrations or level) on the original measured variables (*e.g.* spectra or temperature), the properties are regressed on the principal component scores of the measured variables, (which are orthogonal and, therefore, well conditioned).

3. REAL TIME MULTI-STAGE EPD ALGORITHM WITH PCR

In the process, we often decide when we should stop the plasma at a reactor with small information, and usually use the former wafer information for capturing signals to apply to the next wafer. For this real time detection we used the loading vectors of the former wafer and predicted by product with the data

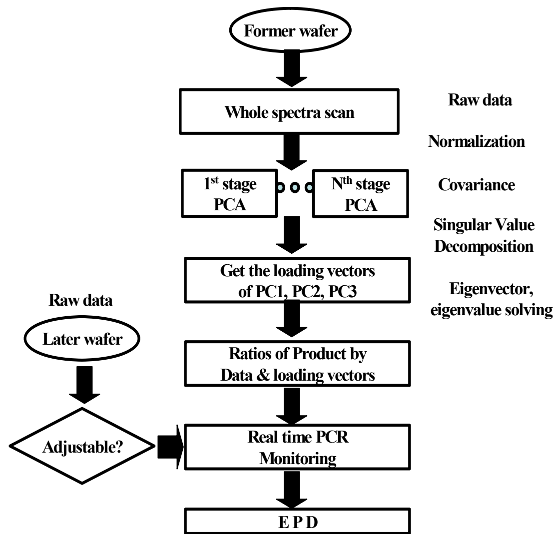


Fig. 2. Multi-stage EPD algorithm using the PCR

of next wafer. If there are many stages for the etching, we need many different PCA models for each stage.
We used this principal component regression for the score vector ($\mathbf{t}_{i,new}$) prediction of new data matrix ($\mathbf{X}_{new}$) by the production with loading vector ($\mathbf{p}_{i,old}$).

$$\mathbf{X}_{new}\mathbf{p}_{i,old} = \mathbf{t}_{i,new} \quad (5)$$

For the real time process there is not enough time for the normalization for the coming data. So we used the original data (OES intensity) for this real time prediction.
Initially, the entire range of OES signals from the first wafer was captured and normalized. The covariance of this normalized data was obtained and a singular value decomposition (SVD) performed. And the loading vectors were obtained from solving of eigenvalue problem of the result of its SVD.
Finally, the entire range of OES signals of the second wafer was captured in real time for multiplying with the loading vector of the first wafer. Over 80% of information can be represented only with 3 products of first 3 PCs in most cases. The sensitivity can be enhanced further by using the ratios of these 3 products. This modelling can be applied each stage of process simultaneously.

4. CASE STUDY

This experiment was performed using a Quantum Plasma Service's HICP polysilicon etch chamber^{TM}. As shown in figure 3, the high density plasma is generated by a main RF power (13.56MHz) to an inductive coil around a dielectric window, and a separate bias RF power (12.56MHz) is applied to the cathode equipped with a ceramic electrostatic chuck (ESC) and thermally controlled using helium backside cooling. The view port is located at the left side of the wafer's flat zone. The chamber is evacuated by a 2300 L/s turbo-molecular pump backed by a dry mechanical pump. Process gases are introduced from a top gas inlet through the center of dielectric window with controlled flow rate, and

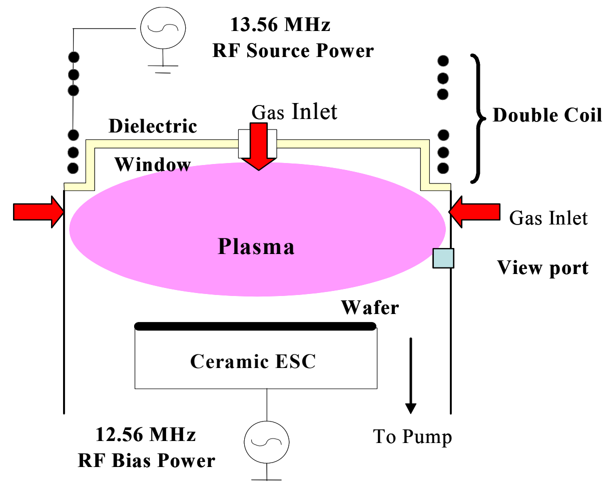


Fig. 3. Double coiled HICP

chamber pressure is controlled with throttling gate valve (TGV).
The signals (2754 wavelengths of whole spectra from 200 to 1,100 nm) can be collected by the OES spectrometer (AvaSpec-3468 of AVANTES Co.).
We did Bottom Anti-Reflective Coating (BARC) and polysilicon etching of with continuous six 200mm wafers, which have the film 4650Å DUA PR / 620Å BARC / 1600Å Polysilicon / 20Å Gate Oxide / Si substrate, as shown in figure 4.
We operated this process at 80℃ cathode, 65℃ wall temperature, 8torr wafer backside He cooling. And we made the process condition at 8mT/300W source power, 90W Bias power, 80HBr/10O_2/40Ar for the BARC etching gas inlet and 20mT/300W source power, 60W Bias power, 120HBr/20Cl_2/12O_2 for the polysilicon etching gas inlet.
At first we compared the product ratio of PC1 (which contains 65.13% information) and PC2 (which contains 6.97% information) of first wafer's data itself and the two important wavelengths (307.7 and 495.9 nm) as shown in figure 5. Because they changed during the same period, we could confirm that using this PCR method would show the exact EPD time. In figure 6 we compared two product values. Estimate product ratio of PC1 and PC2 was from the first wafer's loading vector and the second wafer's real time data. Actual product ratio of PC1 and PC2 was from the loading and data of second data itself as same manner as figure 5 of first wafer.

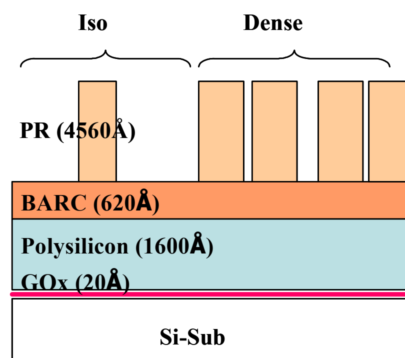
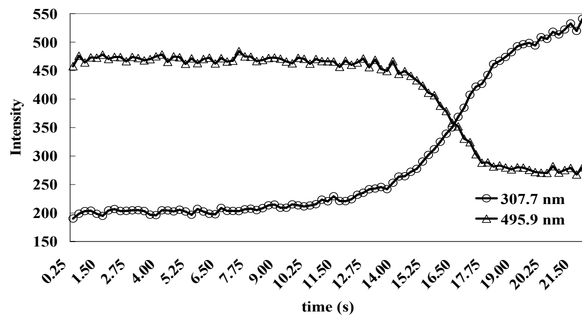
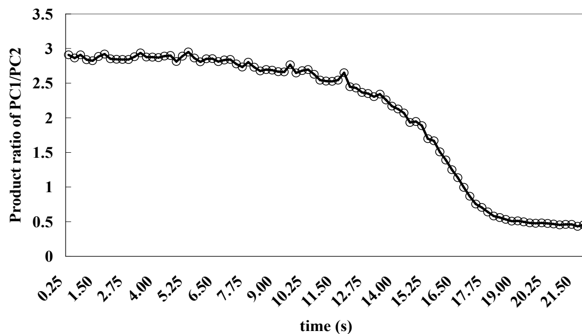


Fig. 4. Wafer composition for the multi-stage etching



(a)



(b)

Fig. 5. Comparison of EPD with two single wavelengths (a) and PCR model (b)

In this experiment we decided to use the ratio of PC1 (which contains 65.14% information) product and PC2 (which contains 5.63% information) product for the explicit behaviours. In this diagram we made a decision of EPD at 15.758s with simple clear estimated curves which is not different from the actual curve. So we could select the time when the curve changed significantly.
After this BARC etching, we did polysilicon etching and the PCR prediction was done by our multi-step modelling. We could capture at about 15s by the change of the product ratio curve as BARC etching as shown in figure 7 (end of the time scale). This figures were from the loading vector value of the second wafer and the data of third wafer of PC1 (which contains 36.91% information) and PC2 (which contains 14.29% information).

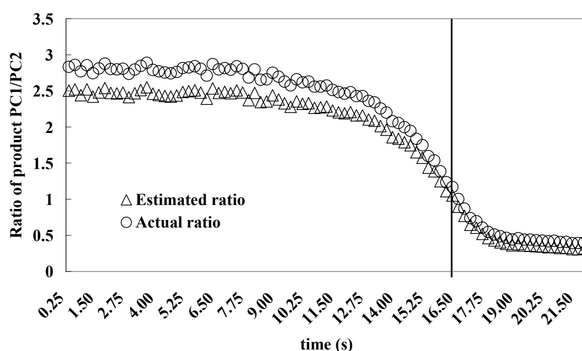


Fig. 6. Endpoint detection of the second wafer with estimated product ratio (BARC etching)

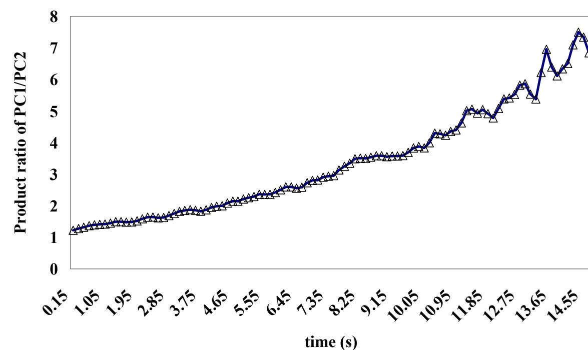


Fig. 7. Endpoint detection of the third wafer with estimated product ratio (polysilicon etching)

We could makes effective EPD of multi-step etching by our PCR modelling with whole emission spectra data and this method gave us simple and easy information of the process state by the observation of the product ratio curve change.

CONCLUSION

In this paper we developed endpoint detection algorithm of the multi-step plasma etching process using the PCR methods. We used the loading vector of former wafer after doing PCA to the data of later wafer. We could determine from the observation that when the curve of this estimated value (ratio of product of PC1 and PC2) changes significantly. In the case study we adjusted this algorithm to the BARC and polysilicon multi-stage etching process in real time and we could get the exact EPD time in each process.
The result shows that this chemometric method is very useful for the simple EPD for the multi-stage etching process and this method is also very important because there are many chambers with one controller which should do EPD control simultaneously.

REFERENCES

Dreeskornfeld, L., *et al*. (2000). "Reactive ion etching end point detection of microstructured Mo/Si multilayers by optical emission spectroscopy", *Microelectronic engineering*, **54**, 303

Gallagher, *et al*. (1997), "Development and benchmarking of multivariate statistical process control tools for a semiconductor etch process: Improving robustness through model updating," in *Proc. ADCHEM*, Banff, Canada, 78

Grill, A. (1994). *Cold Plasma in Materials Fabrication, Ch.5.3*, IEEE Press

Mundt, R. (1997) "Model based training of a neural network endpoint detector for plasma etch application," *in Proc. ADChem*, Banff, Canada

Nello, Cristianini & John, Shawe-Taylor (2001). *An introduction to support vector machines, Ch.6.2*, Cambridge University Press, Cambridge

Pearton, S. J., *et al*. (1999). "Optical emission end point detection for via hole etching in InP and

GaAs power device structures", *Materials Science Engineering*, **B23**, 36
Qin, S. J., et al (1997), "Fault detection and classification of plasma etchers via optical espectroscopy analysis," in *Proc. AIChE Annu. Meet.*, Los Angeles, CA
Qin, S. J., *et al* (2006). "Semiconductor manufacturing process control and monitoring : A fab-wide freamwork", *Journal of Process Control*, **16**, 179
Selwyn, G. S. (1993). *Optical Diagnostic Techniques for plasma Processing.* Manchester, U.K.:AVS
Seo, S. T. (2006). *et al*., "Run-tu-run control of inductively coupled C2F6 plasma of SiO_2 : Multivariable controller design and numerical application.", *Korean J. Chem. Eng*. **23(2)**, 199
Shadmehr, R., *et al*. (1992), "Principal component analysis of optical emission spectroscopy and mass spectrometry: Application to reactive ion etch process parameter estimation using neural networks," J. Electrochem. Soc., 139, 907
Vladmir, Cherkassky and Filip, Mulier (1998). *Learning from data concepts, theory, and methods*, John Wiley & Sons, Inc., New York
White, D., Boning, D., Butler, S. W. and Bama, G. G. (1997). "Spatial charaterization of wafer state using principal component analysis of optical emission spectra in plasma etch," *IEEE Trans. Semicond. Manuf.* **10**, 52
White, D., Goodlin, B. E., Gower, A. E., Boning, D. S., Chen, H. Sawin, H. H. and Dalton T. J. (2000). *IEEE Trans. Semiconductor Manufacturing*, **13**, 193
Wise, B. M., *et al*. (1997) "Development and benchmarking of multivariate statistical process control tools for a semiconductor etch process: Impact of measurement selection and data treatment on sensitivity," in *Proc. IFAC SAFEPROCESS, Hull*, U.K.
Wise, B. M., *et al*. (1999). "A comparison of principal component analysis, multiway principal component analysis, trilinear decomposition and parallel factor analysis for fault detection in semiconductor etch process", *Journal of Chemometrics*, **13**, 379
Wise, B. M. et al. (2004). *PLS_Toolbox 3.5 for use with MATLAB™*, Eigenvector Research
Yue. H. H., *et al*. (2001). "Plasma etching endpoint detection using multiple wavelengths for small open-area wafers", *Journal of Vacuum Science & Technology*, **A 19**, 66