

Copyright © 1997, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

**OES-BASED SENSING FOR PLASMA
PROCESSING IN IC MANUFACTURING**

by

Roawen Chen

Memorandum No. UCB/ERL M97/90

15 December 1997

**OES-BASED SENSING FOR PLASMA
PROCESSING IN IC MANUFACTURING**

Copyright © 1998

by

Roawen Chen

Memorandum No. UCB/ERL M97/90

15 December 1997

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ABSTRACT**OES-based Sensing for Plasma Processing in IC Manufacturing**

by

Roawen Chen

Doctor of Philosophy in Engineering -
Electrical Engineering and Computer Sciences

University of California at Berkeley

Professor Costas J. Spanos, Chair

Commercial semiconductor technology is expected to reach the 0.18 μm device generation using 300mm wafers by the year 2001. In order to meet this expectation, we need better temporal and spatial process uniformity. This issue is particularly true for plasma etching because it is often considered one of the most critical steps in the manufacturing of deep sub-micron IC devices. The need for better control suggests that we must have a robust model-based controller to ensure the IC manufacturers can meet ever tightening specifications.

In this work we are developing a methodology for estimating etching performance using real-time sensors. Specifically, the sensor readings collected throughout plasma etching processes are used in combination with statistical techniques to model etch rate, within-wafer uniformity, aspect-ratio dependent etching (ARDE), and critical dimension (CD) reduction. Most of this effort has been devoted to developing reliable models that relate process performances to signals acquired from optical emission spectroscopy (OES), a sensor providing spatially resolved, *in-situ*, real-time readings without disturbing the plasma or interfering with the process. Our study is based on the OES sensor systems

installed on various commercial plasma etchers including dielectric etchers (AMAT 5300 Centura), polysilicon etchers (Lam 4400 and 9400), and metal etchers (Lam 9600).

Another aspect of this work is to model the spatial variation of critical dimension (CD) using the spatial and temporal information provided by the OES sensor. Both physical and statistical approaches are employed to extract the spatial and temporal variation from the original OES sensor readings for each wafer at selected wavelengths. An additional goal of this work is to demonstrate a novel technique for monitoring the spatial CD variation and detecting abnormal wafers using OES signals. This is accomplished by a statistic describing the temporal variation from the real-time OES signals collected from three spatially resolved beams, and can be employed to indicate the nature of the CD spatial non-uniformity.

Conventional endpoint detection techniques often provide temporally-resolved information of the etch process, but without spatial resolution. With the emerging of the 300mm wafer production, the potential utility of a spatially-resolved endpoint detector is expected to increase greatly. In this project we propose a novel spatially resolved sensor to detect the endpoint and monitor spatial uniformity for a plasma etching process. A scanning spatially-resolved optical emission spectroscopy (SROES) system was built and installed in the Berkeley Microfabrication Laboratory. This sensor system consists of a stepper motor, controller, and a monochromator, which provides an *in-situ* real-time monitoring of the etching endpoint spatially. The most interesting feature of collected plasma spatial images is their temporal dependence. This information can be then used to detect the spatially-resolved endpoints and monitor the etch uniformity. This concept also extends to a scanning SROES system with a full spectral-range spectrometer. The results suggest that full color plasma emission spatial profiles are useful for characterizing and monitoring the plasma conditions. It is our expectation that these signals will play a useful role in designing the next generation equipment controllers.

~~_____~~ C. J. SPANOS 12/19/97
Chair Date

Table of Contents

Chapter 1. Introduction

| | |
|-----------------------------------|---|
| 1.1 Motivation | 1 |
| 1.2 Thesis Organization | 4 |

Chapter 2. Background

| | |
|--|----|
| 2.1 Previous Work | 7 |
| 2.2 Plasma Etch Process | 9 |
| 2.2.1 Etch Mechanisms | 9 |
| 2.2.2 Etching Processes | 11 |
| 2.2.2.1 Poly-Silicon Etching | 11 |
| 2.2.2.2 Oxide Etching | 12 |
| 2.2.2.3 Metal Etching | 14 |
| 2.3 Plasma Etch Equipment | 15 |
| 2.3.1 Capacitive Planar Etcher | 16 |
| 2.3.2 High-Density Plasma Etcher | 17 |
| 2.4 In-Situ Sensors for the Etch Process | 19 |
| 2.5 Optical Emission Spectroscopy | 20 |
| 2.5.1 Quantitative Description of Optical Emission | 21 |
| 2.5.2 OES Instrumentation | 22 |
| 2.5.3 Actinometry | 24 |
| 2.6 Multivariate, Statistical Data Compression | 25 |
| 2.6.1 Principal Component Regression | 25 |
| 2.6.2 Partial Least Squares Regression | 27 |

Chapter 3. Wafer-State Modeling Using OES for Oxide Etch

| | |
|---|----|
| 3.1 Introduction | 33 |
| 3.2 Experiment to Link OES Signals to Oxide Etch Variability | 34 |
| 3.2.1 Data Acquisition | 34 |
| 3.2.2 Experimental Design | 36 |
| 3.3 Modeling Results | 37 |
| 3.3.1 Regression models using the intensities of the pre-selected wavelengths | 38 |
| 3.3.2 Regression models using PCA reduced variables | 41 |
| 3.3.3 Regression models using PLS reduced variables | 45 |
| 3.4 Discussion | 47 |
| 3.5 Summary | 49 |

Chapter 4. Wafer-State Modeling Using OES for Metal Etch

| | |
|--|----|
| 4.1 Introduction | 52 |
| 4.2 Experiment to Link OES Signals to Metal Etch Variability | 52 |
| 4.3 Data Pre-Processing | 56 |

| | |
|--|------------|
| 4.3.1 OES Temporal Variation | 56 |
| 4.3.2 OES Variable Selection | 58 |
| 4.4 Results | 59 |
| 4.4.1 Relating OES and Machine Variables | 59 |
| 4.4.2 Relating OES and Wafer-state Variables | 60 |
| 4.5 Discussion | 62 |
| 4.6 Summary | 64 |
| | |
| Chapter 5. CD Modeling and Monitoring Using SROES for Poly-Si Etch | |
| 5.1 Introduction | 67 |
| 5.2 Experiment to Link OES Signals to CD Variability | 69 |
| 5.3 Models of Average CD Reduction | 72 |
| 5.3.1 Principal Component Regression Model | 72 |
| 5.3.2 Partial Least Squared Regression Model | 75 |
| 5.3.3 Comparison of Models | 76 |
| 5.4 Monitoring and Modeling of CD Spatial Variation | 77 |
| 5.4.1 Extraction of temporal variation of OES signals | 78 |
| 5.4.2 Extraction of spatial variation of OES signals for a three-beam sensor | 82 |
| 5.4.3 Modeling Results | 84 |
| 5.4.4 Monitoring of CD Spatial Variation | 86 |
| 5.5 Discussion and Summary | 87 |
| | |
| Chapter 6. Scanning Spatially Resolved OES | |
| 6.1 Introduction | 91 |
| 6.2 Previous Work | 92 |
| 6.3 Instrument Set-up at the Berkeley Microfabrication Laboratory | 95 |
| 6.4 Experimental Design Used to Collect Scanning OES Data | 100 |
| 6.5 Data Analysis | 101 |
| 6.5.1 Phase 1: Single Wavelength | 101 |
| 6.5.2 Phase 2: Multiple Wavelengths | 106 |
| 6.5.2.1 Signature Analysis | 107 |
| 6.5.2.2 Ratio Analysis | 110 |
| 6.5.2.3 Principal Component Analysis | 111 |
| 6.6 Discussion | 114 |
| 6.7 Summary | 116 |
| | |
| Chapter 7. Conclusion and Future Work | |
| 7.1 Thesis Summary | 119 |
| 7.2 Future Work | 123 |
| | |
| Appendix A Splus Code | 125 |
| | |
| Appendix B Matlab Code | 137 |
| | |
| Appendix C Experimental Design | 143 |

| | |
|---|-----|
| Appendix D Locally Weighted Regression Smoothing | 148 |
| Appendix E Element vs. Wavelength | 150 |

List of Figures

| | | |
|------------|--|----|
| Figure 1.1 | Historical and perspective trend of productivity improvement (modified from [1.4]) | 2 |
| Figure 2.1 | Four basic plasma etching mechanisms: (a) sputtering; (2) pure chemical etching; (c) ion energy-driven etching; (d) ion-enhanced inhibitor etching [2.9] | 10 |
| Figure 2.2 | Typical cross-section for logic CMOS IC..... | 12 |
| Figure 2.3 | The illustrations of (a) capacitive RF discharges in plane parallel geometry; (b) high-density inductively coupled plasma.[2.8] | 16 |
| Figure 2.4 | Schematic diagram of atomic emission spectrum..... | 21 |
| Figure 2.5 | Schematic illustration of the OES instrumentation..... | 23 |
| Figure 2.6 | Principal component transformation in two dimensioned space (• denotes data point) | 27 |
| Figure 2.7 | The algorithmic flow of Principal Component Regression..... | 28 |
| Figure 2.8 | PLSA iteration | 30 |
| Figure 3.1 | Wafer state modeling scheme using OES sensor readings. | 34 |
| Figure 3.2 | Schematic illustration of the experimental set-up on AMAT 5300 oxide etcher..... | 35 |
| Figure 3.3 | Cross section of the test structure for a monitor wafer. | 36 |
| Figure 3.4 | A typical spectrum collected from the Applied Materials 5300 during the main etch (etchant is C_2F_6) with spectral lines labeled with their corresponding species. | 38 |
| Figure 3.5 | Wafer-to-wafer plots of intensities for three different wavelengths (upper: 251nm, middle: 288nm and bottom:440nm)..... | 40 |
| Figure 3.6 | Scree plot of the principal component analysis of OES data (the number shown on the top of each bar indicates the cumulative summation of their explained percentage of variations of total OES data)..... | 42 |
| Figure 3.7 | The loading vectors of the first seven principal components..... | 43 |
| Figure 3.8 | The plots of PRESS for all wafer responses (a) oxide etch rate; (b) uniformity; (c) ARDE at center; (d) ARDE at edge | 46 |
| Figure 3.9 | The flow chart of wafer state modeling | 47 |
| Figure 4.1 | An overview of this study. | 53 |
| Figure 4.2 | The TiN/Al-Cu/TiN/Oxide stack (a) before etch, and (b) after etch. | 54 |
| Figure 4.3 | The identification of the optical emission spectra collected from a Lam 9600 TCP metal etcher (etchant: Cl_2/BCl_3) | 55 |
| Figure 4.4 | Typical contour plots of (a) aluminum CD measurements; (b) oxide loss measurements. (the unit of the contour plots is angstrom)..... | 56 |

| | | |
|-------------|--|----|
| Figure 4.5 | The temporal trajectory of optical emission intensity for one wafer at the Cl line..... | 57 |
| Figure 4.6 | Cl line intensity versus machine settings (each line denotes the fitted linear regression model) | 59 |
| Figure 4.7 | Univariate correlation between emission intensity and TCP power over all 2042 wavelengths in terms of their R-square value..... | 60 |
| Figure 4.8 | The wafer-to-wafer trajectory of oxide etch rate and emission intensity values from Cl line..... | 61 |
| Figure 4.9 | Oxide etch rate versus machine settings (each line denotes the fitted linear regression model) | 62 |
| Figure 4.10 | Information captured by machine-state, process-state, and wafer-state sensors..... | 64 |
| Figure 5.1 | A typical spectrum collected from the Lam 4400 during the main etching (etchant is Cl ₂) with spectral lines labeled with the corresponding species..... | 68 |
| Figure 5.2 | Schematic illustration of three-beam spatially-resolved optical emission spectroscopy | 69 |
| Figure 5.3 | Top view of spatially-resolved OES..... | 70 |
| Figure 5.4 | Cross sections of the test structures before and after etching (PR stands for photoresist)..... | 71 |
| Figure 5.5 | (a)The screeplot of OES data; (b) PRESS-statistics | 73 |
| Figure 5.6 | Loading plots of the first 4 principal components (the spectral lines marked are difficult to observe in the original spectrum) | 74 |
| Figure 5.7 | PRESS vs. the number of variables in model | 76 |
| Figure 5.8 | Typical OES trajectories for different spectral lines. | 79 |
| Figure 5.9 | Poly-Si etched spatial profile vs. OES trajectory. | 80 |
| Figure 5.10 | Temporal variance extraction from the OES trajectories..... | 81 |
| Figure 5.11 | The OES trajectories and CD plots for (a) wafer #2; and (b) wafer #30 | 82 |
| Figure 5.12 | Extraction of OES spatial variance. | 85 |
| Figure 5.13 | The SPC chart of (a) CD spatial variance; (b) OES temporal variance. . | 87 |
| Figure 6.1 | Schematic illustration showing the top view of the discharge chamber for (a) references [6.1][6.2][6.3][6.4][6.5]; (b) reference [6.6]..... | 93 |
| Figure 6.2 | Schematic illustration of the scanning SROES set-up and discharge chamber (a) side view; (b) top view..... | 96 |
| Figure 6.3 | Geometrical illustration of the etch chamber (Lam 4400) and optical paths of the scanning mirror | 97 |
| Figure 6.4 | Schematic “resolution” plots for three different SROES systems with (a) high spatial resolution and low temporal resolution; (b) low spatial resolution and high temporal resolution; (c) high spatial resolution and high temporal resolution. (• denotes data point at time t and spatial point s) | 99 |

| | | |
|-------------|---|-----|
| Figure 6.5 | A typical spectrum collected from the Lam 9400 during main etching (etchant is Cl_2/HBr) with spectral lines labeled with their corresponding chemical species..... | 101 |
| Figure 6.6 | Typical plasma lateral profiles collected by scanning SROES with monochromater turned at (a) 405nm; (b) 262nm..... | 102 |
| Figure 6.7 | The temporal evolution of spatial plasma images (at 262nm) before smoothing..... | 103 |
| Figure 6.8 | (a) OE intensity ratios v.s. time, and (b) its corresponding Si plot | 104 |
| Figure 6.9 | Control chart based on (a) statistical score S_i , and (b) over-etch uniformity | 105 |
| Figure 6.10 | Arrangement and decomposition of a three-way array of scanning OES data set..... | 106 |
| Figure 6.11 | Typical $I(\theta, t)$ trajectory at (a) 252nm; (b) 838nm..... | 108 |
| Figure 6.12 | Typical $I(l, t)$ plots at the central angular point (a) full spectral range plot, (b) 500-1000nm spectral range plot, (c) full spectral range image plot (NOTE: for image plot, black marks maximum intensity and white marks minimum intensity) | 109 |
| Figure 6.13 | Typical $I(l, \theta)$ plots at the 10th scanning cycle (a) full spectral range plot, (b) 500-1000nm spectral range plot, (c) full spectral range image plot | 109 |
| Figure 6.14 | $I(l, \theta)$, $I(l, t)$, and $I(\theta, t)$ plots at different top electrode power levels (a) 270 W, (b) 330 W..... | 110 |
| Figure 6.15 | A typical ratio plot of normalized spatial-temporal profiles of $I(252\text{nm})/I(838\text{nm})$ | 111 |
| Figure 6.16 | Temporal profiles at the four principal components after PCA compression (the first four PCs explain 94% of the variability of optical emission) | 112 |
| Figure 6.17 | Spatial profiles at the four principal components after PCA compression (the first four PCs explain 98% of the variability of optical emission) | 113 |
| Figure 6.18 | $\text{PC}(\theta, t)$ plots of the spatial profiles of the first four principal components | 113 |
| Figure 6.19 | (a) oxide etch rate uniformity of wafer #4 and its contour plot, (b) its corresponding spatial-temporal plot based on 3rd and 4th PCs, and (c) its corresponding ratio plot of wavelength 252nm and 838nm..... | 115 |
| Figure 7.1 | OES “resolutions”..... | 121 |
| Figure 7.2 | A schematic illustration of the relations among etch variability, plasma variability, and OES detection capability (not in scale). ... | 122 |

List of Tables

| | | |
|-----------|---|-----|
| Table 2.1 | Typical feed gas and mechanisms for plasma etch [2.9]..... | 15 |
| Table 2.2 | Range of parameters for RIE and high-density discharge[2.8] | 17 |
| Table 2.3 | Commercially available sensors for plasma etch (modified from Ref.[2.14])..... | 20 |
| Table 3.1 | Change in Percent from Nominal | 37 |
| Table 3.2 | The distinct wavelengths selected for the ordinary least squares regression model. | 40 |
| Table 3.5 | The ANOVA Tables for PCR models of all responses. | 44 |
| Table 3.4 | Proportion of predictor and response variability explained by principal com- ponents (Bold: the PCs selected for PCR models) | 44 |
| Table 3.6 | The ANOVA tables of PLSR models for all responses. | 46 |
| Table 3.7 | Summary of the results of chamber state models. | 49 |
| Table 5.1 | Central Composite Experimental design - nominal and extreme settings | 71 |
| Table 5.2 | Proportion of predictor and response variability explained by principal com- ponents. | 74 |
| Table 5.3 | Comparison of modeling methods for CD reduction. | 77 |
| Table 5.4 | Possible physical causes of the decomposed OES variation. | 86 |
| Table 6.1 | Summary of the previous work related to spatially resolved OES | 94 |
| Table 6.2 | The optimized scanning SROES system settings | 98 |
| Table 6.3 | Machine settings for a 2 ⁴ experimental design | 100 |

Acknowledgments

I would like to thank my research advisor, Professor Costas J. Spanos, for his guidance, support, and encouragement during my graduate years. I would like to thank Professor Andy Neureuther for his professional, support and assistance. I would like to thank Professor David Brillinger for his helpful comment throughout my graduate years in Berkeley. I would like to acknowledge Professor Nathan Cheung for serving the chairman of my Qualification Examination. Many thanks to Professor Lieberman and Professor Flamm for their valuable suggestion and discussion concerning my research work.

I acknowledge Michael Gatto for mentoring my summer stay at AMD. I would like to especially thank John Plummer of National Semiconductor for his mentorship during my two years as a SRC Fellowship recipient. I acknowledge Uzi Vierny, Manus Wong, and Liang-Guo Wang of Applied Material, Randall Mundt of Lam Research, Rick Dill of IBM Corporation, Dan White, Gabriel Barna, and Stephanie Bulter of Texas Instruments and Linda Milor of AMD, for their valuable discussion in many occasions. I would like to thank the Semiconductor Research Corporation for their generous monetary support. I would like to thank Piotr Zalicki of SC Technology for his support and help in this work.

I would like to thank Katalin Voros, Robert Hamilton, David Mudie, John Simon, and the other staff of the Berkeley Microfabrication Laboratory for their expertise and gracious assistance.

I would like to acknowledge the past and present members of the Berkeley Computer-Aided Manufacturing research group for making my life in Cory Hall colorful. I want to thank Crid Yu, Sean Cunningham, Sherry Lee, Xinhui Niu, Eric Boskin, Sundeep Rangan, Tony Mirana, Nickhil Jakatdar, Mareike Claassen, Sov Leung, Hsing-yee Ma, Jeff Lin, Greg Luurtsema, Anna Ison, John Musacchio, Mason Freed, Haolin Zhang, Mien Li, Dong Wu, Nikhil Vaidya, Mark Hatzilambrou, Robert Tu, Manolis Terrovitis, Junwei Bao, Clement Wann, Jeff Ou, George Chien, Eric Minami, Hui Zhang, Joe King, John Deng, James Chen, Jae-wook Lee and Kai Chen for their friendship. The lessons which I

have learned with many of you, have definitely enriched my life. I would like to especially thank Mason Freed and Sundeep Rangan for their help in the scanning OES project.

I would like to thank my family for their encouragement and support. I would like to thank Kung's family for the support and trust you have lent me for these years. Finally, the most thanks to my wife, Cheryl, who has helped to keep my spirit high and perspective broad throughout my graduate school years. Without her love and support, this would not be possible. I share this degree with her.

CHAPTER 1

Introduction

1.1. Motivation

Commercial semiconductor technology is expected to reach the 0.18 μ m device generation using 300mm wafers by the year 2001 [1.1]. The fabrication of integrated circuits (ICs) is becoming very capital-intensive, and if the past exponential trend in fab cost continues, the cost of a state-of-the-art facility for their manufacture is approaching US\$3 billion by the year 2000 and US\$10 billion by the year 2005 [1.2]. Moreover, a fab's life span is reducing over the past decade. In the mid-'80, a fab cost about \$100 million and had an expected life span of nearly ten years. Now we have \$1 billion fabs with a life span of only five years. The depreciation alone is approximately \$4 million a week [1.3]. It is expected that semiconductor industry can not afford to go on investing at these level while remaining profitable. Thus, to offset this enormous investment, IC manufacturers must achieve high yield, throughput and productivity quickly.

Historically, the IC industry has reduced manufacturing cost by improving yields, increasing wafer size, shrinking linewidths and improving equipment productivity. However, the cost of funding the transition to larger wafer size (e.g., > 300mm) and smaller linewidths (e.g., < 0.18 μ m) is probably too high to justify the advantage of reducing the overall cost by increasing productivity. Moreover, a portion of the productivity improvement due to larger wafer size is often offset by the increasing technological challenge for achieving a good spatial uniformity. It has also been shown [1.3][1.4] that one area where we have more room to improve is the area of equipment productivity (see Figure 1.1).

Additionally, it is extremely difficult to measure each wafer after each step in high volume semiconductor factories, which produce chips with over 100 manufacturing steps and high throughput. Present strategy is to measure and monitor wafers periodically after changing the machine settings, however, at the penalty of cost and throughput. Although this method gives an indication of how the machine is operating, it offers no guarantee that the subsequent production wafers will be processed properly.

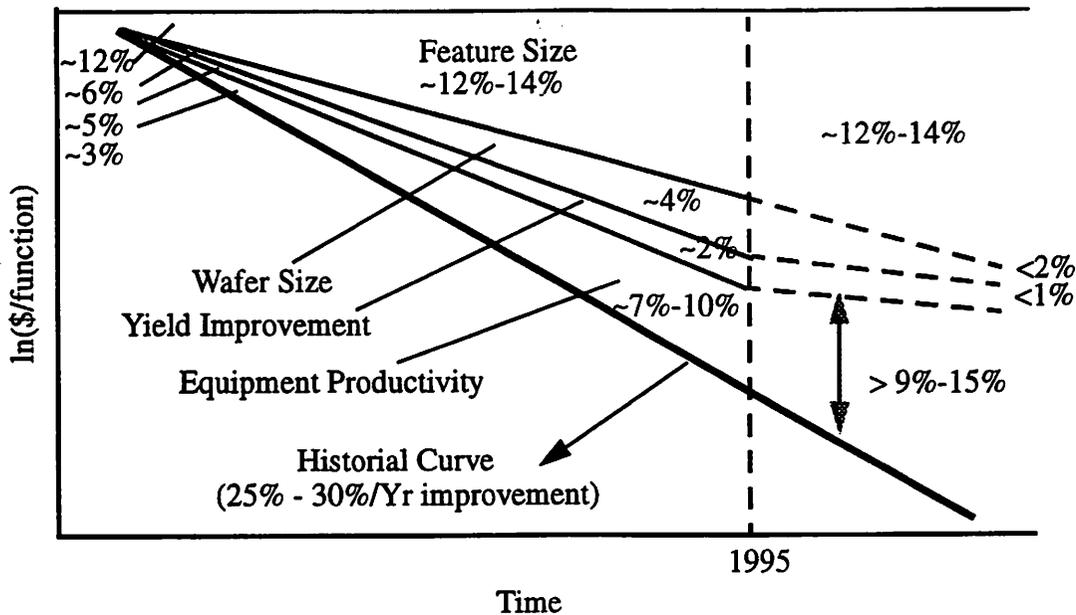


Figure 1.1 Historical and perspective trend of productivity improvement (modified from [1.4])

Therefore, to achieve short cycle times and low-cost wafer fabrication, sensor-aided manufacturing (SAM) of ICs is emerging as an approach to optimize the cost-effectiveness of IC production [1.5]. One key SAM concept is to monitor the equipment and wafer conditions *in-situ* to ensure that the semiconductor wafers are processed properly during each step. The test wafers, machine calibration and maintenance can be thus minimized by the means of SAM so that the overall equipment effectiveness can be improved.

Plasma etching has been widely used in the manufacturing of sub-micron IC devices, and it is also considered to be one of the yield limiters. In the past, empirical optimization of plasma parameters has produced many successful etching and deposition “recipes”. A lot of effort has been spent developing reliable empirical models that relate the response of process outputs (such as etch rate) to variations in input parameters (such as pressure, RF power, or gas flow). These models are required by semiconductor manufacturers in order to predict etch behavior under an exhaustive set of operating conditions to a high degree of precision. Traditionally, this approach has relied on statistical techniques without detailed reference to plasma physics and microscopic etch mechanisms. By contrast, numerical simulations based on fundamental physical theories have been used in academia, but these are computationally expensive and typically take too long for real-time manufacturing applications. Also, these efforts have focused on explaining plasma parameters with little description of spatial and temporal etch performances.

Optical emission spectroscopy (OES) has been widely used in plasma assisted process for *in-situ* process diagnostics and monitoring. Generally, a complete OES system should be capable of resolving three components of plasma variability: spectral resolution (i.e., chemical resolution), temporal resolution, and spatial resolution. However, the ideal OES system with high spectral, temporal, and spatial resolutions is very expensive. Therefore, the commercial OES system often gives up one of these information “dimensions” to

reduce the cost. For example, the full-range OES commercial system only provides a limited spatial resolution. My research objective is to employ *spatially-resolved* optical emission spectroscopy (SROES) as a novel real-time fault detection tool for plasma etch. The OES sensor with spatial resolution has a unique advantage in plasma processing because it provides *in-situ* spatial, temporal, and spectral information without disturbing the plasma or interfering with the process. This study is based on the SROES sensor systems installed on various commercial plasma etchers including polysilicon etchers (Lam 4400, 9400), oxide etchers (AMAT 5300 Centura™) and metal etchers (Lam 9600 TCP™).

There are several important issues in implementing the OES system for *in-situ* process monitoring and diagnosis. First, since a typical spectrum usually contains thousands of wavelengths, an efficient way to compress and extract spectral information from OES readings is vital to implement the OES system in an IC production environment. Secondly, the information included in the temporal and spatial signatures of OES signals collected from one wafer is difficult to interpret. Thirdly, the relation between the OES signals and wafer characteristics is still under investigation (Note that the optical emission intensity of discharge is only a *direct* indicator of plasma states, not wafer states). This thesis will try to resolve these issues.

1.2. Thesis Organization

This thesis starts from a qualitative description of the plasma etch process and mechanisms, plasma etcher designs, optical emission theory, optical emission spectroscopy, and several statistical multivariate techniques in Chapter 2. Chapter 3 through Chapter 5 develop the wafer-state prediction models using the *spectral* information provided by the OES system. Several statistical data compression techniques are employed to filter out the spectral information from a large number of spectral variables. The beauty of these techniques is that they provide an efficient way to extract the important spectral information *without* identifying the chemical species and their spectral signatures prior to the data

analysis. In other words, we propose a novel way to explain the wafer-states by using *statistical* identification procedures. In addition, we interpret these results with the help of the physical understanding of the process, and discuss the constraints while implementing these wafer-state modeling techniques. Up to this point, no within-wafer temporal and spatial information of OES signals is included in this framework. In Chapter 5, I start to estimate and monitor the *spatial* variability of wafer-states using within-wafer *temporal* and *spatial* information provided by a three-beam spatially-resolved OES system. Both physically based and statistically based approaches are employed to extract the spatial and temporal variation from the original SROES sensor reading for each wafer. Chapter 6 develops a novel scanning SROES system which significantly enhances the spatial resolution of OES readings. *Spatial* information provided by this scanning SROES can be then used to estimate the *spatial* variability of wafer-states and, most importantly, is used as a “spatially-resolved endpoint detector” while observing the temporal dependence of the spatial profiles of plasma emission. Finally, conclusions and future work related to this research are given in Chapter 7.

References for Chapter 1

- [1.1] SIA Semiconductor Technology - Workshop Working Group Reports, 1997
- [1.2] R. R. Schaller, "Moore's Law: Past, Present, and Future", IEEE Spectrum, June 1997, p.53.
- [1.3] S. McIntosh, "Conquering semiconductor's economic challenges by increasing capital efficiency", 1997 IEEE International Symposium on Semiconductor Manufacturing Conference, Oct. 6-8, 1997, San Francisco, California.
- [1.4] J. Owens, "Prospects for Maintaining the Historical Reduction in Cost/Productivity", Manufacturing Science and Technology Group, Technical Sessions at the AVS 43rd National Symposium, Oct. 14-18, 1996.
- [1.5] G. Gettel, "Future Factory Level Issues and Needs", AEC/APC Workshop, September 20-24, 1997, Lake Tahoe, Nevada.

CHAPTER 2

Background

2.1. Previous Work

Sensor-aided statistical process control (SPC) is emerging as a novel technique for monitoring, diagnosing, and modeling the plasma etch process in the modern semiconductor manufacturing environment. There are a number of sensors, either in commercialized or still developing stage, which have been investigated by several research groups for plasma etchers. For example, Lee and Spanos [2.1] [2.2] at the University of California, Berkeley have demonstrated the feasibility of using the machine-state real-time sensor signals (e.g., power, pressure, and gas flow) to monitor the health of the commercial plasma etcher in real-time and model the etch rates accurately.

Optical emission spectroscopy (OES) is probably the most widely used in-situ sensor for plasma etching process. Nonetheless, its usage has been mainly limited to monitoring a *single wavelength* of light emitted from a chemical species of interest such as etch byproducts for endpoint detection or contamination control. For example, Selwyn at IBM [2.3] has used optical emission (OE) to detect tool contamination and malfunction for

plasma etching by monitoring emission from discharge-excited impurities. In addition to species identification and endpoint control, a *full spectral-range* OES is also used as a promising tool to characterize the wafer or machine states. However, the most vital challenge is how to analyze a complex spectrum acquired by the multichannel OES rapidly and accurately. One solution is to analyze the OES data by multivariate statistical techniques. Shadmehr *et al* at IBM [2.4] have proposed to use principal component analysis (PCA) and neural networks (NN) to characterize the effects of machine parameters (e.g., pressure, power, and gas mixture) on the optical emission and mass spectra of CHF_3/O_2 plasma. Similarly, Anderson *et al* at the University of New Mexico [2.5][2.6] have used OES to characterize the oxide etch process. Their results show that the emission fingerprint, as determined by the Chemometrics, correlate well with the oxide etch rate for the CF_4/CHF_3 plasma. Recently, White *et al* [2.7] at MIT have used some non-linear statistical techniques such as nonlinear partial least squares (PLS) and multilayer perceptron (MLP) neural networks to estimate the aluminum linewidth reductions at different sites on the wafer by combining the spectra collected from different regions in plasma using spatially-resolved OES (SROES).

Although these techniques show a promising way to describe the wafer-to-wafer behavior of equipment/wafer states, no real-time and only limited spatial information is included in their frameworks. The typical approach of applying these sensor signals is to model the wafer-state measurements from one selected site by one spectrum collected during main etch period (that is, one OES sample per wafer) [2.4][2.7], or to model the spatial mean of the wafer-state measurements using the corresponding temporal average of sensor signals acquired from one wafer [2.5][2.6]. Although these techniques can predict some of the wafer-states (mostly etch rate) under certain circumstances, the temporal and spatial information collected from sensors is mostly sacrificed in order to simplify the analysis. Moreover, few explicit physical interpretations have been made to relate optical emission intensities to equipment or wafer state, which limits the utility of these techniques.

2.2. Plasma Etch Process

Plasma etching is considered a bottleneck process in integrated circuit (IC) manufacturing lines. It usually involves complex reactions/collisions of a large number of energetic electrons, photons, ions and chemically reactive gas molecules. The success of a plasma etch process is determined by some complex tradeoffs in order to optimize each of a long list of process parameters, including etch rate, selectivity, damage, residue, microloading, and profile and linewidth control. Moreover, continued device scaling is making plasma etch even more difficult to manage. In order to control this “mysterious” process, comprehensive understandings of etch mechanisms and etcher designs are necessary.

2.2.1. Etch Mechanisms

There are four basic plasma etch mechanisms: sputtering, pure chemical etching, ion energy driven etching, and ion inhibitor etching [2.8]. *Sputtering* is a process of removing the substrate material by purely physical processes, as illustrated in Figure 2.1. Since sputter etching is purely physical and requires high ion energy and low pressure, it is the least selective mechanism. Sputtering rates are generally low because the yield is typically of order one atom per incident ion, regardless of the etched materials. However, sputtering is the most anisotropic etch process.

Chemical etching is simply due to the fact that gas-phase etchant atoms or molecules chemically react with the surface. Thus this process can be highly selective, and the etch rate can be quite large because the flux of etchants to the substrate can be high in processing discharges. However, the most important requirement for this kind of process is that a volatile reaction by-product has to be formed. Moreover, chemical etching is usually isotropic, which is undesirable for achieving good linewidth control.

Ion-enhanced energy-driven etching is a combined etching process with both etchant atoms (chemical etching) and energetic ions (ion-bombardment). Generally, the

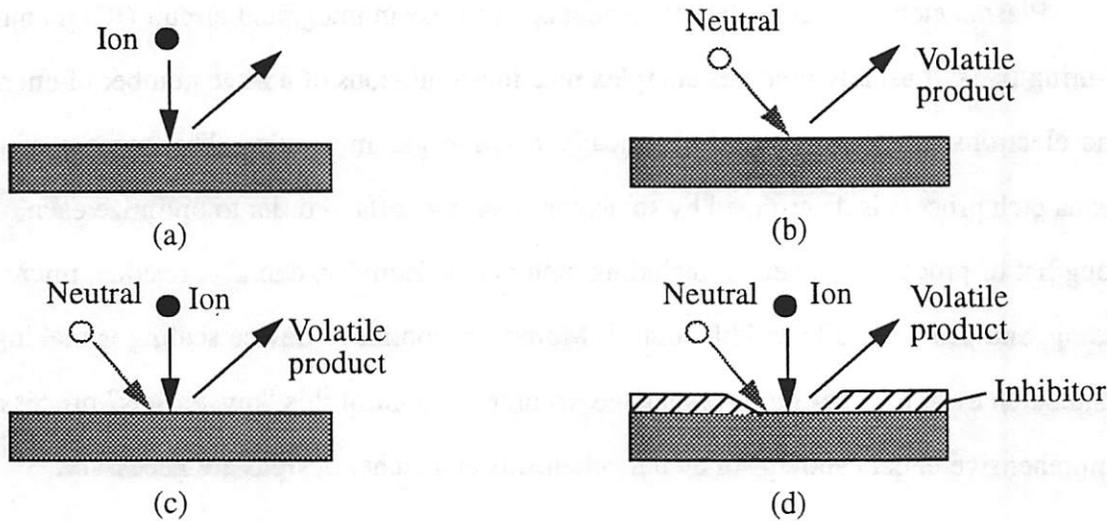


Figure 2.1 Four basic plasma etching mechanisms: (a) sputtering; (2) pure chemical etching; (c) ion energy-driven etching; (d) ion-enhanced inhibitor etching [2.9]

etch rate of this etch process is much larger than that of either pure chemical etching or sputtering alone, since the impinging ions can damage the surface so that it increases the etchant atom's reactivity significantly. The etch product must be volatile, as for pure chemical etching. Because the energetic ions have a highly directional angular distribution when striking the substrate, the etching can be highly anisotropic. However, ion-enhanced energy-driven etching may have poor selectivity compared to pure chemical etching. Thus the trade-off between anisotropy and selectivity is an important issue in designing this etch processes.

Ion-enhanced inhibitor etching, illustrated in Figure 2.1(d), is a conceptually more different etch mechanism, compared with the previous etch mechanisms. The discharge supplies etchants, energetic ions, and inhibitor precursor molecules that adsorb or deposit on the substrate to form a protective layer of polymer film. The ion-bombarding flux prevents the inhibitor layer from forming on horizontal surfaces so as to expose the surface to the chemical etchant. Where the ion flux is low, such as on the vertical feature sidewalls,

the inhibitors form a very thin film to protect the surface from the etchant, thereby making the process anisotropic.

Although four etch mechanisms have been distinguished, their use for a particular film etch often involves parallel or serial combinations of these mechanisms, as will be noted for the Al etch in the next section.

2.2.2. Etching Processes

The materials that need to be etched in a typical silicon-based IC generally fall into one of three areas: polysilicon, dielectrics or metals, each of which requires a dedicated etch process and equipment set-up. The difficulties of each etch process are primarily caused by the push to smaller features, but also by new transistor design, DRAM cell structures, isolation technologies, planarization and interconnect schemes. Figure 2.2 illustrates a typical cross-section of CMOS IC and its required etch processes. In the following sections, these three etch processes are described in detail.

2.2.2.1. Poly-Silicon Etching

Poly-Si can be etched in the chlorine discharge. For undoped poly-Si, the etch rate is very slow unless undoped poly-Si is etched by Cl in the presence of energetic ion bombardment. For heavily n-type doped poly-Si, it can be rapidly and spontaneously etched by Cl atoms. Thus, etched features are often severely undercut. This can be prevented by sidewall inhibitor chemistry. It is also important to note that Cl also attacks the doped poly-Si but at a much slower rate [2.9].

There are several critical requirements for poly-silicon etching that must be addressed. First, selectivity to the oxide is extremely important for poly-silicon etching,

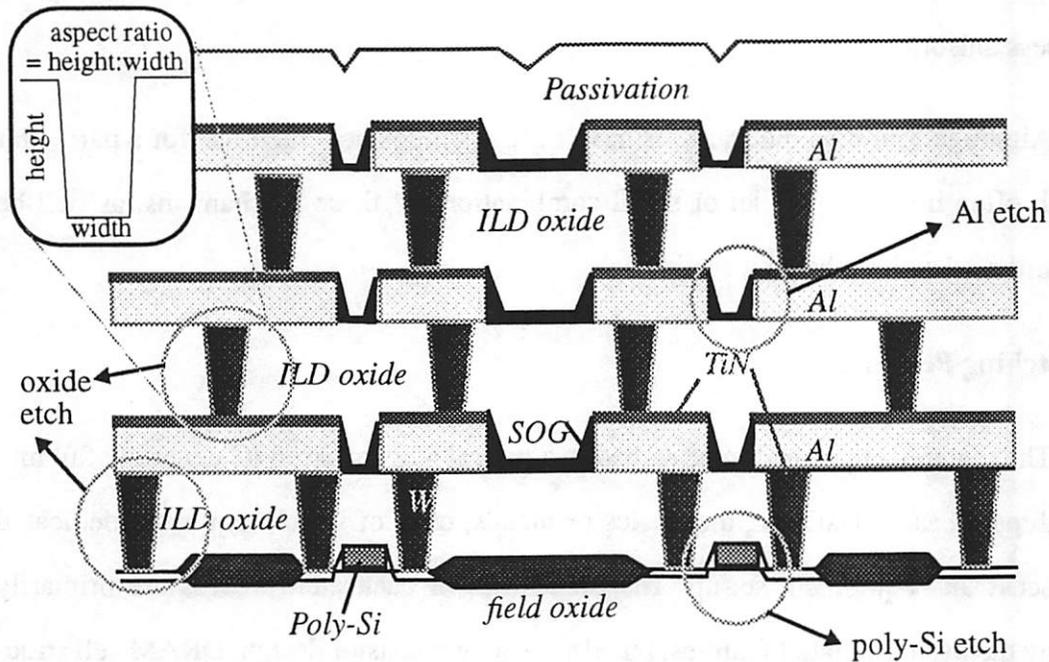


Figure 2.2 Typical cross-section for logic CMOS IC

because the etch must not punch through the thin gate oxide, keeping the oxide loss small, even with a 50-100% overetch. Secondly, the gate profile control is also vital. The modern gate design usually includes tungsten silicide over poly-Si application. This type of structure requires two different materials to be etched while maintaining all the important parameters such as profile control, minimal undercutting, and selectivity between layers and the substrate.

2.2.2.2. Oxide Etching

The etching of SiO_2 is different from poly-Si etching because the silicon-oxygen bond is about twice as strong as the silicon-silicon bond that is broken in poly-Si etching¹. Thus, to break the silicon-oxygen bond, ion and neutrals are directed at the wafer surface with more than twice as much as power as in other types of etch systems. Although F atoms

1. The silicon-silicon bonding is 78 kcal/mole, whereas the silicon-oxygen bonding is 190 kcal/mole.

are known to etch SiO_2 , the pure chemical etch rate based on F atoms is slow and therefore impractical in a real etch system. As such, a more complex fluorocarbon etch chemistry is commonly used for oxide etching. It is known that CF_x radicals do not spontaneously etch SiO_2 because these radicals do not dissociatively absorb in SiO_2 . Therefore, no pure chemical etching of SiO_2 takes place in fluorocarbon plasma, and all observed SiO_2 etching is ion-energy driven [2.8][2.9]. The SiO_2 etching is anisotropic, and the etch rate correlates with the ion bombarding energy and is independent of the substrate temperature [2.8]. High selectivity can be achieved for CF_x radical etchants produced using unsaturated fluorocarbon feedstocks or by adding hydrogen to saturated feedstocks [2.8]. In both cases, the protective polymer layer is formed on the surfaces other than silicon oxide (the oxygen released in the etching of SiO_2 prevents the formation of the polymer)[2.10].

In the area of oxide etching, the most demanding application is high aspect-ratio contact/via etch (as shown in Figure 2.2). It is due to the fact that contacts have different depths at a single masking layer, so the shallow contact clears first while the deep contact continues to etch. Another challenge of oxide etching is to obtain high selectivities to the photoresist masking layer as well as the underlying material once it is exposed. Because the modern high-density plasma is very effective at dissociating the feedstock gases, which creates excessive amounts of aggressive fluorine atoms, free fluorine tends to rapidly eat away organic photoresist and thereby reduce the photoresist selectivity.

Another problem in oxide etch is the etch-stop, which means the etching action slows or stops altogether for some unknown reasons [2.10]. This problem is even more severe for very small and deep features. One hypothesis to explain etch-stop is that there is an imbalance between polymer formation, passivation, and ion-bombardment processes [2.10] so that the etch process stops due to the excess of polymer formation.

2.2.2.3. Metal Etching

Today, metal etching still means mostly aluminum etching, although tungsten is often used for contact and via fill. Aluminum etching is essentially a chemical-reaction driven process [2.9][2.11]. However, since aluminum is not etched in fluorine atom plasma (because aluminum fluoride is completely involatile), aluminum has to be etched in atomic and molecular chlorine or bromine (both aluminum chlorides and aluminum bromides are volatile). It is also important to note that ion bombardment has a limited effect on the aluminum etching rate because the spontaneous chemical attack is very fast [2.9]. Therefore, the *inhibitor* chemistry is necessary for linewidth control in aluminum etching. In the case of Cl_2/BCl_3 etchant, the aluminum layer is etched by chlorine atoms and molecules while the addition of BCl_3 is used to provide the inhibitor chemistry that enhances the etch anisotropy [2.9]. Moreover, since oxygen and water are desorbed on the surface of aluminum film which will seriously interfere with the etching process, the addition of the BCl_3 can react preferentially with H_2O and O_2 in the process chamber, thereby providing reproducible processing when added to Cl_2 [2.9]. As a result, plasmas based on BCl_3/Cl_2 mixtures have become the standard for aluminum etching [2.11]. After aluminum etching, a timed over-etch process is usually performed to ensure the aluminum film has been cleared from the wafer. The oxide loss due to the over-etch process is primarily driven by the ion-bombardment mechanism [2.8].

One challenge with aluminum etching is that there are varying amounts of copper added to the aluminum (to prevent the “spiking” effect). This makes etching more difficult because copper chlorides have low volatility. Photoresist degradation is another problem in chlorine-based plasma since the etch product, i.e., AlCl_3 , accelerates degradation of resist. In addition, etch rate microloading (also known as aspect-ratio dependent etch) is of

increasing importance in metal etch, especially for device geometries shrinking below half-micron [2.12]. The Al etch rate microloading is a tradeoff against Al etch rate. In an etchant deficient environment, Al etch rate in the open area is slower than that in the dense area, due to the chemical loading effect. When the etchant concentration is sufficient, the etch rate in the open area is enhanced, while the etch rate in the dense area is limited by species transportation.

Table 2.1 summarizes the typical feed gas and mechanisms for plasma etching various materials.

Table 2.1
Typical feed gas and mechanisms for plasma etch [2.9]

| Materials Etched | Source gas | Additive | Mechanism |
|-------------------|-----------------|---------------------------------------|----------------------------|
| Poly-Si (n-doped) | Cl ₂ | None C ₂ F ₆ | Chemical Ion-inhibitor |
| SiO ₂ | CF _x | H ₂ | Ion-enhanced energy-driven |
| Al | Cl ₂ | BCl ₃ | Ion-inhibitor |

2.3. Plasma Etch Equipment

Today almost all commercial plasma etchers are single wafer systems, that is, the system presents an identical environment to every wafer, compared with the batch wafer systems in which the etch conditions are varied dependent on the wafer locations. Although there are many variations of chamber configuration with major differences in electrode structure for commercial single-wafer plasma etcher, two most commonly used plasma reactors are introduced in this section: capacitive planar etcher and high-density discharge etcher.

2.3.1. Capacitive Planar Etcher

Capacitively driven RF discharge (also called RF diodes) etchers are commonly used for materials processing. The basic geometry consists of a vacuum chamber containing two planar electrodes separated by a spacing l and driven by an RF power source, as illustrated in Figure 2.3 (a). When operated at low pressure, with the wafer placed on the powered electrode for removing substrate materials, this reactor is commonly called reactive ion etcher (RIE). In this case, the etching is a chemical process enhanced by energetic ion bombardment of the substrate. Table 2.2 lists the typical range of parameters for a capacitive RF discharge etcher.

However, the crucial limiting feature of capacitive RF discharge etcher is that the ion-bombarding flux and bombarding energy cannot be varied independently. This is similar to the lack of independent voltage and current control in semiconductor pn junction. To accomplish the etching performance needed for the deep sub-micron devices, the ion-bombardment energy has to be controllable independently of the ion/neutral flux [2.8], which presents more options to the process engineers to optimize the etching conditions. This

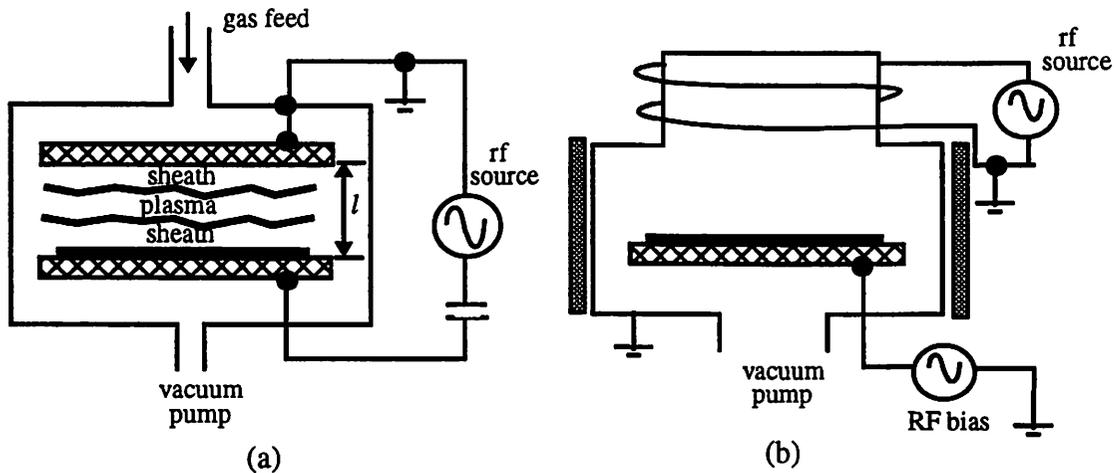


Figure 2.3 The illustrations of (a) capacitive RF discharges in plane parallel geometry; (b) high-density inductively coupled plasma.[2.8]

requirement leads to the development of high-density plasma etcher, as described in the next section.

Table 2.2
Range of parameters for RIE and high-density discharge[2.8]

| Parameter | RIE | High-Density |
|------------------------------------|------------------------------------|------------------------------------|
| Pressure (mtorr) | 10-1000 | 0.5-50 |
| Power (W) | 50-2000 | 100-5000 |
| Frequency (MHz) | 0.05-13.56 | 0-2450 |
| Plasma Density (cm ⁻³) | 10 ⁹ -10 ¹¹ | 10 ¹⁰ -10 ¹² |
| Electron temperature (V) | 1-5 | 2-7 |
| Ion acceleration energy (V) | 200-1000 | 20-500 |
| Fractional ionization | 10 ⁻⁶ -10 ⁻³ | 10 ⁻⁴ -10 ⁻¹ |

2.3.2. High-Density Plasma Etcher

The low pressure plasma source is desirable in the modern etcher design because it is difficult to diffuse etchants in and reaction byproducts out of openings that are smaller than 0.25 μ m. The solution is to go to lower pressures where the mean free paths of gas ion and molecules are longer, which also reduces scattering collisions that can cause low profile control [2.13]. These needs led to the development of the new generation of high-density plasma source, which is capable of generating enough ions to achieve acceptable etch rate at reduced pressure (high etch rate is important in single wafer systems for reasonable throughput). Nonetheless, as mentioned in the last section, the need of high-density plasma etch is not only driven by a need to operate at lower pressure or with high ion density, but rather the need to have independent control over the ion bombarding energy and the ion and neutral density [2.8][2.10][2.13]. This “plasma decoupling” allows better control of the various etch mechanisms (*e.g.*, physical ion-bombardment and chemical etching) that competes with each other in plasma etching.

To achieve a high density plasma source, a common feature is that RF is coupled to the plasma across a dielectric cylindrical wall, rather than by direct connection to an electrode in the plasma [2.8], as for the RIE etcher. This noncapacitive power transfer is the key to achieving a high ion-density efficiently. To control the ion energy, the electrode on which the wafer is placed can be independently driven by a capacitively coupled RF source. Hence, an independent control of the ion/radical fluxes (through the source power) and the ion-bombarding energy (through the wafer electrode bias power) is possible. There are many different high-density plasma source designs in terms of the way by which power is coupled to the plasma. The examples include electron cyclotron resonance (ECR), helical resonator (HR), and inductive (or transformer) coupled plasma source (ICP or TCP)[2.8]. Among these different designs, ICP (or TCP) is the most commonly used high-density plasma source in current fabrication lines. An inductive coupled plasma etcher is shown in Figure 2.3(b). In this case, plasma acts as a single-turn, lossy conductor that is coupled to a multiturn nonresonant RF coil across the dielectric discharge chamber, where RF power is inductively coupled to the plasma by transformer action. In contrast to ECR, a dc magnetic field is not required in this efficient power coupling design [2.8].

Despite the obvious needs for developing high-density plasma sources with low pressure, high ion/radical fluxes, and controllable ion energies, there are many other issues needed to be resolved. The first critical issue is to achieve the required process uniformity over 200 to 300mm wafer diameters. In contrast to the nearly one-dimensional geometry of a typical RIE etcher, high-density cylindrical sources can have length-to-diameter ratios of order or exceeding unity [2.8]. Plasma formation in such geometries is inherently radically nonuniform. The second critical issue is to achieve an efficient power-coupling across a dielectric wall over a wide operating range of plasma parameters. The third issue is that the degradation of and deposition on the wall can lead to irreproducible source behavior and the need for frequent costly cleaning cycles. The fourth issue is that the low-pressure opera-

tion usually leads to severe pumping requirements for high etching rate, and thus to the need for larger and costlier vacuum pumps.

For more detailed and quantitative description of these plasma etchers, one can refer to Chapter 11 and 12 in the book of “Principles of Plasma Discharges and Materials Processing” by Lieberman and Lichtenberg [2.8].

2.4. In-Situ Sensors for the Etch Process

Barna et al. [2.14][2.15] state that any IC manufacturing tool requires three distinct types of sensors to completely characterize the tool and its effect on the wafer: machine-state (e.g., RF power, gas flow, pressure), process-state (e.g., plasma density, delivered power), and wafer-state (e.g., etch rate, film thickness) sensors. Of these three sets, machine-state parameters are the easiest to monitor and control. Usually, a closed-loop controller continuously adjusts machine parameters to preset targets. These types of sensor readings are available in most modern etchers. In contrast, process-state sensor parameters are dependent on both machine-state parameters and uncontrolled perturbations in the system. This type of sensor has to be *in-situ*, real-time and non-intrusive. Wafer-state sensors are often the most direct way to monitor the effect of the process on the wafer. However, due to the scarcity of *in-situ* and real-time wafer-state sensors, an *ex-situ* and post-process approach is typically used in manufacturing sites. Consequently, such sensor data can only be used to provide run-to-run control of the process. The typical approach being pursued for *in-situ* control is to characterize the etch process using process state sensors, rather than wafer-state sensors. A major roadblock to this type of real-time and *in-situ* control is the complicated or sometimes, unclear relationship between process-state and wafer-state parameters.

There are many different types of *in-situ* commercially available sensors that could be used for monitoring the parameters specified in plasma etch [2.14], as listed in Table 2.3.

Table 2.3
Commercially available sensors for plasma etch (modified from Ref.[2.14])

| Monitoring Parameters | Sensor | Comments |
|-------------------------------|---|---|
| Equipment state | | |
| Chamber/substrate temperature | Thermocouple | All these machine-state sensors are commonly installed on the commercial etchers by equipment vendors |
| Pressure | Ionization gauge | |
| Gas flow | Mass flow controller | |
| rf power and coil power | rf sensors | |
| bias voltage | rf sensors | |
| Process state | | |
| Reactant concentration | Optical emission Laser-induced fluorescence Mass spectroscopy | Standard technique Not robust enough for production, Installation expensive |
| Endpoint detection | Optical emission Residual gas analysis (RGA) | Routinely used Time delay, expensive |
| Ion density | Langmuir probe | Intrusive |
| rf plasma parameters | I/V/phase at multiple frequencies | Difficult to correlate to specific machine/wafer/process states |
| Wafer temperature | Thermocouple Pyrometer Fluorescence | Wafer temperature inferred Emissivity and light interference Intrusive probe |
| Wafer state | | |
| Etch rate (thickness) | Ellipsometer Interferometer | Widely implemented on R&D basis |
| Critical dimension (CD) | Scatterometer | Complex implementation |

2.5. Optical Emission Spectroscopy

Optical emission spectroscopy (OES) is the most widely used optical analytical and diagnostic instrument in plasma-assisted processing, where gas-phase species are promoted to excited electronic states by collisions with energetic electrons. The emission processes in OES is illustrated in Figure 2.4. It is important to note that only excited species in plasma can be detected by OES, and thus the observed spectrum only gives information about the excited-state density and does not directly reflect the ground-state population. The densities

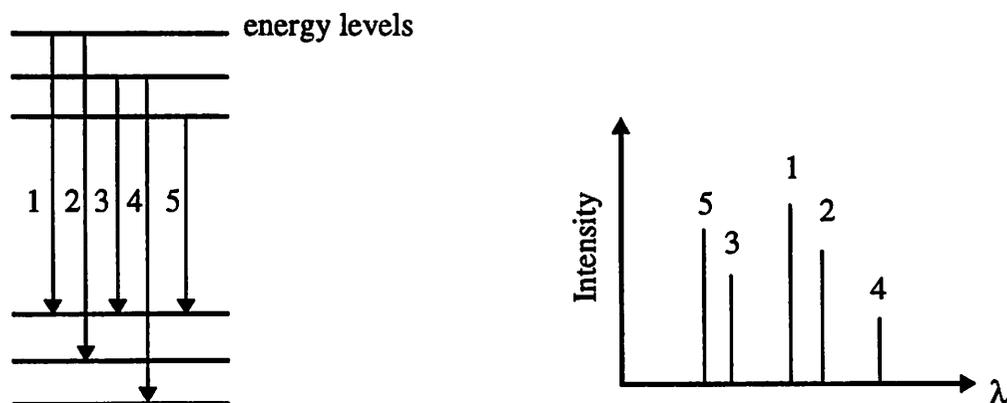


Figure 2.4 Schematic diagram of atomic emission spectrum

of excited-state species are typically $< 10^{-4}$ of the ground state density [2.8][2.9]. Furthermore, while emission from specific intermediates and products may dominate optical emission spectra, emission from the chemically dominant species (e.g., etchant) and interesting species (e.g., etch byproduct) may not be observable. Still, when such emission is present, OES has proven to be a very powerful tool and is routinely used to monitor gas-phase population during process development and for endpoint detection during IC manufacturing.

2.5.1. Quantitative Description of Optical Emission

Based on the assumption that the excited state is formed solely by electron impact excitation and that this excitation occurs in a single step from the ground state, and that emission must occur once the state is excited, optical emission intensity is written [2.9] as

$$I(\lambda_{ij}) = NP_{ij}A_{ij}(\lambda_{ij})K \quad (2.1)$$

where i indicates the ground state, j indicates the excited state, λ_{ij} is the transition wavelength between state i and state j , N is the ground state density, A_{ij} is the Einstein emission

probability, K is a correction factor which describes the effect of view volume and alignment, and P is the electron impact excitation function which represents the probability of exciting the state j by electron impact, starting from the ground state. P is a complex function of electron temperature T_e , and is given by [2.8]

$$P = \int_0^{\infty} 4\pi v_e^2 dv_e \sigma_{\lambda}(v_e) v_e f_e(v_e, T_e, n_e) \quad (2.2)$$

where v_e is the electron velocity, σ_{λ} is the cross section for emission of a photon of wavelength λ due to electron impact excitement, and f_e is the electron distribution function which depends on electron temperature and electron density [2.8]. It is also important to note that the electron temperature is a monotonically increasing function of $1/Pr$, where Pr is the pressure [2.8][2.9]. Among these variables in Eq.(2.1), only the ground state density N and the electron impact excitation function P are dependent on the machine settings, and only N is wavelength-independent.

2.5.2. OES Instrumentation

Optical emission spectroscopy involves the collection, spectral dispersion, and detection of light. Typical experimental setups for OES are shown in Figure 2.5. Emission from a specific volume in the plasma chamber is imaged onto the entrance slit of a spectrometer by a series of UV-grade fused-silica lens and optical fiber.

The light can be dispersed by a diffraction grating system. One common design for grating spectrometer is Czerny-Turner arrangement [2.16], as shown in Figure 2.5. A curved mirror images the source from the entrance slit onto a parallel beam that is directed to the diffraction grating. A second curved mirror focuses that diffracted beam onto the exit

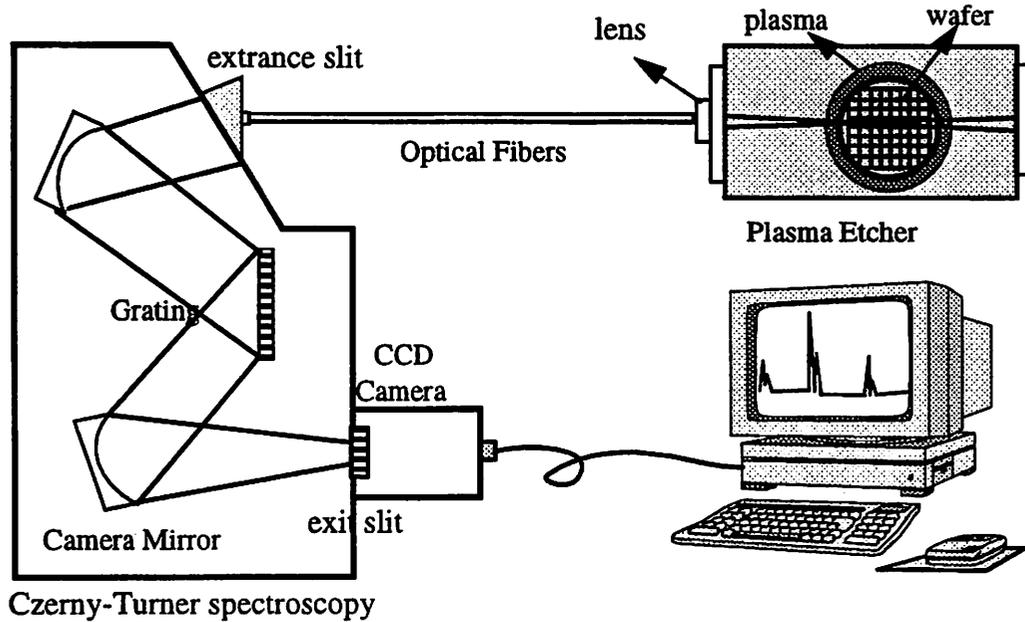


Figure 2.5 Schematic illustration of the OES instrumentation

slit. The diffraction grating is commonly a square, of length L , with ν grooves per millimeter that run normal to the plane of incidence of light. The groove density determines the spectral dispersion, which is given by a grating equation [2.17]

$$m\lambda = d(\sin\theta_m - \sin\theta_i) \quad (2.3)$$

where m is the diffraction order, d is the groove separation, and θ_m and θ_i are the angles of incidence and diffraction. The grooves are blazed to maximize first-order diffraction ($m=1$) at a specified wavelength. The spectral resolution due to dispersion is given by [2.17]

$$\Delta\lambda = \frac{w}{m\nu f} \cos\theta_m \quad (2.4)$$

where w is the exit slit widths, and f is the spectrometer focal length which is the distance from the exit plane to the last focusing mirror (also the distance from the entrance plane to the first focusing mirror). It is also important to note that the efficiency of a grating system and optical fiber is a function of wavelength. Therefore, the optical emission intensity detected by OES is rewritten as

$$I(\lambda_{ij}) = NP_{ij}A_{ij}Q(\lambda_{ij})K \quad (2.5)$$

where $Q(\lambda_{ij})$ is a correction factor to account for the apparatus collection efficiency at the emission wavelength λ_{ij} .

The dispersed light can be detected by a parallel detector such as a photodiode array, or a charged-coupled device (CCD) array detector. This technique can detect a wide spectral range without scanning the grating, thereby achieving a rapid data acquisition. However, one potential disadvantage in using a parallel detection scheme is relatively low spectral resolution when a broad spectral region is being monitored at one time. Array detectors often have 1000 to 500 pixels limiting the spectral resolution of an OES system. That is, the spectral resolution of OES system is often limited by the light detector, instead of the grating's groove density. For example, if the spectral range 250-750nm is monitored with a 1000-channel detector, the maximum resolution is about 0.5nm, which may not be satisfactory to reveal enough etch information. The spectral resolution can be improved by examining a smaller spectral range with the same detector, or using a high-density CCD array [2.16].

2.5.3. Actinometry

As mentioned, the observed intensity of a spectral line is not comparable to a measurement of the concentration of the chemical species associated with this spectral line.

Therefore, actinometry [2.18] has been used to determine the quantitative concentration of interesting species. Actinometry involves the addition of a known amount of a gas phase species such as Ar to the discharge. The assumptions of actinometry are that the excited state is formed only by single step electron excitation from the ground state, that the electron velocity dependence of the cross section for this excitation is similar to that of the species of interest (see Eq. (2.2)), that the wavelengths of the two lines are similar, and that the emission probabilities are truly constant. Under these circumstances, the variation of the electron energy distribution function with features of the discharge like pressure and power can be compensated by taking the ratio of emission of the discharges of interest, Z, to that of the inert gas, M:

$$\frac{I_Z}{I_M} = \frac{N_Z P_Z Q(\lambda_Z) A_Z K}{N_M P_M Q(\lambda_M) A_M K} \equiv C \cdot \frac{N_Z}{N_M} \quad (2.6)$$

If N_M is known and I_Z and I_M are measured, an absolute value of N_Z can be determined. Even if C is not known, the relative variation of N_Z with variation of discharge parameters can be found.

2.6. Multivariate, Statistical Data Compression

Since typical sensor data sets contain many correlated variables (for example, one spectrum contains about a thousand correlated variables), it is not practical to use all of them directly. Therefore, several statistical data-compression and multivariate regression techniques are presented here to extract valuable information from these multivariate sensor signals and relate these reduced data set to the wafer states.

2.6.1. Principal Component Regression

A statistical method called Principal Component Analysis (PCA) is often used to create a few statistically significant, uncorrelated weighted sums from the original corre-

lated data [2.19]. PCA is used to explain most of the variance in the original data set by only a few principal components (PCs). PCA transforms the input variables to a set of orthogonal variables known as principal components, which are linear combinations of the original variables [2.19]. Specifically, PCA decomposes a matrix X ($I \times J$), where there are J measurements variables for I objects, into a series of R principal components

$$X = \sum_1^R z_r p_r^T \quad (2.7)$$

where p_r is the loading vector and z_r is a score vector (*i.e.*, the coordinates in the transformed space). Since the covariance matrix of the original input matrix X is symmetric, it can be decomposed into $X^T X = p \Lambda p^T$, where the diagonal elements of the Λ are the eigenvalues which summarize the variances explained by the PCs, and the columns of p are the eigenvectors of the covariance matrix of X . The coefficients of the original variable are the eigenvectors p or loadings. The PCA transformation is often

$$Z = p^T X \quad (2.8)$$

where X has been scaled by subtracting from each column its mean and dividing by its standard deviation. The transformed variables, Z , are the principal components of X or t -scores. The i th principal component is

$$z_i = p^T x \quad (2.9)$$

The loading vectors p provide the directions with maximum variability. Typically, the first few principal components can explain most of the variation of input data set (*i.e.*, X). The number of principal components retained in the model is sometimes determined by a “scree” plot, which characterizes the proportion of total population variance explained by the first R principal components where R is from 1 to the total number of observations.

Another way to understand PCA is to illustrate it by a simple case, as shown in Figure 2.6. The original axes x_1 and x_2 are transformed to a new orthogonal principal component axes z_1 and z_2 . Because z_1 explains most of the variation, we can sometimes visualize all data points are along one axis (i.e. in a one dimensional space).

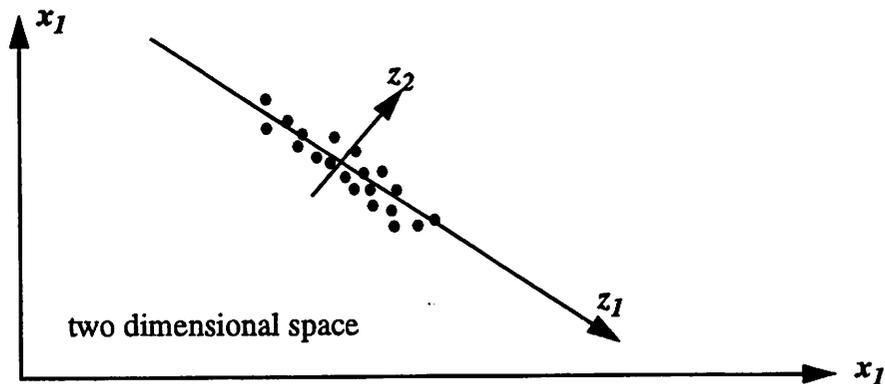


Figure 2.6 Principal component transformation in two dimensional space (• denotes data point)

Once this principal component analysis is complete, the reduced PC variables can be used as the input matrix for regression. This is called Principal Component Regression (PCR). An overview of these procedures is shown in Figure 2.7.

2.6.2. Partial Least Squares Regression

Like Principal Component Analysis, Partial Least Square Analysis (PLSA) is also based on projecting the information in high-dimensional data spaces down onto low-dimensional spaces, defined by a small number of eigenvectors, known as latent variables[2.20][2.21]. These new latent variables summarize much of the important information contained in the original data set. However, in the PLSA algorithm, both the response

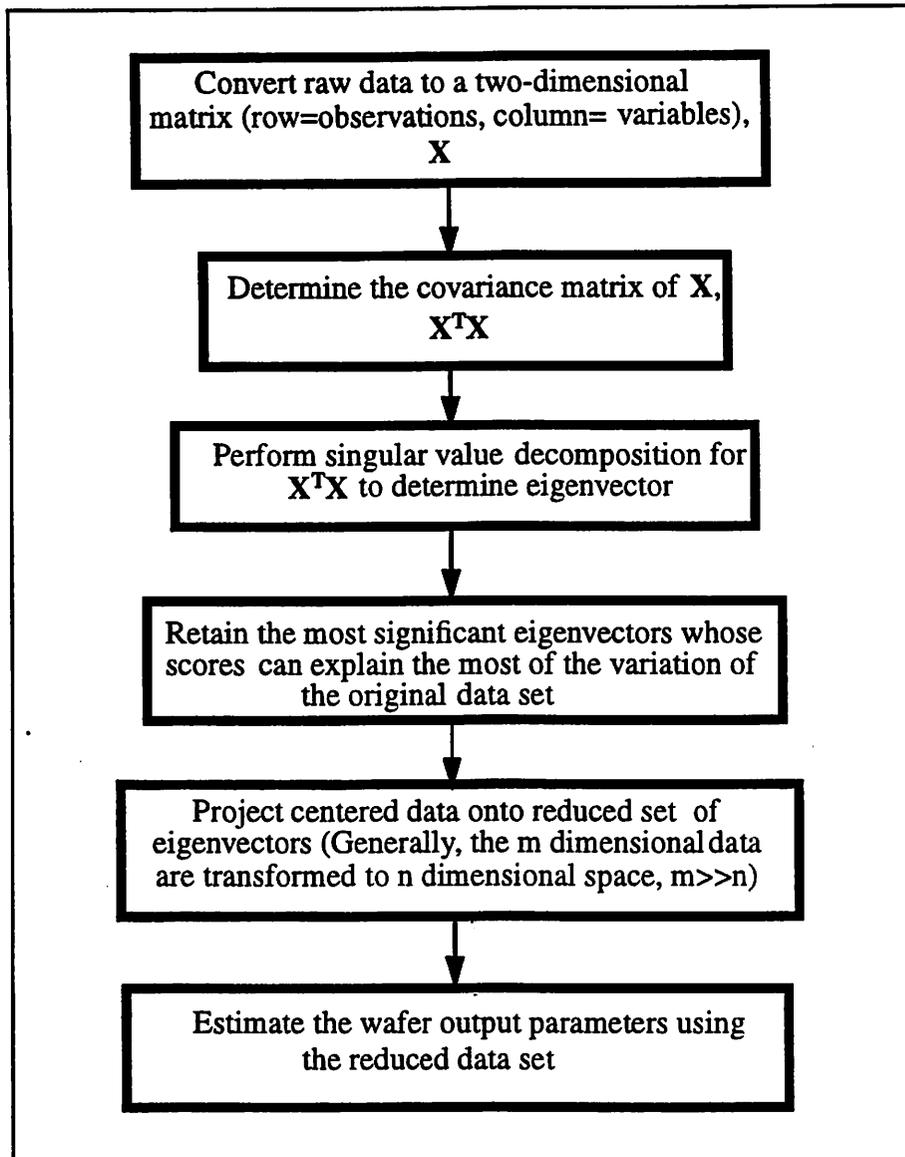


Figure 2.7 The algorithmic flow of Principal Component Regression

variables (Y) and input variables (X) are used. The scaled and mean-centered X and Y can be decomposed as:

$$X = \sum_{a=1}^A t_a P_a^T + E \quad (2.10)$$

$$Y = \sum_{a=1}^A t_a q_a^T + F \quad (2.11)$$

where t_a is the latent vector, E and F are the residuals, and w_a and q_a are loading vectors whose elements w_{aj} and q_{aj} express the contribution of each variable x_j and y_j , respectively. A common PLSA algorithm is summarized below and illustrated in Figure 2.8:

Step 1: Set u equal to any column of Y .

Step 2: Regress columns of X on u to get loadings: $w^T = u^T X / u^T u$.

Step 3: Normalized w to unit length.

Step 4: Calculate scores $t = Xw / w^T w$.

Step 5: Regress the columns of Y on t : $q^T = t^T Y / w^T t$.

Step 6: Calculate the new score vector for Y : $u = Yq / q^T q$.

Step 7: Check convergence of u : if yes go to 8; if no go to 2

Step 8: Calculate X matrix loadings by regressing columns of X on t : $p^T = t^T X / t^T t$.

Step 9: Calculate residual matrices: $E = X - tp^T$; $F = Y - tq^T$

The total number of PLS components (*i.e.*, A) needed to extract the information from X and Y is usually low (typically 2 to 5 in our application) and can be estimated by cross-validation in order to prevent model overfitting. This is done by minimizing a criterion called the prediction error sum of squares (PRESS) [2.22] defined as

$$PRESS = \sum_{i=1}^m (\hat{Y} - Y)^2 \quad (2.12)$$

where m is the total number of variables used in the model. The first few dimensions are usually the most important and dominate the model. Hence, the final model of Eq.(2.11) can be expressed in terms of the X data as the regression model:

$$Y = X\beta + F \quad (2.13)$$

where the matrix of regression coefficients is given by:

$$\beta = W(p^T W)^{-1} Q^T \quad (2.14)$$

where W , P , and Q are the matrices whose columns are the vectors w_a , p_a and q_a ; $a=1,2,\dots,A$, as defined above.

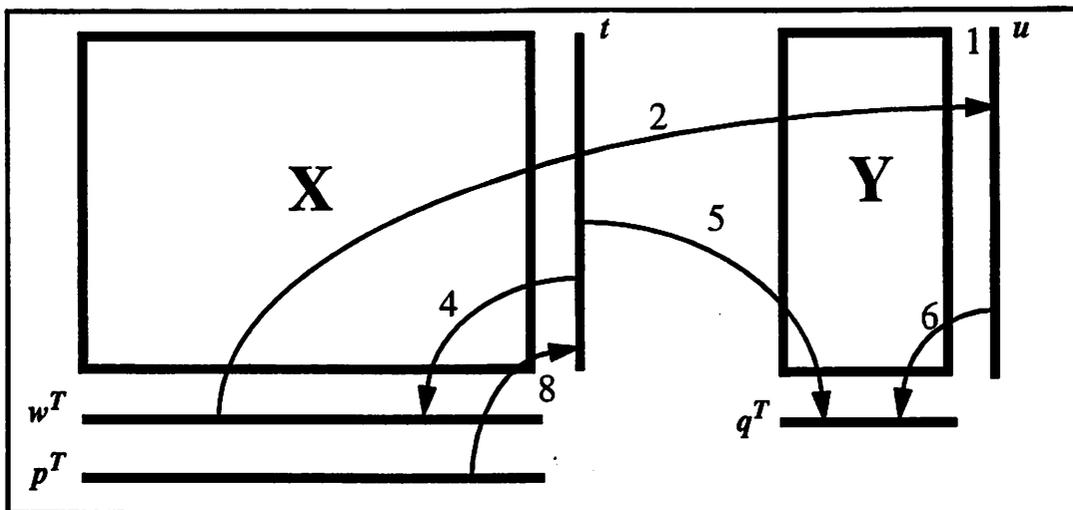


Figure 2.8 PLSA iteration

References for Chapter 2

- [2.1] S. F. Lee and C. J. Spanos, "Prediction of Wafer State After Plasma Processing Using Real-time Tool Data", *IEEE Transaction on Semiconductor Manufacturing*, vol. 8, no. 3, 1995, pp. 252-261.
- [2.2] S. F. Lee et al., "RTSPC - A Software Utility for Real-Time SPC and Tool Data Analysis", *IEEE Transaction on Semiconductor Manufacturing*, vol. 8, no. 1, 1995, pp. 17-25.
- [2.3] G. S. Selwyn, *Optical Diagnostic Techniques for Plasma Processing*, AVS Press, New York, 1993.
- [2.4] R. Shadmehr et al, "Principal Component Analysis of Optical Emission Spectroscopy and Mass Spectrometry: Application to Reactive Ion Etch Process Parameter Estimation Using Neural Networks", *J. Electrochem. Soc.*, vol. 139, no. 3, March, 1992, pp. 907-914.
- [2.5] M. P. Splichal and H. M. Anderson, "Application of Chemometrics to Optical Emission Spectroscopy for Plasma Monitoring", *SPIE*, vol. 1594, 1991, p.1991.
- [2.6] H. M. Anderson and M. P. Splichal, "An Integrated System of Optical Sensors For Plasma Monitoring and Plasma Process Control", *SPIE*, vol. 2091, 1994, p.333.
- [2.7] D. A. White, et al., "Spatial Characterization of Wafer State Using Principal Component Analysis of Optical Emission Spectra in Plasma Etch", *IEEE Transaction on Semiconductor Manufacturing*, vol. 10, no. 1, 1997, pp. 52-61.
- [2.8] M. Lieberman and A. Lichtenberg, *Principles of Plasma Discharges and Materials Processing*, John Wiley and Sons, New York, 1994.
- [2.9] D. Manos and D. Flamm, *Plasma Etching*, Academic Press, London, 1989.
- [2.10] P. Singer, "The Many Challenges of Oxide Etching", *Semiconductor International*, June, 1997, pp. 109-114.
- [2.11] P. E. Riley, et al. "Plasma Etching of Aluminum for ULSI circuits", *Solid State Technology*, vol. 36, no.2, Feb., 1993, pp. 47-55.
- [2.12] D. Ma et al, "Metal Etch Technology for Sub-Half-Micron Devices", *semiconductor International*, July, 1996, pp. 181-188.
- [2.13] P. Singer, "New Frontiers in Plasma Etching", *Semiconductor International*, July, 1996, pp. 152-164.

Chapter 2

- [2.14]G. Barna, et al., "Sensor Needs for IC Manufacturing", *Solid State Technology*, April, 1994, pp. 57-61.
- [2.15]G. Barna, et al., "Dry Etch Processes and Sensors", *Solid State Technology*, January, 1994, pp. 47-53.
- [2.16]I. P. Herman, *Optical Diagnostics For Thin Film Processing*, Academic Press, San Diego, 1996.
- [2.17]R. Petit, *Electromagnetic theory of gratings*, Springer-Verlag, New York, 1980.
- [2.18]J. Coburn and M. Chen, *J. Vac. Sci. Technol.*, vol. 18, 1981, p.683
- [2.19]J. E. Jackson, *A User's Guide to Principal Components*, John Wiley and Sons, New York, 1991.
- [2.20]I. Frank, J. Friedman, "A Statistical View of Some Chemometrics Regression Tools", *Technometrics*, Vol. 35, N.2, May 1993, pp. 109-148.
- [2.21]J. MacGregor, C. Jaeckle, C. Kiparisside and M. Koutoudi, "Process Monitoring and Diagnosis by Multiblock PLS Methods", *Process System Engineering*, Vol. 40, N. 5, May 1994, pp.826-838.
- [2.22]D. M. Haaland and E. V. Thomas, "Partial Least-Squares Methods for Spectral Analysis", *Analytical Chemistry*, Vol. 60, No. 11, June 1988, pp. 1193-1217.

CHAPTER 3

Wafer-State Modeling Using OES for Oxide Etch

3.1. Introduction

As indicated in Section 2.4, the lack of in-situ and real-time wafer-state sensors leads to an effort of developing reliable models that relate the wafer-state measurements (*e.g.*, etch rate or etching uniformity) to variations in real-time process-state sensor readings so that the process-state sensor can be used as a “counterpart” of the wafer-state sensor. These models, so called Wafer-State Models, are needed in order to predict etch behavior under a wide range of operating conditions to a high degree of precision. In this chapter, we present a methodology to estimate several important post-etching wafer characteristics for the interconnect dielectrical etching process using OES sensor readings (see Figure 3.1). All data acquisitions, measurements, and wafer fabrication were performed on state-of-the-art plasma etchers in a high-volume IC production site.

One challenge worth noting is that the spectral data are difficult to interpret using conventional multivariate linear regression because of the large number of variables involved. Consequently, Principle Component Analysis (PCA) and, alternatively, Partial Least Squares (PLS) have been employed to compress the spectral data, and then relate the reduced data set to the etching performances.

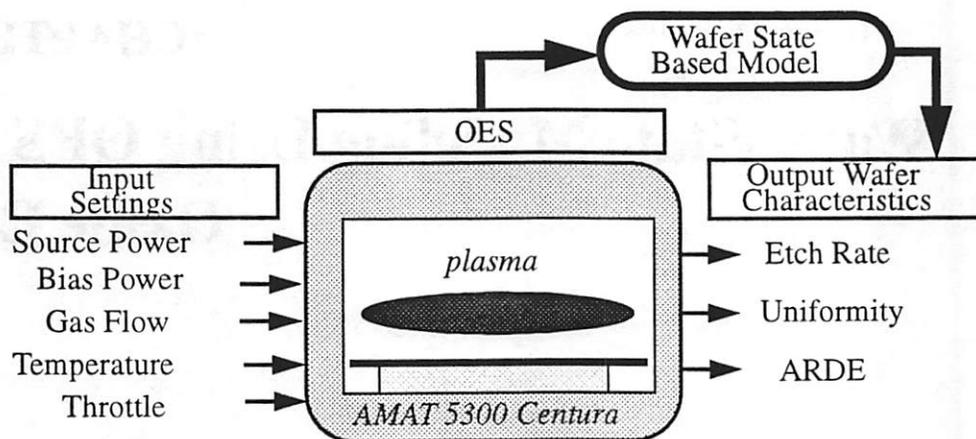


Figure 3.1 Wafer state modeling scheme using OES sensor readings.

3.2. Experiment to Link OES Signals to Oxide Etch Variability

3.2.1. Data Acquisition

An SC Technologies optical emission spectrograph and controller were installed on the Applied Materials 5300 Centura dielectric high-density plasma (HDP) etching chamber at AMD's Fab-25. An optical fiber was mounted on the reactor viewport and directed toward the entrance slit of an image spectrograph, as illustrated in Figure 3.2. The grating system dispersed the incident light and projected the spectrum onto a photodiode sensor array. Because there are 501 pixels on the photodiode array, we had the capability to monitor the intensities of 501 different wavelengths simultaneously during the etching

process. The control of this system and the subsequent data conversion and analysis were done on an IBM 486 PC and Sun SPARC-20 workstation, respectively. The S-plus commercial software package was used for the statistical analysis. Since the intensity ratio between the main C and F lines is strongly related to the etching performance [3.3], the spectral data was acquired in the 200 to 500nm range in order to ensure that most of the C_xF_y spectral features were monitored. In this study we used the 600 groove/mm grating with a 0.6nm/pixel resolution and an acquisition time of less than one second per frame.

For each wafer, five spectral frames were collected throughout the 150 seconds of the main etch step. The signals collected during the main etch step fluctuated whenever the RF power was applied or removed. The unstable features caused by these transient effects must be removed before any further statistical analysis. In this work, the OES data sets were selected by choosing the third spectrum which was collected at about the 80th second of processing for each wafer, that is, only one spectrum was used for each wafer (i.e., sampling rate is only one per wafer).

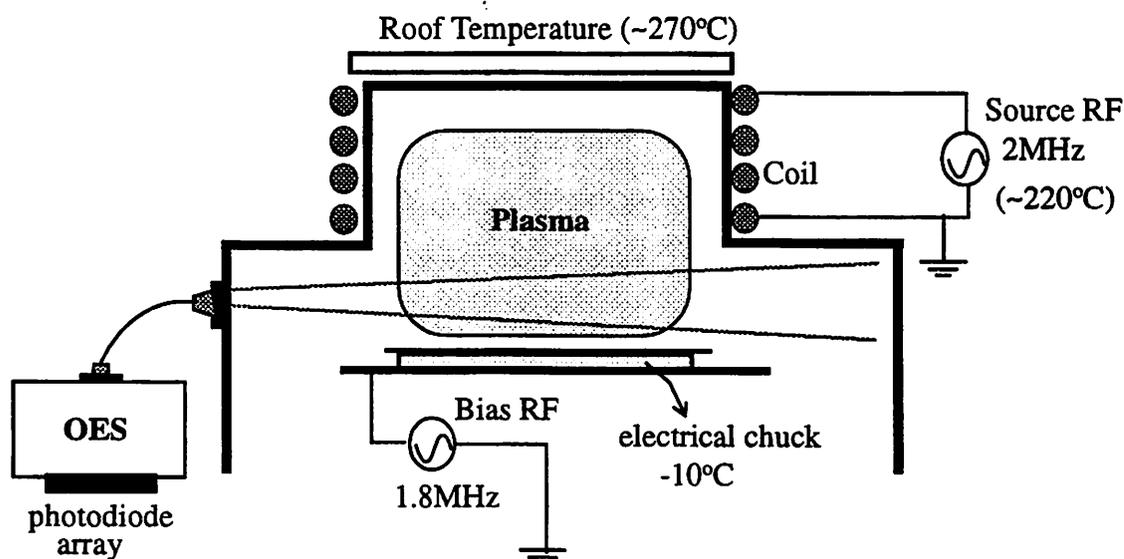


Figure 3.2 Schematic illustration of the experimental set-up on AMAT 5300 oxide etcher

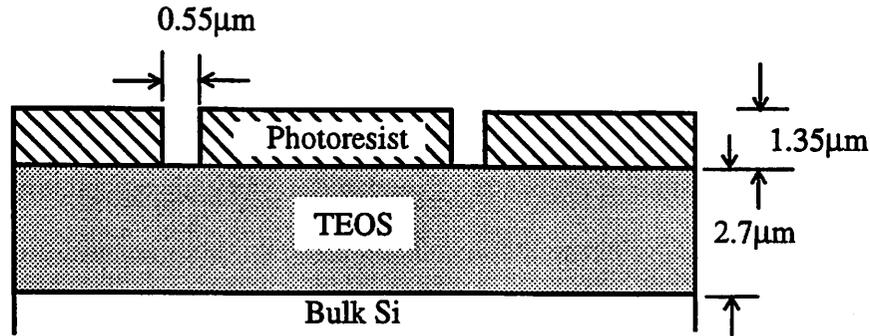


Figure 3.3 Cross section of the test structure for a monitor wafer.

3.2.2. Experimental Design

A designed experiment was conducted in order to develop the wafer state models. The experiment was a five-factor, 26-run central composite design, in addition to four center points [3.4]; this is a resolution V design without blocking. The input variables in this experimental design were source power (Watts), bias power (Watts), C_2F_6 flow rate (sccm), roof temperature ($^{\circ}C$), and throttle opening ratio (%). The deviations from the center values of the input settings used in this experiment were about $\pm 15\%$ (see Table 3.1). The nominal values of input settings are the typical recipe for interconnect dielectrical etching. This range of input settings provides sufficient range of real-time data and output wafer responses. Thirty 8-inch monitor wafers used in this experiment were covered with $2.7\mu m$ TEOS and $1.35\mu m$ photoresist (see Figure 3.3) patterned and developed for narrow line openings. In-line sensor data were captured during the processing of each wafer and the responses were measured for each run. The output wafer measurements (*i.e.*, the responses of this experiment) were oxide etch rate, oxide etching uniformity¹, aspect-ratio dependent etching² (ARDE) near the center of the wafer, and ARDE near the edge of the

1. Within-wafer uniformity is defined as (Standard deviation of etch rate over 17 selected sites on wafer) / (Average of etch rate over the 17 sites on wafer) x 100.

2. Aspect-ratio dependent etching (ARDE) is determined from SEM picture readings and is defined as $100 \times (\text{Depth of contact hole} - \text{Depth of open area}) / (\text{Depth of contact hole})$

wafer. Film thickness was measured both pre- and post-etch on the Prometrix UV 1050 at 17 points on each wafer; etch rate and uniformity were calculated from these thickness measurements. ARDE values were determined from SEM photographs. Due to some sensor failures, only 29 out of 30 wafers were included in the analysis that follows.

Table 3.1
Change in Percent from Nominal

| input variables | Source (Watts) | Bias (Watts) | C ₂ F ₆ flow rate (sccm) | Roof Temp (°C) | Throttle (Percent) |
|--------------------------------|----------------|--------------|--|----------------|--------------------|
| central point | 2600 | 1600 | 30 | 270 | 65 |
| change in percent from nominal | 11.5% | 12.5% | 33.3% | 3.7% | 53.8% |

3.3. Modeling Results

Since the typical OES data set contains 501 correlated variables in each spectrum, it is not practical to use all of them. One notable challenge is to decide which wavelengths should be chosen to represent the entire set of OES variables. Several data filters are introduced in this study. The easiest method is simply to filter the wavelengths based on spectral identification prior to the data analysis. Additionally, Principal Component Analysis (PCA) and Partial Least Squares (PLS) are two statistical multivariate data reduction techniques for compressing a large number of variables down to a small number (<10) of orthogonal variables, as described in Section 2.6. These reduced variables can be then used as the input matrix for a regression model. Ordinary Least Squares Regression [3.7][3.8] has been employed to relate the three reduced OES data sets to the etching performance. All three methods have been applied in this study, and they will be discussed in some detail next.

3.3.1. Regression models using the intensities of the pre-selected wavelengths

A typical optical emission spectrum acquired from the Applied Materials 5300 with C_2F_6 etchant is shown in Figure 3.4, along with the corresponding chemical species associated with chosen spectral lines. Only those spectral peaks associated with the important plasma species are chosen as the input variables for regression modeling. An F-distribution test is also employed in order to confirm whether these intensities vary significantly during the experiment. This test compares the variance of the real-time signals collected from the factorial experiment and those collected from the baseline runs (*i.e.*, center points of the DOE). Those variables that have a substantial variation relative to the baseline data are considered to be sensitive to the equipment settings. More specifically, the F-statistic is calculated by

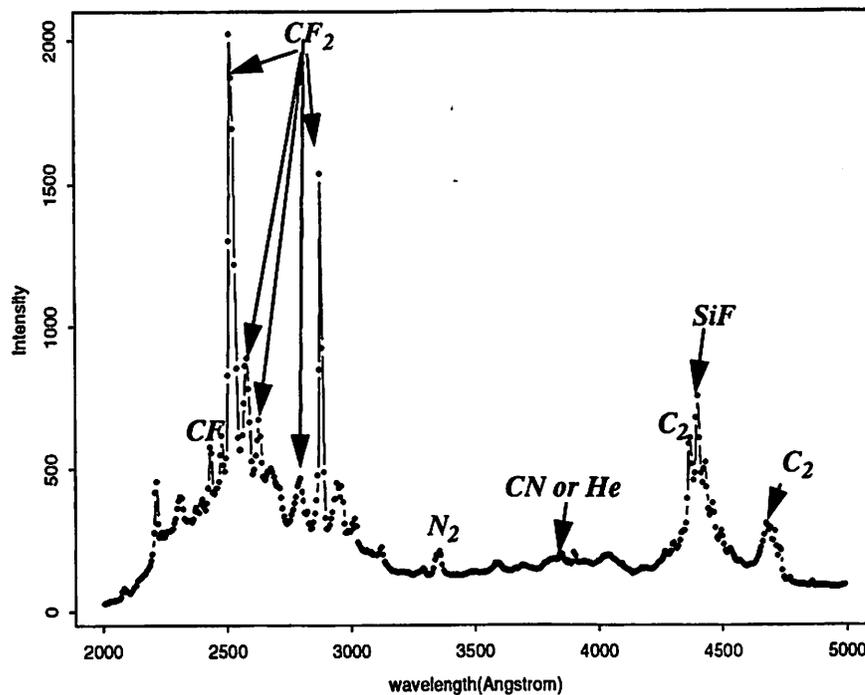


Figure 3.4 A typical spectrum collected from the Applied Materials 5300 during the main etch (etchant is C_2F_6) with spectral lines labeled with their corresponding species.

$$\frac{s_{all}^2/v_{all}}{s_{ctr}^2/v_{ctr}} \sim F_{\alpha, v_{all}, v_{ctr}} \quad (3.1)$$

where s_{all}^2 is the estimated variance of the signals collected during the factorial experiment, s_{ctr}^2 is the estimated variance of the signals collected during the centerpoint runs, α is the probability of type I error (*i.e.*, the probability that a process producing acceptable values of a particular quality characteristics will be rejected as performing unsatisfactorily), v_{all} signifies the degrees of freedom in the factorial experiment, and v_{ctr} signifies the degree of freedom in the centerpoint runs. In our case, v_{all} is 29 and v_{ctr} is 3. Those wavelengths having high F-test values are determined to have statistically significant variation. Using this approach, we selected eight particular wavelengths representing the dominant chemical species, listed in Table 3.2, as the input variables for building regression models.

Table summarizes the modeling results for different wafer responses. The results reveal that OES signals have a strong correlation with oxide etch rate and uniformity because of good R^2 values³ (greater than 0.8), while the OES signals can only explain 50% of the variation in ARDE. Nonetheless, despite the high R^2 value, the models resulting from this method are not suitable for the purpose of prediction. This is because the high degree of correlation among these input OES variables induced a multicollinearity problem, and as a result, the prediction capability of the model could be very poor [3.5]. For example, Figure 10 displays typical wafer-to-wafer time series of intensities for three different wavelengths (*i.e.* 251, 288 and 440nm). It apparently shows a strong correlation among these three variables. As a result, in the model of oxide uniformity, the correlation between coefficients of first and second variables is 0.8892, so they are almost correlated. Moreover, we only select eight wavelengths and ignore the rest of the OES variables in

3. The value of the R^2 statistic is a common measure for regression model goodness of fit. It is a measure of the proportion of total variation explained by the regression model. A perfect model fit is indicated by an R^2 statistic of 1.

Chapter 3

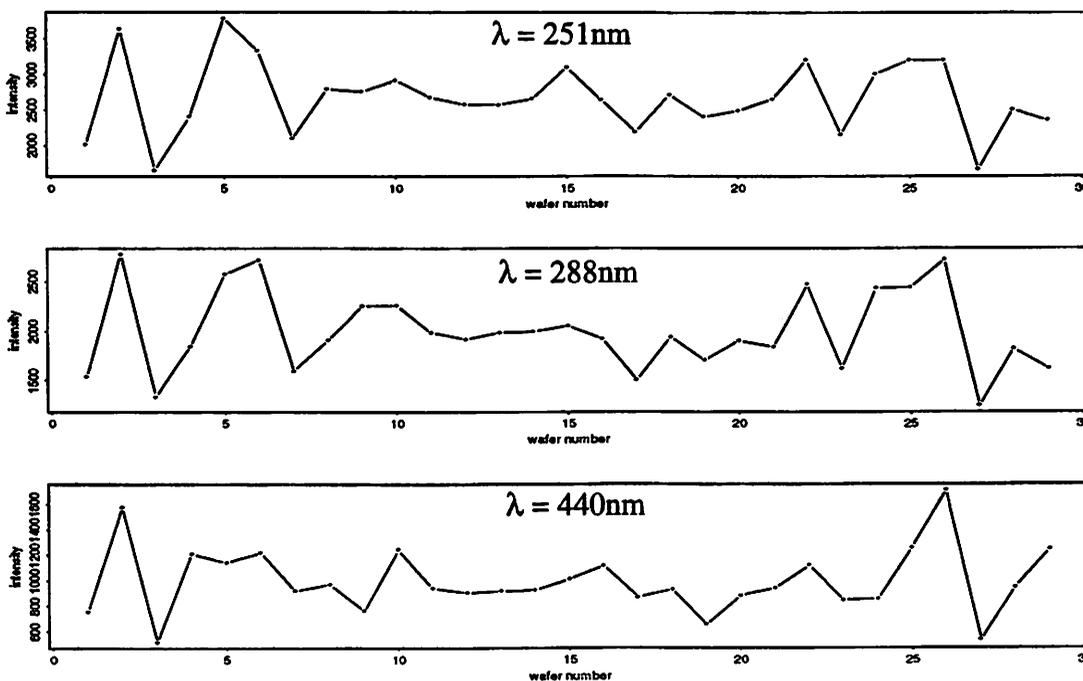


Figure 3.5 Wafer-to-wafer plots of intensities for three different wavelengths (upper: 251nm, middle: 288nm and bottom:440nm)

this OLSR modeling. This assumption might not be true and may result in incomplete models. In order to overcome these problems, other statistical methods including PCR and PLSR are used as alternatives.

Table 3.2

The distinct wavelengths selected for the ordinary least squares regression model.

| wavelength(nm) | possible species |
|----------------|------------------|
| 248 | CF |
| 251.6 | CF ₂ |
| 258 | CF ₂ |
| 288 | CF ₂ |
| 385 | CN or He |
| 437 | C ₂ |
| 440 | SiF |
| 467 | C ₂ |

Table 3.3 ANOVA Tables of OLSR models for all responses

| Response: Oxide Etch Rate | | | | |
|---------------------------------------|-----|----------|---------|----------|
| Source | DOF | SS | MS | F |
| Regression | 8 | 38363120 | 4795390 | 18.11299 |
| Residual | 20 | 5294974 | 264749 | |
| $R^2 = 0.88$; Adjusted $R^2 = 0.834$ | | | | |

| Response: Oxide Etch Uniformity | | | | |
|---------------------------------------|-----|----------|----------|----------|
| Source | DOF | SS | MS | F |
| Regression | 8 | 94.88443 | 11.86055 | 40.52149 |
| Residual | 20 | 5.85396 | 0.29270 | |
| $R^2 = 0.942$; Adjusted $R^2 = 0.92$ | | | | |

| Response: ARDE_Center | | | | |
|---------------------------------------|-----|----------|----------|----------|
| Source | DOF | SS | MS | F |
| Regression | 8 | 6247.014 | 780.8767 | 3.234707 |
| Residual | 20 | 4828.115 | 241.4057 | |
| $R^2 = 0.56$; Adjusted $R^2 = 0.392$ | | | | |

| Response: ARDE_Edge | | | | |
|--|-----|----------|----------|----------|
| Source | DOF | SS | MS | F |
| Regression | 8 | 2868.347 | 358.5434 | 4.498748 |
| Residual | 20 | 1593.97 | 79.6985 | |
| $R^2 = 0.643$; Adjusted $R^2 = 0.507^a$ | | | | |

a. This is a modified version of the R^2 statistic which considers the number of parameters of the model and is given by $AdjR^2 = 1 - [1 - R^2] \frac{n}{n-p}$, where n is the number of observations and p is the number of parameters.

3.3.2. Regression models using PCA reduced variables

As detailed previously in Section 2.6, PCA reduces the dimensionality of the data by projecting them onto a low-dimensional space and converting them into a uncorrelated data set. One important task is to determine how many principal components (PCs) should be retained in the model. An empirical method is to make a screeplot which indicates the

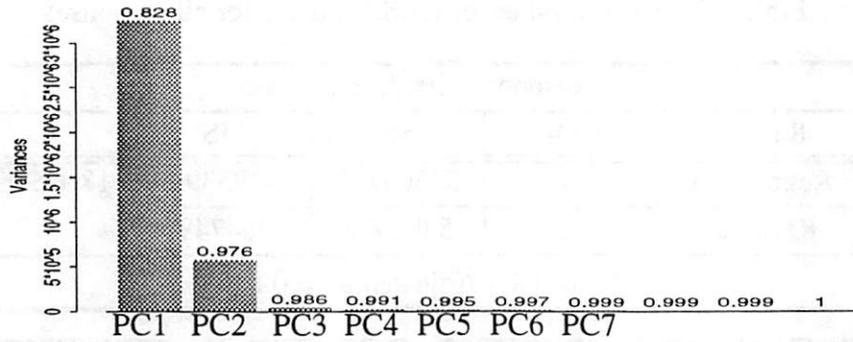


Figure 3.6 Scree plot of the principal component analysis of OES data (the number shown on the top of each bar indicates the cumulative summation of their explained percentage of variations of total OES data)

percentage of variation explained by each principal component in the study [3.6]. We found that seven PCs are adequate to explain 99.9% of process variation (see Figure 3.6.).

The loading plots of the first seven PCs (Figure 3.7) reveal that these PCs consist mostly of contributions from only 10~20 spectral peaks. The rest of the wavelengths have only negligible weights in the PCA modeling, which agrees with our previous observation that only 10~20 spectral lines among the 501 monitored wavelengths can be related to the variation of the process outputs.

Using the seven PCs with the largest eigenvalues, the PCA models can explain 85% and 95% of the variation in etch rate and uniformity, respectively. Also, an R^2 of 0.65 is achieved for the ARDE models. Table 3.4 summaries the variability of the predictors (i.e., OES signals) and responses (i.e., wafer-state measurements) explained by the principal components. The results reveal that the 2nd, 4th, and 7th PCs explain the most of variabilities of oxide etch rate and uniformity. By contrast, the 2nd, 5th, and 6th PCs are more related to ARDE. The loading plots (Figure 3.7) indicate that CF_2 lines, especially for 258nm and 288nm spectral lines, are more responsible features than SiF lines.

Since only a few PCs show a strong correlation with wafer states, one variable selection technique based on the student-t test at the 0.05 significance level is conducted to reduce the number of PCs used for principal component regression modeling. The results

show that only 4, 5, 3, 5 PCs are sufficient to explain oxide etch rate, uniformity, ARDE_center, and ARDE_edge, respectively. Table 3.5 summaries the ANOVA tables for the models based on these selected PCs.

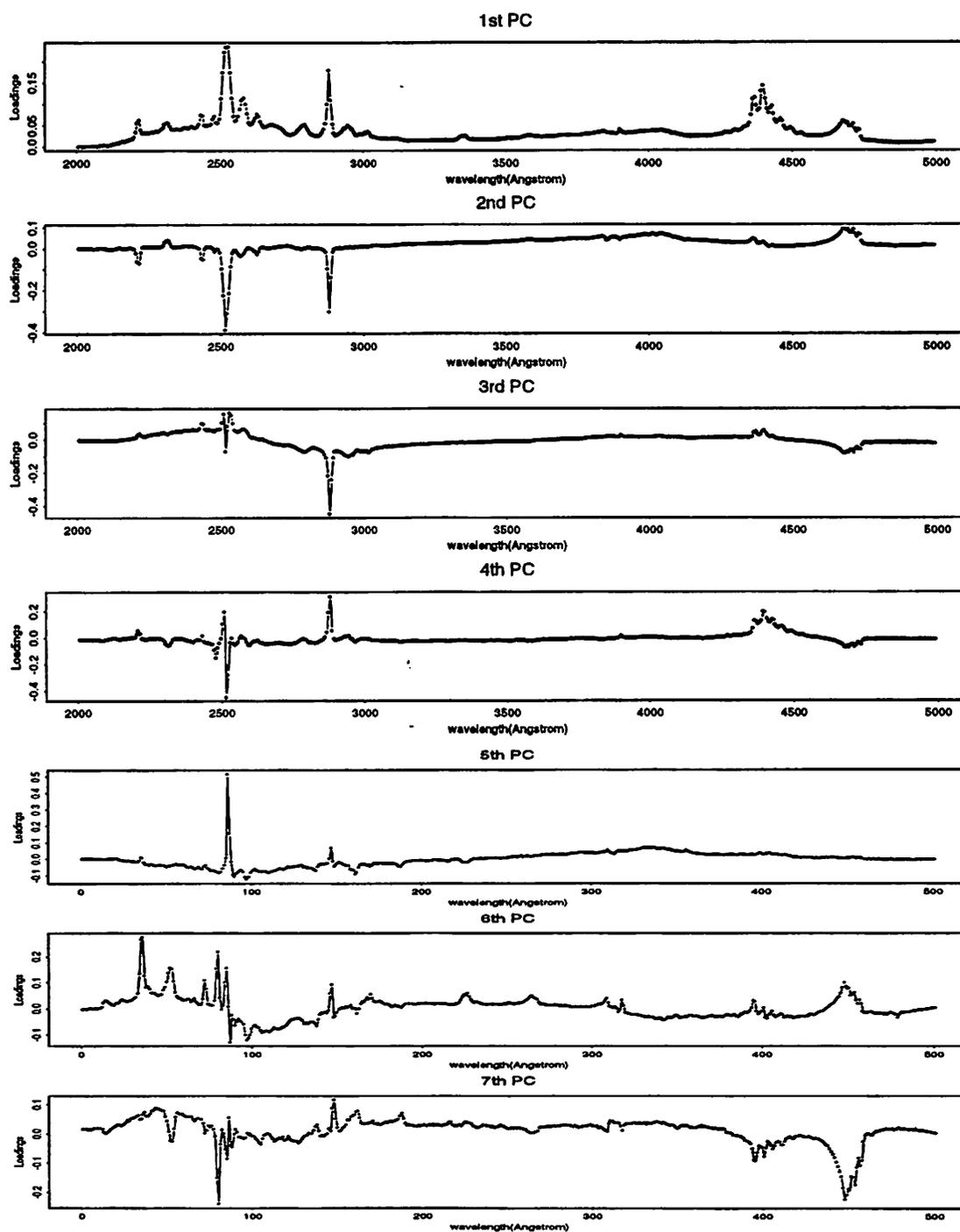


Figure 3.7 The loading vectors of the first seven principal components

Table 3.4
Proportion of predictor and response variability explained by principal components (Bold: the PCs selected for PCR models)

| | X | Y(etch rate) | Y(uniformity) | Y(ARDE_cen) | Y(ARDE_edg) |
|--------------|-------|--------------|---------------|-------------|--------------|
| Explained by | | | | | |
| z_1 | 82.8% | 12.5% | 15.6% | 1.22% | 3.3% |
| z_2 | 14.8% | 55.1% | 61.8% | 40% | 36% |
| z_3 | 1% | 0.5% | 0.65% | 2% | 5.8% |
| z_4 | 0.5% | 9.4% | 5.7% | 0.5% | 0.2% |
| z_5 | 0.4% | 0.04% | 5.5% | 8.5% | 8.5% |
| z_6 | 0.2% | 1.3% | 0.45% | 9.1% | 12.2% |
| z_7 | 0.2% | 9.8% | 6.45% | 0.68% | 11.2% |
| Unexplained | 0.1% | 11.3% | 3.7% | 38% | 25.8% |

Table 3.5
The ANOVA Tables for PCR models of all responses.

| Response: Oxide Etch Rate | | | | |
|--------------------------------------|-----|----------|---------|----------|
| Source | DOF | SS | MS | F |
| Regression | 4 | 37918939 | 9479735 | 39.64236 |
| Residual | 25 | 5739155 | 239131 | |
| $R^2 = 0.87$; Adjusted $R^2 = 0.85$ | | | | |

| Response: Oxide Etch Uniformity | | | | |
|---------------------------------------|-----|----------|----------|----------|
| Source | DOF | SS | MS | F |
| Regression | 5 | 95.86826 | 19.17365 | 90.55085 |
| Residual | 24 | 4.87013 | 0.21174 | |
| $R^2 = 0.951$; Adjusted $R^2 = 0.94$ | | | | |

| Response: ARDE_Center | | | | |
|--------------------------------------|-----|----------|----------|----------|
| Source | DOF | SS | MS | F |
| Regression | 3 | 5421.432 | 1807.144 | 7.990983 |
| Residual | 26 | 5653.697 | 226.148 | |
| $R^2 = 0.49$; Adjusted $R^2 = 0.43$ | | | | |

| Response: ARDE_Edge | | | | |
|------------------------------------|-----|----------|----------|----------|
| Source | DOF | SS | MS | F |
| Regression | 5 | 3154.624 | 630.9248 | 11.09685 |
| Residual | 24 | 1307.692 | 56.8562 | |
| $R^2 = 0.71$ Adjusted $R^2 = 0.65$ | | | | |

3.3.3. Regression models using PLS reduced variables

PLS can also reduce the number of terms in the final model. The main difference is that while in PCA the transformation is only dependent on the variability in the input matrix X (i.e., OES data), in PLS the transformation depends on both the input X and the response Y (i.e., wafer state measurements). To determine the appropriate number of variables to retain in the modeling, we minimize a criterion called prediction error sum of square (PRESS) (see Eq.(2.12)). These minimized PRESS statistics indicate that 2-6 variables (dependent on the different wafer responses) are sufficient to describe the input data and then used as the regressor for the wafer responses. Figure 3.8 shows the plots of PRESS versus the number of variables used in PLSR models. Like PCA, the results of PLS models exhibit good R^2 values for oxide etch rate and uniformity, and only moderate R^2 values for ARDE. Table 3.6 shows the ANOVA summaries of these PLSR models.

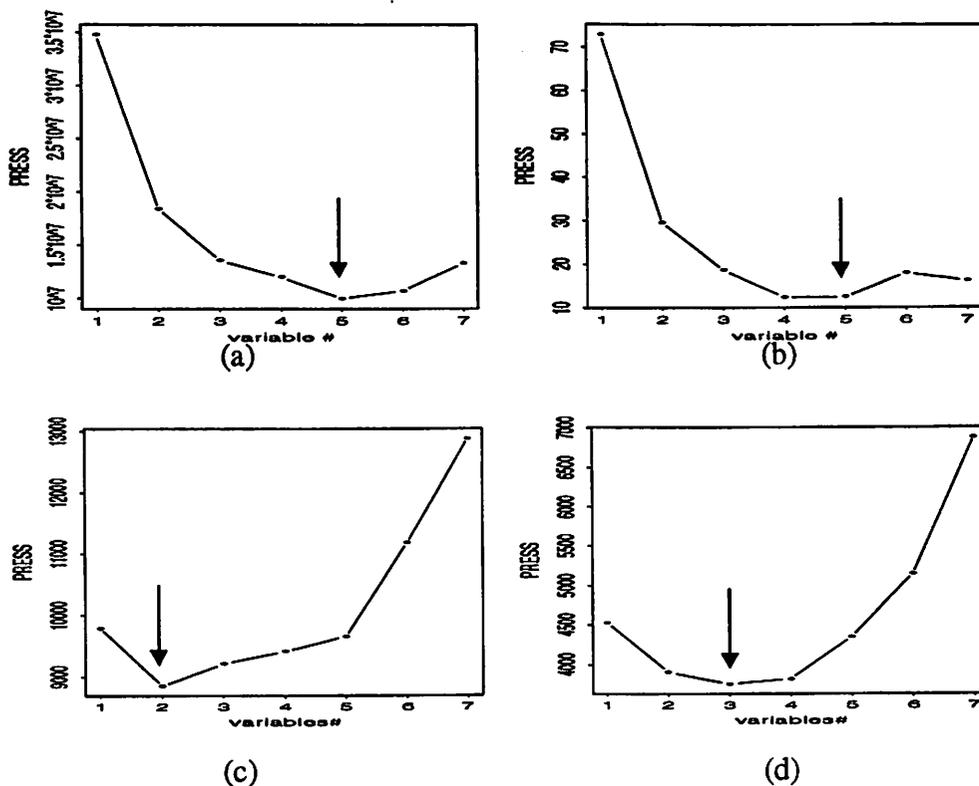


Figure 3.8 The plots of PRESS for all wafer responses (a) oxide etch rate; (b) uniformity; (c) ARDE at center; (d) ARDE at edge

Table 3.6
The ANOVA tables of PLSR models for all responses

| Response: Oxide Etch Rate | | | | |
|--|-----|----------|---------|----------|
| Source | DOF | SS | MS | F |
| Regression | 5 | 39026329 | 7805266 | 38.75869 |
| Residual | 23 | 4631765 | 201381 | |
| $R^2 = 0.894$; Adjusted $R^2 = 0.872$ | | | | |

| Response: Oxide Etch Uniformity | | | | |
|---|-----|----------|----------|----------|
| Source | DOF | SS | MS | F |
| Regression | 5 | 96.99312 | 19.39862 | 119.1285 |
| Residual | 23 | 3.74527 | 0.16284 | |
| $R^2 = 0.9628$; Adjusted $R^2 = 0.955$ | | | | |

| Response: ARDE_Center | | | | |
|--------------------------------------|-----|----------|----------|-------|
| Source | DOF | SS | MS | F |
| Regression | 2 | 4697.992 | 2348.996 | 9.577 |
| Residual | 27 | 6377.137 | 245.275 | |
| $R^2 = 0.42$; Adjusted $R^2 = 0.38$ | | | | |

| Response: ARDE_Edge | | | | |
|-------------------------------------|-----|---------|----------|--------|
| Source | DOF | SS | MS | F |
| Regression | 3 | 2849.2 | 949.7351 | 14.719 |
| Residual | 26 | 1613.11 | 64.5244 | |
| $R^2 = 0.64$; Adjusted $R^2 = 0.6$ | | | | |

A flow chart of the modeling scheme using the OES signals is illustrated in Figure 3.9.

3.4. Discussion

As indicated in the previous sections, an OES sensor can capture the wafer-to-wafer variations of the oxide etch rate and uniformity. This can be explained as follows: oxide etch is essentially an ion-bombardment sputtering process, as mentioned in Chapter

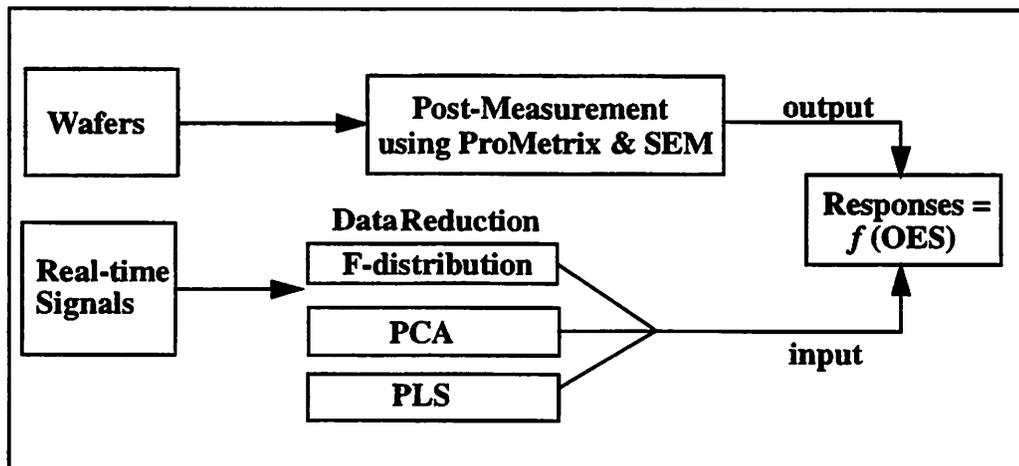


Figure 3.9 The flow chart of wafer state modeling

2. The oxide etch rate is dependent on the ion bombarding activity, and thereby relates to the ion energy (controlled by bias power) and ion flux (controlled by source power and gas flow). Because optical emission intensity of discharge is strongly related to source and bias power levels as indicated in Section 2.5, OES sensors are capable to capture the changes of ion-bombardment behaviors and thus explain oxide etch rate variations.

The overall limited success in predicting ARDE from OES signals might be due to several reasons. First, ARDE is only measured at two separate sites on each wafer (one is in the center and another is in the edge). It may be insufficient to represent the entire wafer. Secondly, the ARDE measurements are conducted by an off-line cross-section SEM micrographic readings, and thus may not be accurate. Thirdly, ARDE is not only related to the ion energy/flux, but rather related to the polymer forming, and chamber wall and roof temperatures (NOTE: polymer deposition is extremely sensitive to the wall/roof temperatures [3.9]), which may not be revealed by the OES sensors.

Table 3.7 gives a comparison of all three data reduction and modeling methods. It shows the number of variables used in each model and the performance of each model in terms of adjusted R^2 values. This comparison reveals that performance of PLSR is similar to that of PCR.

It is also important to examine the role of scaling in the OES data sets. Scaling usually plays a vital role for the data files with different scale or unit. However, in the OES data file, most of the data (say, 60-70%) appear to be noise. Scaling, as a result, could give equal weights to the noise features as that for meaningful spectral features, and thereby degrade the R-square value of the model. For the good models with a high correlation between response and OES data such as etch rate, the use of the scaled X matrix could result in a worse R-square than unscaled X matrix does. For the models with low correlation ($R^2 < 0.6$), scaling does not show any significant effect on R-square.

Table 3.7
Summary of the results of chamber state models.

| Response: Oxide_ER | | | | Response: Oxide_Uniformity | | | |
|------------------------|----------------------|----------------|---------------------|----------------------------|----------------------|----------------|---------------------|
| Data reduction method | # of input variables | R ² | Adj. R ² | Data reduction method | # of input variables | R ² | Adj. R ² |
| Species Identification | 8 | 0.88 | 0.834 | Species Identification | 8 | 0.94 | 0.92 |
| PCA | 4 | 0.87 | 0.85 | PCA | 5 | 0.95 | 0.94 |
| PLS | 5 | 0.9 | 0.872 | PLS | 5 | 0.96 | 0.955 |

| Response: ARDE at center | | | | Response: ARDE at edge | | | |
|--------------------------|----------------------|----------------|---------------------|------------------------|----------------------|----------------|---------------------|
| Data reduction method | # of input variables | R ² | Adj. R ² | Data reduction method | # of input variables | R ² | Adj. R ² |
| Species Identification | 8 | 0.56 | 0.392 | Species Identification | 8 | 0.64 | 0.507 |
| PCA | 3 | 0.49 | 0.43 | PCA | 5 | 0.74 | 0.66 |
| PLS | 2 | 0.42 | 0.38 | PLS | 3 | 0.64 | 0.6 |

3.5. Summary

Run-to-run chamber state modeling using real-time OES signals is effective in capturing the process variation and explaining the final wafer characteristics, especially for oxide etch rates and their within-wafer uniformity. In this chapter, three data reduction techniques are compared. First, Ordinary Least Squares Regression is performed on wavelengths selected based on species identification. Two other modeling techniques, Principal Component Analysis and Partial Least Squares were also introduced to eliminate the correlation among input variables and reduce the input matrix size. The resulting models of oxide etch rate and within-wafer uniformity were very good in the sense that 85% of the

Chapter 3

etch rate variation and more than 95% of uniformity variation were explained by these models. However, only less than 60% of ARDE variation can be captured by OES signals, which might be due to the inaccuracy of ARDE measurements or/and the insufficient information collected by OES. No modeling technique is overwhelmingly better than the others, in terms of their R^2 values. Nevertheless, PLSR generally involves a more complex computation, and therefore is computationally expensive.

References for Chapter 3

- [3.1] J. A. Chan, "Response Surface Methodology Improves Batch RIE Processing", *Semiconductor International*, Vol. 18, N. 3, March 1995, pp. 85-90.
- [3.2] C. P. Ho, J. D. Plummer, S. E. Hansen, R. W. Dutton, "VLSI Process Modeling-SUPREM III", *IEEE Transaction on Electron Devices*, Vol. ED-30, No. 11, November 1983, pp.1438-1453
- [3.3] S. Wolf and R. Tauber, *Silicon Processing for the VLSI Era*, vol. 1, Lattice Press, 1986, pp.550-551.
- [3.4] G. E. P. Box, W. G. Hunter, and J. S. Hunter, *Statistics for Experimenters*, John Wiley and Sons, 1978.
- [3.5] Sherry Lee, *Semiconductor Equipment Analysis and Wafer State Prediction System Using Real-Time Data*, Ph.D. Dissertation, Dec. 1994.
- [3.6] J. E. Jackson, *A User's Guide to Principal Components*, John Wiley and Sons, 1991.
- [3.7] I. Frank and J. Friedman, "A Statistical View of Some Chemometrics Regression Tools", *Technometrics*, Vol. 35, N.2, May 1993, pp. 109-148.
- [3.8] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*, Academic Press, 1979.
- [3.9] P. Singer, "The Many Challenges of Oxide Etching", *Semiconductor International*, June, 1997, pp. 109-114.

CHAPTER 4

Wafer-State Modeling Using OES for Metal Etch

4.1. Introduction

In this chapter we examine the optical emission spectra collected from a commercial transformer coupled plasma (TCP) high-density plasma (HDP) metal etcher at Texas Instruments (TI) under various machine input settings. Our goal is to describe the relation between emission intensity variations and machine parameter variations and, most importantly, the relationship between emission variation and wafer state variation, as illustrated in Figure 4.1. These optical emission spectra collected during aluminum etch and over etch of the underlying oxide can be employed in combination with several statistical and physical techniques to model the wafer states such as etch rate and critical dimension (CD), and machine states such as coil power and chamber pressure.

4.2. Experiment to Link OES Signals to Metal Etch Variability

This work was conducted on a commercial Lam 9600 TCP metal plasma etch tool with BCl_3/Cl_2 etchant. The experimental wafers have a 200\AA TiN/ 6000\AA Al-0.5%Cu/

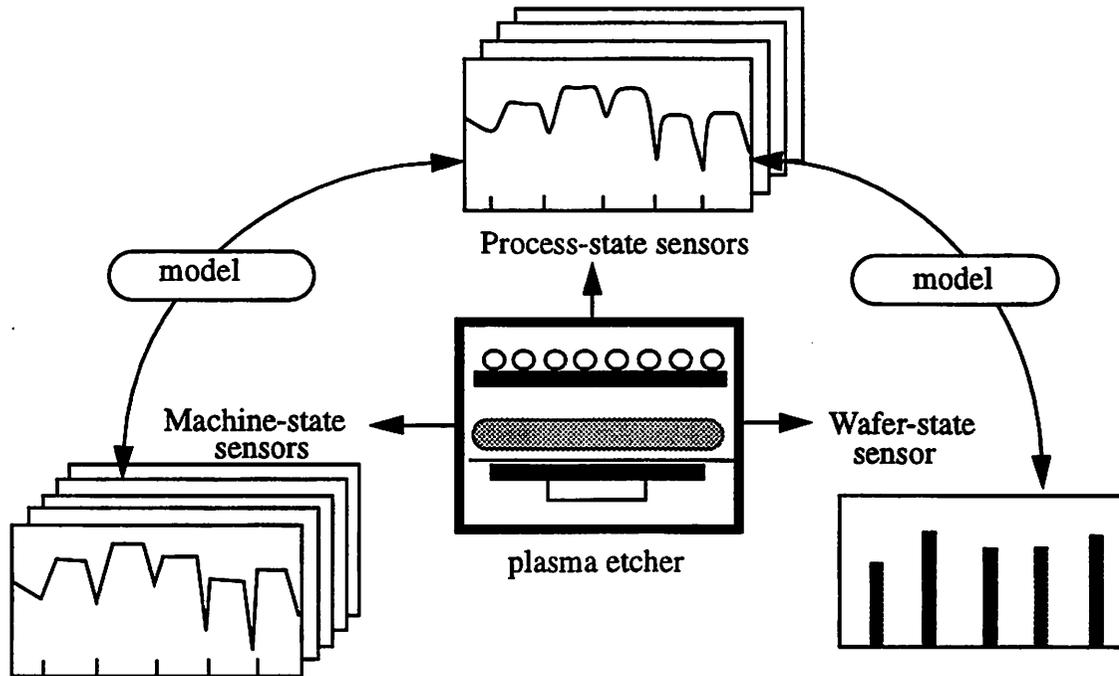


Figure 4.1 An overview of this study.

200Å TiN/SiO₂ stack covered by patterned photoresist with 0.5μm linewidth (see Figure 4.2). The main etch process was terminated by the intensity drop of the AlCl spectral line ($\lambda=262\text{nm}$) detected by an interference filter endpoint system, followed by a timed over-etch process for the underlying TiN and SiO₂ layers. Several sensors were monitored in this project, including machine-state sensors such as pressure, gas flow and power sensors, and process-state sensors such as OES. The machine-state sensors were built into the etcher tool to collect the available machine data at an acquisition rate of 1Hz during the processing of each wafer. A Chromex spectrograph with a Princeton Instrument 1024 × 256 CCD camera was used in this study to obtain the temporally and spatially resolved process-state information. Three optical fibers are connected between the reactor viewport and the spectrograph. The spectrograph is set up to view the plasma across three distinct regions arranged laterally above the wafer being etched. The spectrograph was configured

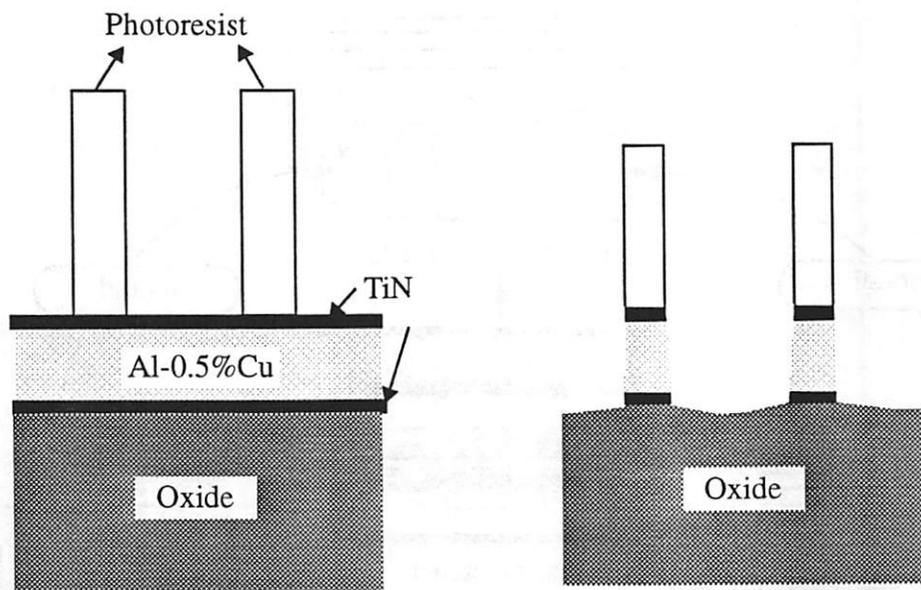


Figure 4.2 The TiN/Al-Cu/TiN/Oxide stack (a) before etch, and (b) after etch.

with a split grating, which enabled the entire spectrum to be imaged onto the CCD in two separate stripes per beam. This split grating provides the desired bandwidth/resolution combination without having to mechanically scan the grating. In this experiment, the spectral data collection was performed at an acquisition rate of 0.33Hz (i.e., 3 second cycle time). Each spectra has 2042 intensity values in the 245 to 520nm and 530nm to 800nm range, corresponding to a spectral resolution of less than 1nm. A typical optical emission spectrum acquired from the etcher is shown in Figure 4.3, along with the corresponding chemical species associated with chosen spectral lines.

Since one goal of this effort was to explain the plasma etch process for a wide variety of process conditions and across a wide range of setpoints, an experimental design was created to span the range of setpoints of interest. A blocked, five-level central composite experiment was conducted on 70 wafers for the five variables: TCP coil power (Watts), RF bias power (Watts), pressure (mTorr), total gas flow (sccm), and C_2/BCl_3 ratio. The deviation from the nominal values of the input settings used in this experiment was +/-

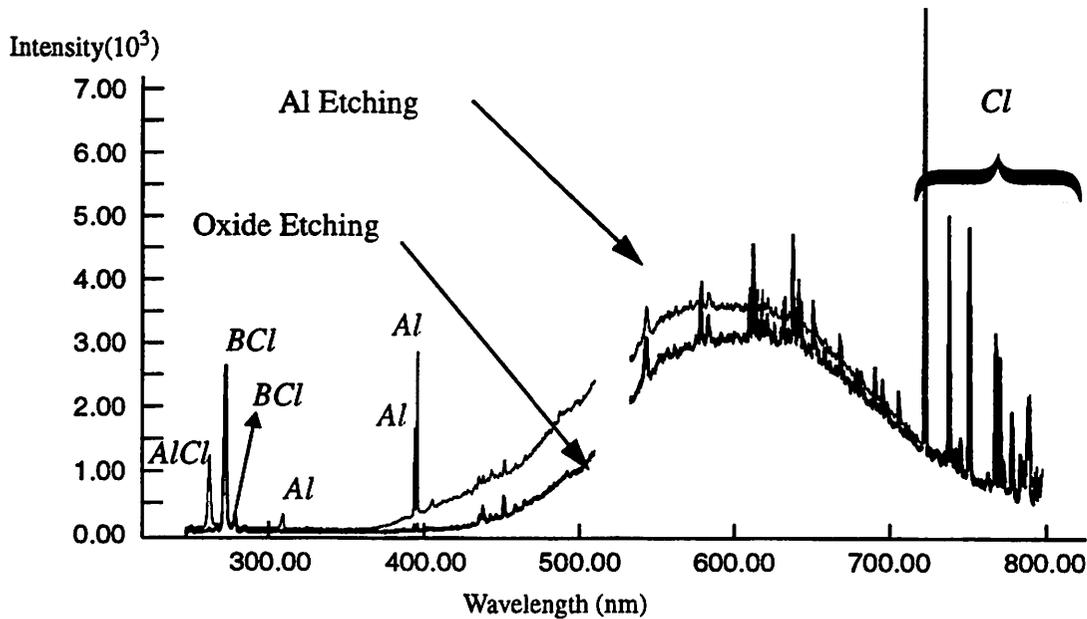


Figure 4.3 The identification of the optical emission spectra collected from a Lam 9600 TCP metal etcher (etchant: Cl_2/BCl_3)

10~20%. Several wafer-state parameters were determined after the etch process: Electrical critical dimension (CD) measurements were taken on a Tokyo Seimitsu tool for the aluminum linewidth after etch process (see Figure 4.4(a)). The oxide lost during the over-etch process was determined by the difference in oxide thickness measured after oxide deposition and after etch (see Figure 4.4(b)). As a result, oxide etch rate was calculated as the total oxide loss divided by the oxide etch time (determined by the endpoint trajectory), from which we determined the oxide etch rate. Both oxide loss and CD were measured for each of the 32 dies on the wafer. Due to several sensor and measurement failures, only 51 wafers were used in this analysis. In the following sections, “CD” is used to represent the post-etch aluminum linewidth, and “oxide etch rate” is used to represent the over-etch rate for the underneath oxide layer.

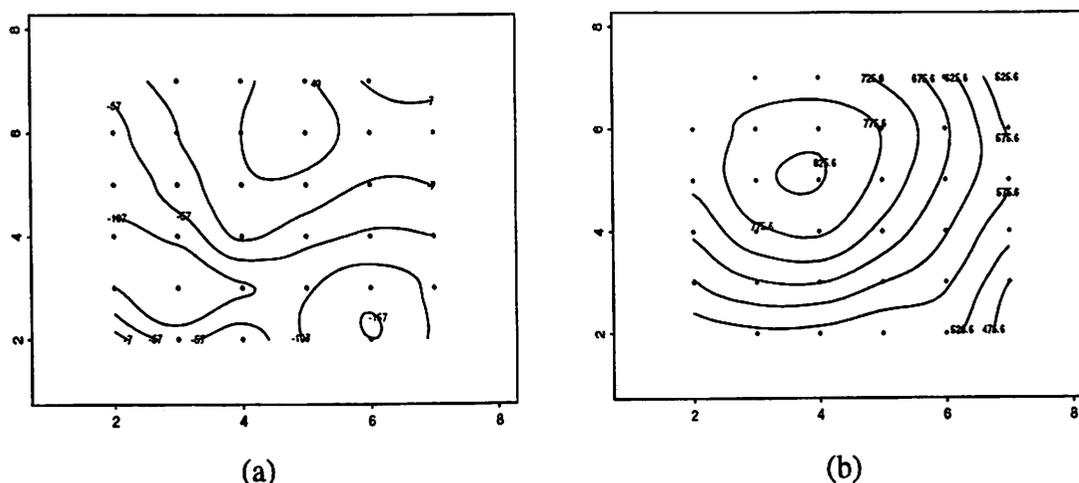


Figure 4.4 Typical contour plots of (a) aluminum CD measurements; (b) oxide loss measurements. (the unit of the contour plots is angstrom)

4.3. Data Pre-Processing

The variation of OES data at one wavelength has two components: temporal variation (due to the optical emission intensity changes over time), and spatial variation (due to the intensity difference of three spatially-resolved spectra at the same spectral line). In this chapter, we only consider the temporal variation without taking into account the spatial variation among the three beams of OES signals in order to simplify the analysis. The details of using the information of the spatial variation and temporal-spatial interaction variation will be addressed in the later chapters.

4.3.1. OES Temporal Variation

A typical temporal trajectory of optical emission at the Cl line (725nm) from one wafer is shown in Figure 4.5. Note that it is a three-stage process. Stage 1 is the main etch of the Al layer terminating by the machine endpoint detector (operating at 261nm, *i.e.*, AlCl line), stage 2 is a brief over-etch process for the underlying thin TiN layer, and stage 3 is an over-etch process for the underlying oxide layer. While described in different

stages, it is important to note that this is a single recipe etch process, that is, the process recipe is identical during stage 1 through 3. This within-wafer temporal profile is due to the fact that different amounts of chemical species are generated or consumed throughout the etch process. Figure 4.5 also clearly shows that the within-stage variation is much smaller than the stage-to-stage variation. Therefore, the optical emission intensity during the aluminum etch (stage 1) and the oxide etch (stage 3) for each wafer is approximated by the average of emission intensity over stage 1 and stage 3, respectively. (The TiN etch stage is ignored in this study)

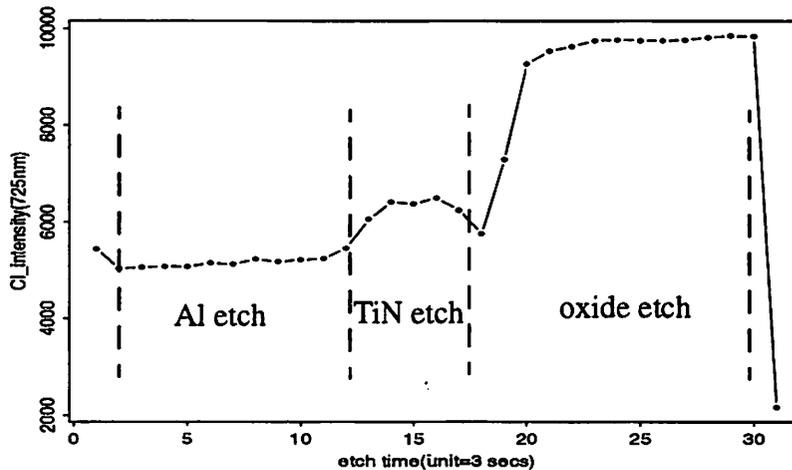


Figure 4.5 The temporal trajectory of optical emission intensity for one wafer at the Cl line

In addition, because each wafer was intentionally processed under different machine settings, wafer-to-wafer variation of sensor signals is much larger than within-wafer variation for each etch stage. Therefore, in the following sections, we model the wafer-to-wafer variation at selected wavelengths for the selected stage of the etch process. For instance, the wafer-to-wafer variation for the emission intensity at selected wavelengths for the oxide etch (i.e. over etch) stage can be examined to see whether it is affected

by the wafer-to-wafer chamber pressure, TCP coil power variation, or other process variables.

4.3.2. OES Variable Selection

One challenge worth noting is how to decide which wavelengths should be chosen to represent the entire set of OES variables in order to determine the temporal variation. Different wavelengths (corresponding to different chemical species and different excited states) may reveal totally different temporal profiles. Therefore, much effort has been devoted towards choosing the appropriate spectral lines to represent the entire spectrum. We tried two separate approaches. First, only the wavelengths associated with identified chemical species (e.g., Cl, BCl, and Al) shown in Figure 4.3 are chosen as the input variables for the following regression model. Second, a statistical spectral compression technique, based on principal component analysis (PCA)[4.1], is employed to extract the variance of the original spectral data using only a set of principal components (PCs)[4.1]. PCA often reveals relationships that were not previously suspected and thereby allows novel interpretations, as demonstrated in Chapter 3. As mentioned in Section 2.6, PCA can find a new coordinate system obtained by rotating the original system, and the new axes represent the directions with maximum variability. In a sense, PCA captures the wavelengths which give the highest intensity variability and describes their linear combination as principal components [4.1].

These PCs contain most data variance and are used as the input matrix for modeling. It is important to note that although this PCA technique lacks the insights of spectral identification, it is sufficient to extract the spectral lines giving the most variance by taking the lines with the greatest amplitudes on the loading plots. The objective of this regression modeling is to explain the wafer-to-wafer variation (that is, the OE at one wavelength for one wafer compares with the OE at the same wavelength collected from other wafer). As long as the PCs capture the wafer-to-wafer variation of the plasma emission, this informa-

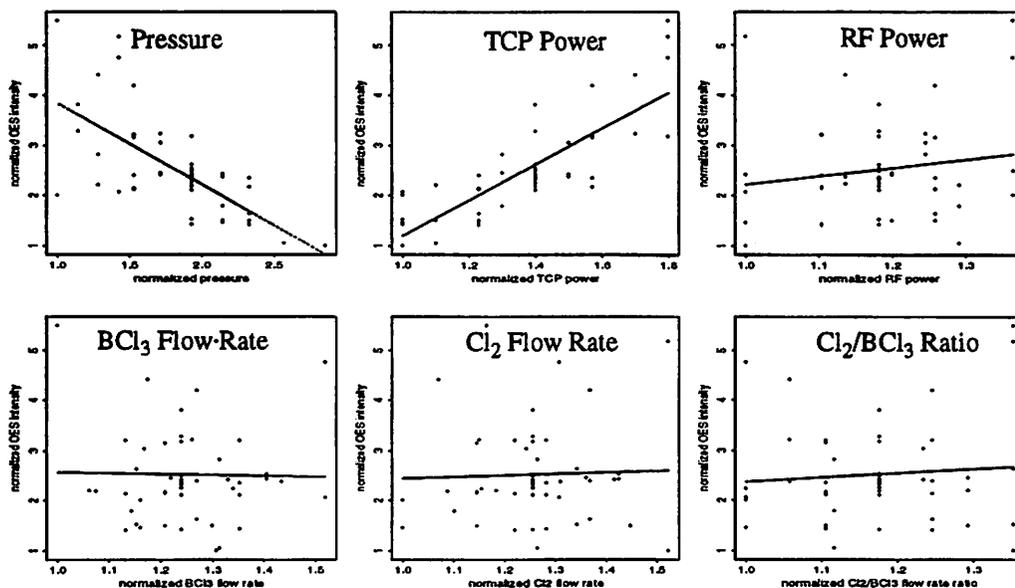


Figure 4.6 Cl line intensity versus machine settings (each line denotes the fitted linear regression model)

tion is sufficient to explain the corresponding wafer state variation. The details of the PCA algorithm are described in Chapter 2.

4.4. Results

4.4.1. Relating OES and Machine Variables

The first aspect of this work is to model the temporal mean of OES trajectories using the machine parameters collected by the machine-state sensors. We first select four spectral lines including 262nm(AlCl), 272nm(BCl), 396nm(Al), and 725nm(Cl) to represent the entire spectrum. A least-squares linear regression model is determined ($y = \text{OES intensity}$, $x = \text{machine settings}$) and a student-t test [4.2] is used to examine the significance of each machine variable. The results show that more than 96% of the emission intensity variation for either oxide etch or Al etch is simply explained by three machine parameters: 1/pressure, TCP coil power and bias power. In fact, 70~85% of emission intensity variation can be explained by the TCP coil power variation alone, whereas gas flow, and Cl_2/BCl_3

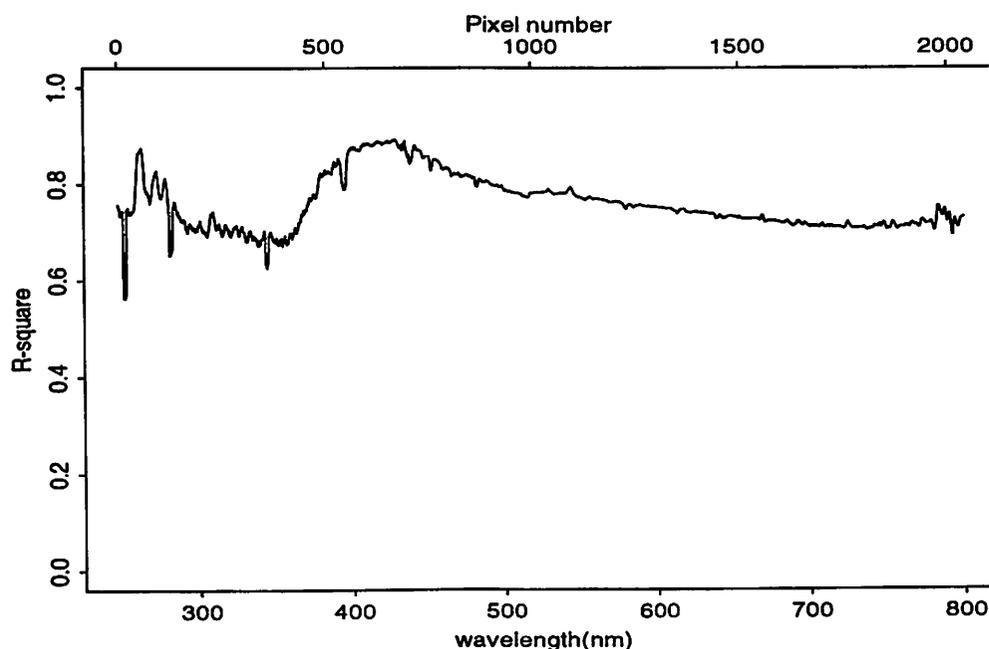


Figure 4.7 Univariate correlation between emission intensity and TCP power over all 2042 wavelengths in terms of their R-square value.

ratio have a negligible effect on the emission intensity. Figure 4.6 displays the effects of machine parameters on the optical emission intensities. Similar modeling results are shown in Section 4.4.2 for the OES data reduced by PCA.

It is worth noting that the correlation between emission intensities and TCP coil power is a weak function of wavelengths, as shown in Figure 4.7. In other words, the enhancement or reduction of emission intensities is strongly sensitive to the TCP coil power, rather than their corresponding chemical species. This will be addressed in the following sections.

4.4.2. Relating OES and Wafer-state Variables

The second aspect of this work is to model the spatial mean of the wafer-state parameters such as etch rates or CDs as a function of their corresponding OES signals acquired from one wafer without taking into account the spatial variation of wafer-state

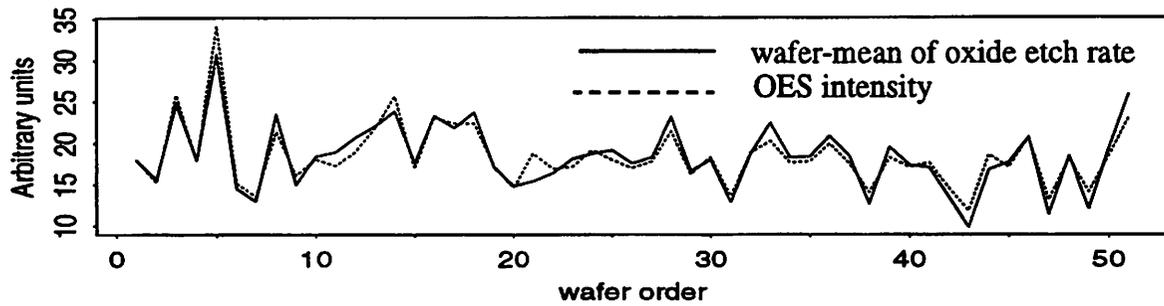


Figure 4.8 The wafer-to-wafer trajectory of oxide etch rate and emission intensity values from Cl line

measurements. The results show that the oxide etch rate can be easily modeled by four selected wavelengths (adjusted R-square ~ 0.9) with a residual standard error (RSE) of only $1.1\text{\AA}/\text{sec}$ (the average oxide etch rate is about $18.4\text{\AA}/\text{sec}$), whereas only an adjusted R-square of 0.036 is achieved for aluminum CD using the same wavelengths, corresponding to a RSE of $0.014\mu\text{m}$.

Figure 4.8 shows that the variations in emission intensities are strongly dependent on oxide etch rate even though only one spectral line, $\lambda=500\text{nm}$ (which is not a distinct feature as shown in Figure 4.3) is used to represent the entire spectrum. This suggests that the changes of oxide etch rate in this etching condition can be fully explained by emission brightness changes ($R^2 > 0.9$) regardless of the chosen wavelength's corresponding chemical species. In other words, the observed correlation between oxide etch rates and emission intensities is almost independent of wavelength.

The models are slightly improved if we use a PCA reduced OES data set as input variables. A student-t test at the 0.1 significance level is used to select the principal components which should be retained in the model to explain the variation of wafer-state measurements. This significance test is conducted for the modeling of the oxide etch rate. Four PCs are selected and the resulting oxide etch rate RSE is only $1\text{\AA}/\text{sec}$, corresponding to an adjusted R^2 of 0.94. However, the wafer-to-wafer variation of the spatial CD mean is still poorly explained by the PCA-based model. An

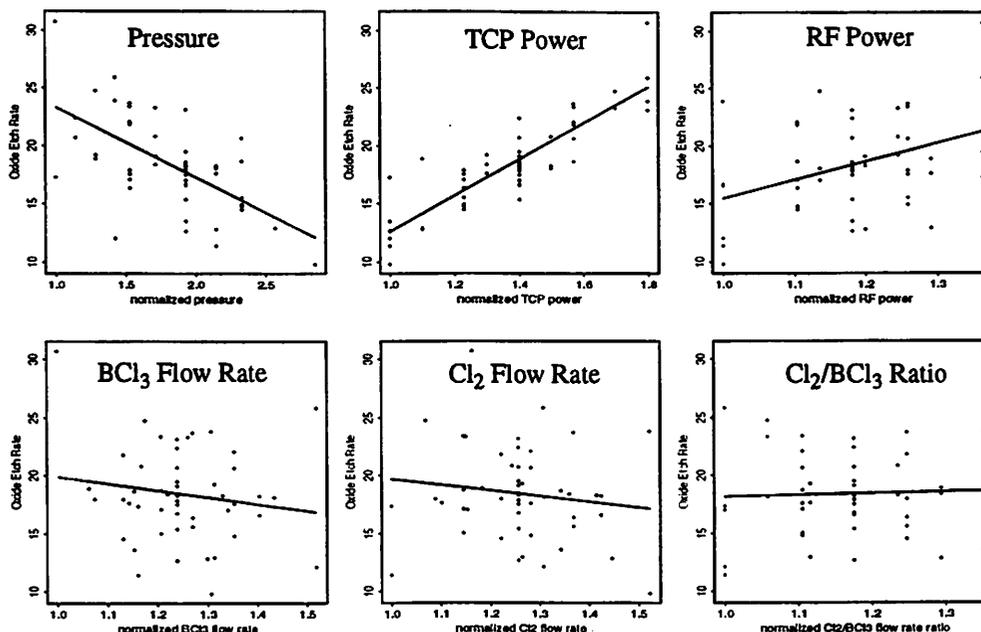


Figure 4.9 Oxide etch rate versus machine settings (each line denotes the fitted linear regression model)

adjusted R^2 of 0.43 is achieved for the model, which relates aluminum CDs to OES sensor signals, corresponding to a RSE of $0.011\mu\text{m}$.

4.5. Discussion

As indicated in the previous sections, in aluminum etch, optical emission intensity is primarily dependent on TCP coil power, and the effects of pressure and RF bias power are less significant. This can be explained as follows: despite the fact that the increase of pressure (P_r) can enhance the ground state density N as shown in Eq.(2.1), P in Eq. (2.1) tends to drop since the electron temperature decreases with decreasing $1/P_r$ [4.5][4.6]. Therefore, the optical emission intensity I ($\propto N$ and P) is less sensitive to the change of pressure. On the other hand, since electron density is proportional to TCP coil power [4.5], the increase of TCP coil power leads to a higher P (electron impact excitation function) in Eq.(2.2). Moreover, a higher concentration of neutrals can be achieved by increasing the TCP coil power due to increased dissociation of the BCl_3 . Therefore, the higher TCP coil

power tends to increase both N and P , resulting in an increase of emission intensity. This hypothesis is strengthened by Figure 4.7: since TCP coil power controls the emission intensities for all wavelengths, the correlations between TCP coil power and emission intensities show little dependence on wavelengths.

Oxide over-etch in this experiment is mostly an ion-enhanced process since low pressure (favors higher ion bombardment energies) and high TCP power (favors higher ion flux) are used to enhance oxide etch rate. By contrast, in the case of a chemically driven process, such as aluminum etch, the concentration of neutral etchant (in this case, the neutral etchants are Cl_2 and Cl) is the key factor for determining the etch rate. Because neutral etchants can also be generated through the dissociation of BCl_3 , higher pressure (more Cl_2 and BCl_3) and higher power (more BCl_3 dissociation) can both lead to an increase in the concentration of neutral etchant.

It is relatively straight forward to explain the poor correlation between CD and OES signals. First, since CD is a complex function of pressure, etchant concentration, wafer temperature, and power (favors the reaction between etchant and photoresist so generate polymer), a linear statistical modeling technique may not be sufficient to explain CD behavior, as is also suggested by White *et al.* [4.3]. Secondly, CD variation is usually an accumulation of the variation from the previous lithography and thin-film deposition processes, as well as the inhibitor-driven etch process. Because there is no pre-etch CD measurement available in this experiment, we are unable to decompose the final CD variation into lithography-induced CD variation and etch-induced CD variation. Yu *et al.* [4.4] have also concluded that CD variation is mostly attributed to the lithography step, rather than etch step. Because the OES signal only captures the variation due to plasma-induced fluctuation, it is incapable to describe the total CD variation. Lastly, CD variation is not only related to the plasma intensity, but is also a function of over-etch time. In this work, the influence of over-etch time has not been considered.

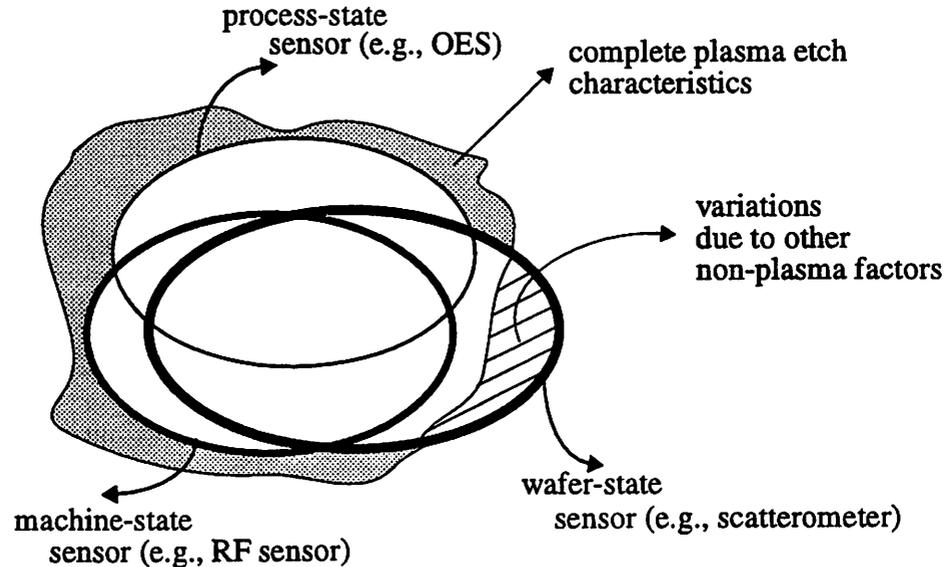


Figure 4.10 Information captured by machine-state, process-state, and wafer-state sensors

4.6. Summary

In this chapter we have demonstrated that the variation of the measured wafer-state parameters can be sufficiently captured by the measured OES only under certain circumstances. Our results suggest the following: (1) optical emission due to the excited species in plasma during Al etching is very sensitive to the variations of TCP coil power, while the portions of etch rate variation driven by chamber pressure and other machine variables are not captured well by the OES readings; (2) the variation in oxide etch rate during over-etch is mainly attributed to the TCP coil power change, and the brightness of plasma emission is strongly related to the oxide etch rate; (3) the correlation between emission intensities and oxide etch rates is not a strong function of wavelength; (4) the variation of OES signals has a direct correlation to the oxide etch rate variation which is driven by ion energy/density change. OES is less directly related to CD variation which is inhibitor controlled.

As such, a process-state sensor [4.7][4.8] such as OES will have a limited success in capturing the variations of aluminum CDs. The success of wafer-state modeling is lim-

ited by the nature of the physical and chemical mechanisms of the etch process, as well as the sensor efficiency. Specifically, OES analysis is only able to capture the changes of plasma variables (e.g., ion density, ion energy, and electron temperature) which might not be the only causes of the etch variations, as illustrated in Figure 4.10. Hence, OES is more able to capture the plasma-induced etch variation, rather than to capture the variation induced by non-plasma factors (e.g., wafer temperature). Sensor fusion [4.7] (that is, the integration of the signals collected by many different types of sensors) is thus needed to enhance the correlation between the sensor signals and wafer-state variables. An *in-situ*, real-time wafer-state sensor may be the ultimate approach to provide a direct measurement of wafer-states (e.g. CDs) for fault detection and process control.

References for Chapter 4

- [4.1] J. E. Jackson, *A User's Guide to Principal Components*, John Wiley and Sons, New York, 1991.
- [4.2] D. C. Montgomery, *Introduction to Statistical Quality Control*, 2nd ed., John Wiley and Sons, New York, 1991.
- [4.3] D. A. White, *et al.*, "Spatial Characterization of Wafer State Using Principal Component Analysis of Optical Emission Spectra in Plasma Etch", *IEEE Transaction on Semiconductor Manufacturing*, vol. 10, no. 1, pp. 52-61, 1997.
- [4.4] C. Yu and C. J. Spanos, "Patterning Tool Characterization by Causal Variability Decomposition", *IEEE Transaction on Semiconductor Manufacturing*, vol. 9, no. 4, pp. 527-535, 1996.
- [4.5] M. A. Lieberman and A. J. Lichtenberg, *Principles of Plasma Discharges and Materials Processing*, John Wiley and Sons, New York, 1994.
- [4.6] D. Manos and D. L. Flamm, *Plasma Etching*, Academic Press, London, 1989.
- [4.7] G. Barna, *et al.*, "Sensor Integration into Plasma Etch Reactors of a Developmental Pilot Line", *J. Vac. Sci. Technol. B*, vol. 12, no. 4, pp. 2860-2867, Jul/Aug 1994.
- [4.8] G. Barna, *et al.*, "Sensor Needs for IC Manufacturing", *Solid State Technology*, pp. 57-61, April, 1994.

CHAPTER 5

CD Modeling and Monitoring Using SROES for Poly-Si Etch

5.1. Introduction

The variation of poly-Si critical dimension (CD) has a significant impact in determining integrated circuit (IC) performance, because of possible device parameter drift and circuit performance degradation introduced by CD non-uniformity. For example, the microprocessor speed is a function of I_{dat} which is strongly correlated to gate CD [5.1]. In this chapter, we develop a novel approach to describe plasma-induced poly-Si CD reduction (also known as undercut) and its spatial variability using the signals collected from a spatially-resolved optical emission spectroscopy (SROES) system. A Chromex 250IS Imaging Spectrograph along with a Princeton Instruments TE-CCD-1024E thermoelectrically cooled CCD camera and an ST-130 controller is installed in the Berkeley Microfabrication Laboratory, and is set up to view the plasma across three distinct regions arranged laterally above the etching wafer. Three full-range spectra are collected simultaneously.

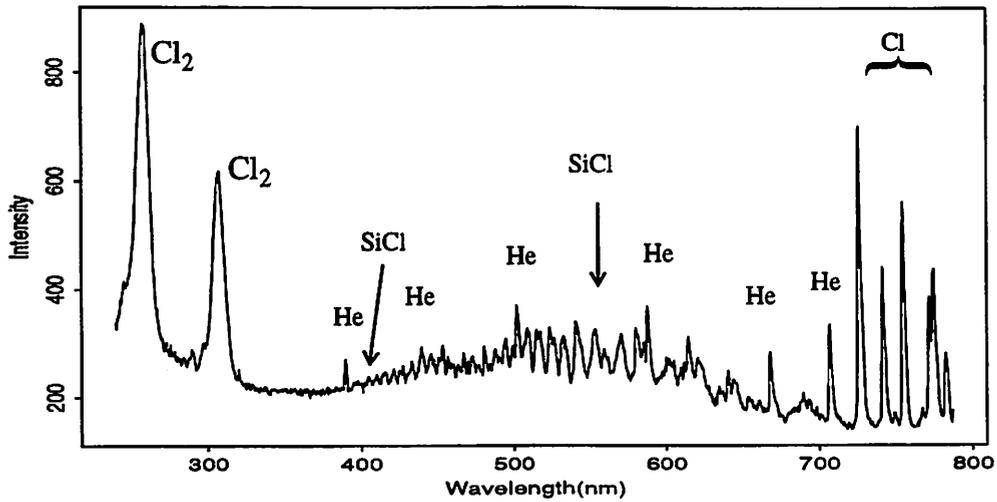


Figure 5.1 A typical spectrum collected from the Lam 4400 during the main etching (etchant is Cl_2) with spectral lines labeled with the corresponding species.

As a result, the data sets collected by this SROES system contain spatial, temporal and spectral information of plasma emission.

The separation of the statistical contribution of OES signals is a key step to determine the spatial uniformity of CDs. However, it is not easy to directly use OES spatial and temporal information because our OES sensors can only acquire signals in real-time from three¹ distinct regions, while CDs are measured off-line at many different positions on each wafer.

In Section 5.3, CD reduction is modeled using a compressed data set of OES without taking into account the spatial variation among the three beams of the OES signals. In Section 5.4, we take advantage of the wealth of spatially-resolved and real-time data that an OES system can provide to characterize the temporal and spatial behavior of the plasma and to estimate the spatial variation of within-wafer CD measurements. In other words, in addition to establishing a model to describe the wafer-average CD using OES intensities,

1. The number of beams used in this experiment is chosen after considering the hardware installation complexity, chamber geometrical limit, vertical CCD resolution, and possible acquired data size.

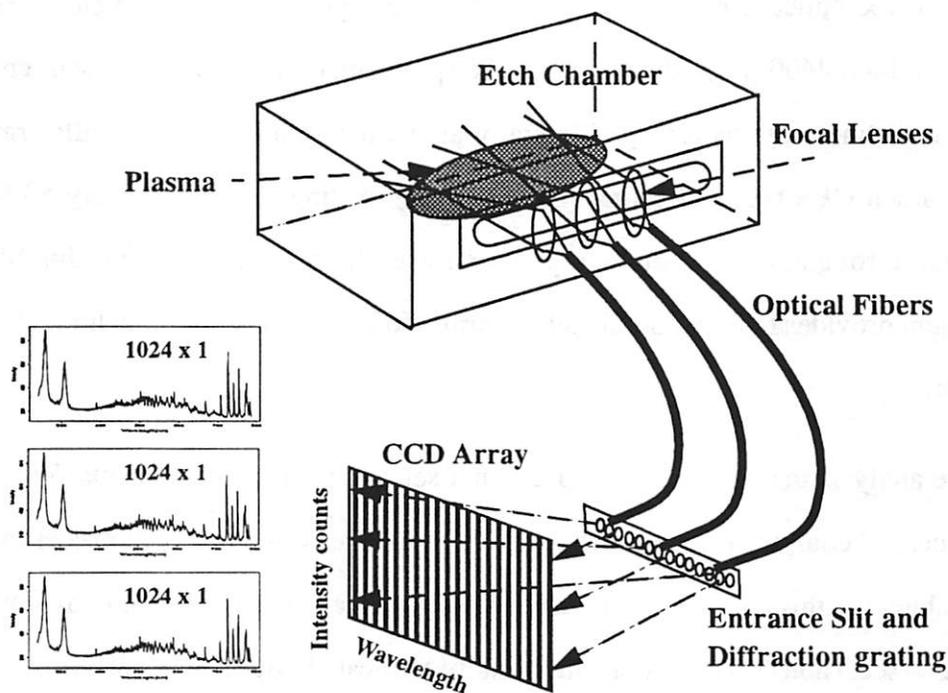


Figure 5.2 Schematic illustration of three-beam spatially-resolved optical emission spectroscopy

we also investigate the spatial and temporal emission profiles and relate them to the CD spatial variances.

Although the optical emission intensities are changing from wafer to wafer, we use the term, "OES temporal variation", to describe the variation of real-time OES signals, collected from the duration of the etch process for one wafer in this Chapter. We also use the term "CD" to describe the etch-induced CD reduction in the following sections.

5.2. Experiment to Link OES Signals to CD Variability

Three-beam *in-situ* SROES sensors have been installed on a Lam 4400 Rainbow plasma etcher. The sensor readings are collected simultaneously throughout the etching process at an acquisition rate of 1 Hz. We have chosen a 150 groove/mm grating in order to monitor the 240nm to 790nm spectra with a resolution of 1.2nm/pixel, as shown in

Figure 5.1. Three optical fibers are connected from the spectrometer to the etcher reactor viewport of a Lam 4400 polysilicon etcher. The spectrometer is set up to view the plasma across three distinct regions arranged laterally above the etching wafer, as illustrated in Figure 5.2. Each OES frame contains 3072 (1024x3) readings. Approximately 50 frames are collected throughout the main etch process for each wafer. As a result, this SROES sensor system provides both temporal and spectral information in addition to limited spatial information.

The analysis that follows was based on a set of designed experiments. We used a fractional central composite design with 9 center points, resulting in a total of 37 runs. The input variables for this experimental design are RF power, pressure, gap, total flow, and Cl_2/He ratio (see Table). The wafers are initially fabricated with a poly-Si film of 4500Å on buffer oxide layer of 1500Å, before further lithography and etching processes are carried out. The photomask was fabricated using a 2-micron design rule. Linewidths are optically measured off-line both before and after the etching process with nine points recorded for each wafer, as depicted in Figure 5.3. Once these measurements are complete, CD reduction can be estimated by subtracting the poly-Si linewidth from its corresponding photoresist linewidth, see Figure 5.4. Electrical CD measurement is also conducted for the poly-Si

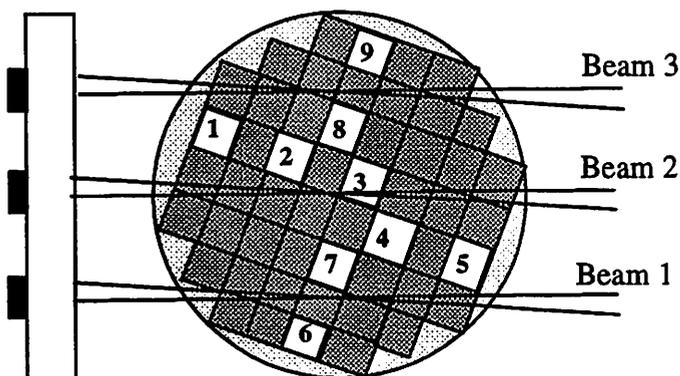


Figure 5.3 Top view of spatially-resolved OES

linewidth to verify and calibrate the optical CD measurements. The spatial variance of CD measurements for one wafer is defined as

$$\frac{\sum_{i=1}^n (CD_i - \overline{CD})^2}{n-1} \quad (5.1)$$

where i is the index of CD measurement ranging from 1 to 9, $n=9$ is the total number of measurements², and \overline{CD} is the “etch-induced CD reduction”. A gauge study [5.2] was also conducted to ensure adequate CD inspection capability. This study showed that approximately 4.5% of the overall CD variation is due to measurement error, while the rest of the variation is induced by the process.

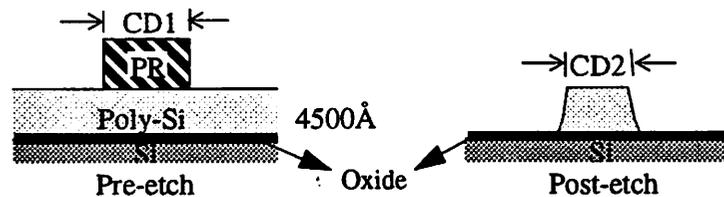


Figure 5.4 Cross sections of the test structures before and after etching (PR stands for photoresist)

Table 5.1
Central Composite Experimental design - nominal and extreme settings

| Design variables | RF Power (Watt) | Pressure (toor) | Gap (cm) | Total flow (sccm) | Cl ₂ /He ratio |
|-----------------------------------|-----------------|-----------------|----------|-------------------|---------------------------|
| Central settings | 275 | 425 | 0.8 | 580 | 0.42 |
| Change in percentage from nominal | 14.5% | 13% | 12.5% | 7% | 7% |

2. Only nine points were measured for each wafer in order to reduce the measurement time while retaining acceptable spatial information., as shown in Figure 5.3

5.3. Models of Average CD Reduction

This section describes the procedures of the data analysis for two different modeling methods. To perform this analysis, all experimental data are transformed into a matrix format. Let a 37×1 array $Y = [y_1, y_2, \dots, y_{37}]^T$ and a 37×1024 matrix X represent the sets of CD measurements of 37 experimental wafers and the measured 1024 optical emission spectral channels corresponding to the 37 wafers (that is, only one spectral frame for each wafer), respectively. X_{ij} is the value of the j th OES channel for the i th wafer, where $i = 1, 2, \dots, 37$ and $j = 1, 2, \dots, 1024$. The X and Y matrices in this study are not scaled.

5.3.1. Principal Component Regression Model

In this analysis, we first estimate the number of principal components that should be retained in the model by performing a screeplot analysis. Figure 5.5(a) shows that the first 4 principal components are adequate to explain 99.5% of process variation. We retain this percentage of variance in order to ensure that some of the variation contributed from the weak³ spectral lines is also included by the selected principal components. Although many weak spectral lines associated with the chemical species (e.g., SiCl) related to the etching performance are relatively weak as observed in Figure 5.1, they might be physically significant and should not be neglected.

The number of principal components retained in the model can be also verified by cross-validation. A Prediction-Error-Sum-of-Squares (PRESS) statistic is employed by a jackknife routine [5.3] shown in Appendix A to confirm that 4 principal components give the least PRESS, as shown in Figure 5.5(b). It also shows that the inclusion of the 5th PC significantly degrades the PRESS in this PCA analysis.

3. In comparison to some of the strongest lines such Cl₂.

Next, these four principal components are employed in a regression model against the original response Y . The model is

$$\hat{Y}_i = \mathbf{z}_{ij}\hat{\beta}_z = 0.66 + 0.00052z_{i1} - 0.0021z_{i2} + 0.0014z_{i3} - 0.00082z_{i4}, \quad (5.2)$$

where $i = 1, \dots, 37$. It is interesting to note that the 4th principal component has the highest coefficient comparing with others. Also, the PCR of our OES data shows that the first principal component accounts for 88.3% of the total variability of the predictors, the second principal component accounts for about 7.7%, the third one accounts for 3% and the last one for 0.5%. However, the fourth principal component explains more variability in the response than in any other principal components, as listed in Table 5.2. Therefore, one should not expect principal components to explain variability in responses in the same order that they do for the variables from which they are formed [2.19]. An account for the variability of the predictors and responses explained by the principal components is given in Table 5.2.

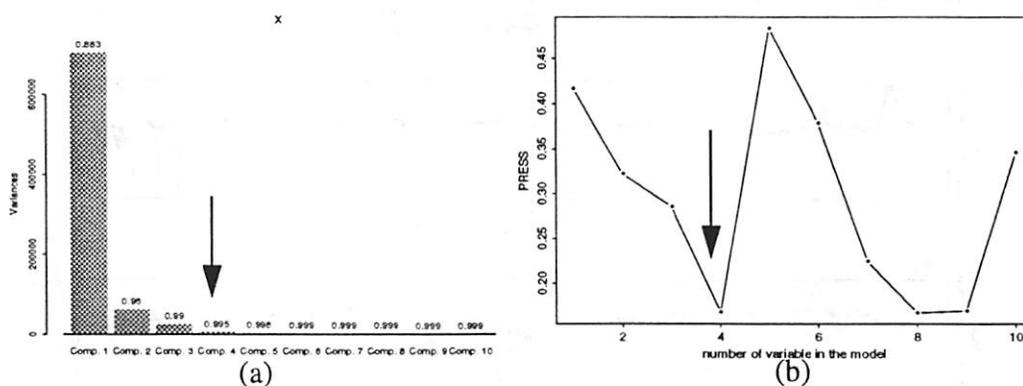


Figure 5.5 (a)The screenshot of OES data; (b) PRESS-statistics

Table 5.2
Proportion of predictor and response variability explained by principal components

| | X | X | Y | Y |
|--------------|------------|------------|------------|------------|
| Explained by | individual | cumulative | individual | cumulative |
| z_1 | 88.3% | 88.3% | 0.5% | 0.5% |
| z_2 | 7.7% | 96% | 35.8% | 36.3% |
| z_3 | 3% | 99% | 2.7% | 39% |
| z_4 | 0.5% | 99.5% | 28% | 77% |
| Unexplained | 0.005% | 100% | 33% | 100% |

Because the 2nd and 4th principal components appear to reveal a strong effect on the responses, it might be worth to see the loading plots (*i.e.*, the coefficients of the original variables) of all four components so as to distinguish the important spectral lines which might be related to the etching performance. Figure 5.6 indicates that 387nm, 500nm,

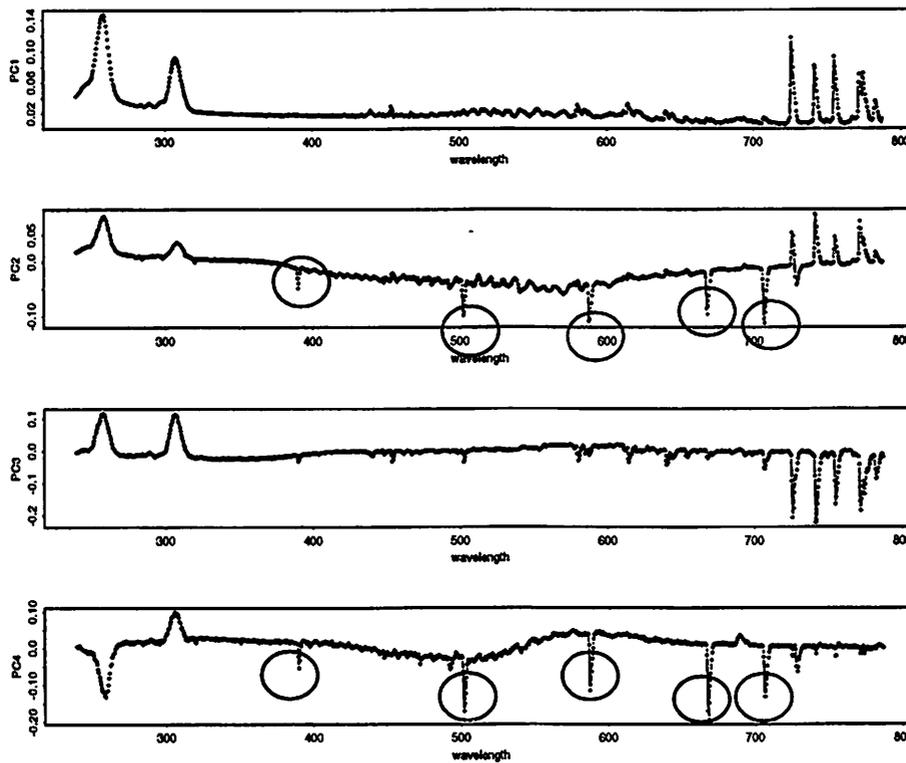


Figure 5.6 Loading plots of the first 4 principal components (the spectral lines marked are difficult to observe in the original spectrum)

583nm, 663nm and 703nm peaks are the most significant features, although they are not revealed by observing Figure 5.1. PCR, in this case, is useful in filtering the spectral data so as to identify the features which are difficult to observe in the original spectrum, as circled in Figure 5.6.

The prediction error sum of squares (PRESS) of this fitted model can be determined by a jackknife iteration [5.3]. The result shows that the value of PRESS of this model is about 0.168, which corresponds to a standard error of prediction (SEP)⁴ of 0.2 μm . By comparison, the average standard deviation of within-wafer and wafer-to-wafer CD measurement are about 0.06 μm and 0.13 μm , respectively. These tests suggest that the PCA model based on the OES signals can not fully explain the CD variations. This will be further addressed in Section 5.5.

Another modeling technique, PLSR, will be used in the next section.

5.3.2. Partial Least Squared Regression Model

The PLSR technique operates in somewhat the same way as PCR in that a set of vectors are obtained from the predictor variables. However, what makes PLSR different is that as each vector is obtained, it is immediately related to the response and the reduction in variability among the predictors. The estimation of the next vector takes that relationship into account. Simultaneously, a set of vectors for the responses is also being obtained that also takes this relationship into account.

To determine how many variables are significant and should be retained in the modeling, a PRESS statistic can be employed. It indicates that 5 latent vectors are needed to describe the predictors in order to achieve minimum PRESS, as seen in Figure 5.7. The PLSR model can be written as

4. $SEP = (PRESS/m)^{1/2}$, where m is the total number of PCs used in the prediction model.

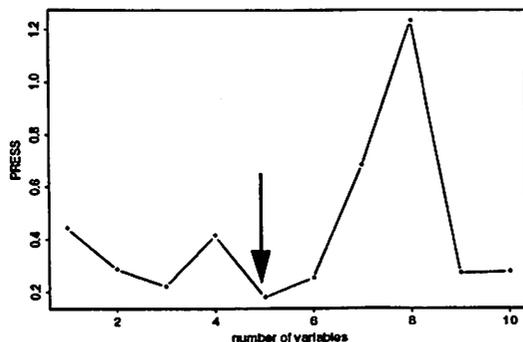


Figure 5.7 PRESS vs. the number of variables in model

$$\hat{Y}_i = l_{ij}\hat{\beta}_l = 0.662 + 0.0001l_{i1} - 0.0001l_{i2} + 0.0005l_{i3} - 0.0003l_{i4} + 0.01l_{i5} \quad (5.3)$$

where $i = 1, \dots, 37$ and l_j is the j th latent variable. The residual sum of squares (RSS) of this PLSR prediction model is only 0.0493 with 31 degrees of freedom, smaller than that derived from the PCR model which is 0.062. This is expected because we use both X and Y to construct the model. It is important to note that the purpose of this model is prediction, while another set of OES sensor data can be projected on the previously constructed latent structure. Although we do not have additional data set to examine the prediction accuracy of this model, jackknife is one approach to estimate the prediction error. Using a jackknife routine shown in Appendix A, the prediction error sum of squares is about 0.2, corresponding to a SEP of $0.2\mu\text{m}$, which is similar to that in the PCR model.

5.3.3. Comparison of Models

For the data set used, no one modeling method clearly stands out as the “best” prediction model for CD reduction. Both PCR and PLSR techniques can reduce the parameter space of the models, although each one of them uses a different approach. The PCR model has a RSS value of 0.0616 and corresponds to an R-square of 0.67, and a PRESS value of 0.168. Four principal components are used and each one has a different correlation to the

response parameters (*i.e.*, CD reduction). This information is helpful in the spectral analysis since many small but important atomic spectral lines are often observed by stronger molecular spectral bands. In our case, we find that the 2nd and 4th principal components capture significant spectral lines which are not prominent in the original data (see Figure 5.6). A comparison of the modeling results is summarized in Table 5.3.

Table 5.3
Comparison of modeling methods for CD reduction

| | RSS | | R^2 | $adj.R^2$ ^a | SEP (μm) |
|------|-----------------------------|---|-------|------------------------|-----------------|
| | # of variable used in model | | | | |
| PCR | 0.062 | 4 | 0.67 | 0.63 | 0.2 |
| PLSR | 0.049 | 5 | 0.8 | 0.77 | 0.2 |

a. Adjusted R^2 is a modified version of the R^2 statistics that considers the number of parameters of the model and is given by $adj. R^2 = 1 - (1 - R^2)n / (n - p)$, where n is the number of observations and p is the number of parameters.

Up to this point, we only use the average plasma emission intensity to describe the etch state for each wafer and no real-time plasma emission information is included in this modeling framework. From the next section, we will start to include the real-time information and relate it to the CD spatial variation.

5.4. Monitoring and Modeling of CD Spatial Variation

In this section, we illustrate how to determine and summarize the statistical spread associated with the acquired OES sensor readings. The variation of OES data can be characterized by several categories including spectral variation (due to the emission intensity difference at different wavelengths), temporal variation (due to the optical emission intensity changes over time), and spatial variation (due to the intensity difference for the spectra collected at different regions in the plasma). In this study, a typical OES data set collected from all wafers contains 1024 spectral variables \times 50 temporal variables \times 3 spatial vari-

ables \times 37 wafers. Obviously, it is not easy to utilize all the information contained in this large number of variables.

We propose a simple approach to decompose the variation of OES signals. This approach is to determine the deviations of the intensities of selected wavelengths from their temporal and/or spatial baseline trends. Since poly-Si is etched by Cl_2 and Cl [5.5] and the S/N ratio of Cl_2 line ($\lambda \sim 262\text{nm}$) is much higher than that of the rest of the spectral features, the intensities of Cl_2 reveal useful etch information. Therefore, the Cl_2 line is selected to represent the entire spectrum to simplify the analysis and enhance the computational efficiency.

In the following sections, both physical and statistical approaches are employed to extract the spatial and temporal variation from the original OES sensor readings for a selected Cl_2 spectral line.

5.4.1. Extraction of temporal variation of OES signals

In the past, an Auto-Regressive Integrated Moving Average (ARIMA) time-series model was employed to filter the non-stationary sensor data time-series by a baseline model and to convert them into a more stationary data series so that a control chart can be employed [5.6][5.7]. However, although this technique can capture the machine faults associated with short-term fluctuation, the difficulty of detecting the optical emission long-term time drift problem (for example, due to window clouding effect) limits its application. In this section, we propose an alternative approach to extract the temporal variation.

The temporal variance for the Cl_2 spectral line is determined as follows. Since the OES trajectory is not a stationary time series (see Figure 5.8), the OES data must be filtered from a baseline trend. The wafer with a minimum CD spatial variance and a nominal recipe is selected as a baseline wafer. The first step in the algorithm is to remove the intensity dis-

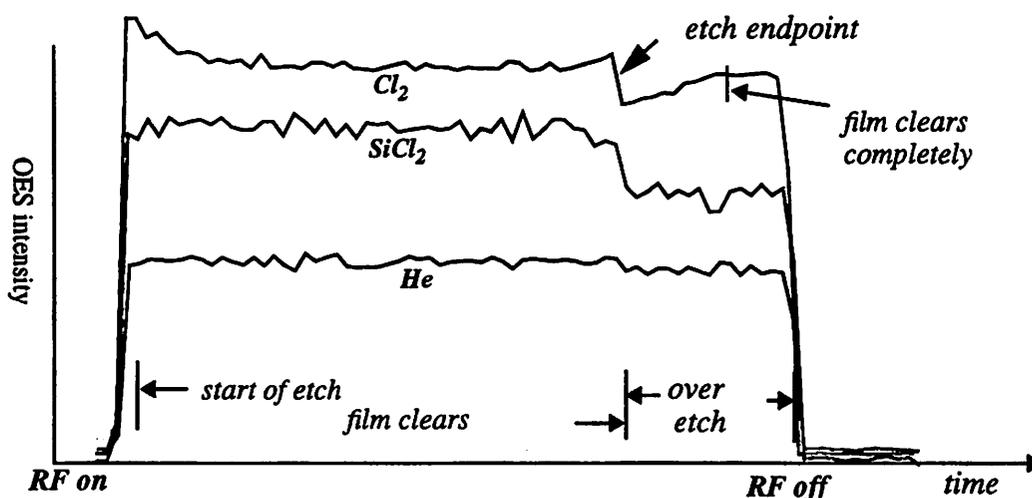


Figure 5.8 Typical OES trajectories for different spectral lines.

continuity due to the recipe change (we reduce the power level for the over etch step) by mean shifting, and then model the baseline trend of the OES signals by a locally weighted regression smoother⁵ [5.8], as shown in Figure 5.10. The baseline trend is then normalized between 0 and 1 to prevent the intensity drift. Once the baseline behavior has been established, the normalized OES data from the other wafers are subtracted from the trend model. The deviations from the trend in the etch time frame suggest a plasma disturbance or a chemical reaction discontinuity. The estimated temporal variance of the OES signals at a selected spectral line can be expressed as

$$\frac{\sum_{i=1}^n (y_i - y_i^{Trend})^2}{n-1} \quad (5.4)$$

where y_i is the OES intensity at time i , n is the total number of sequential observations, and y_i^{Trend} is the trend model at time i , as given by

5. This particular smoothing algorithm uses locally linear fits. A window is placed about each time point and data points that lie inside the window are weighted so that nearby points get the most weight and a weighted regression is used to predict the value at that point. A more complete description of this smoothing algorithm will be provided in Appendix D.

$$y_i^{trend} = \sum_{j=1}^{N(y_i,0)} y_{i,j} W\left(\frac{|y_{i,0} - y_{i,j}|}{\max_{N(y_i,0)} |y_{i,0} - y_{i,j}|}\right) \quad (5.5)$$

where j indexes the near neighbor of each data point, $N(y_{i,0})$ is the total number of the near neighbors of point $y_{i,0}$, W is the weight assigned to each point of a neighborhood with a tri-cube weight function (i.e., $W(u) = (1-u^3)^3$ for $0 < u < 1$; $W(u) = 0$, otherwise). The values derived from Eq.(5.4) can be then used to represent the temporal variation of the OES signals, as illustrated in Figure 5.10.

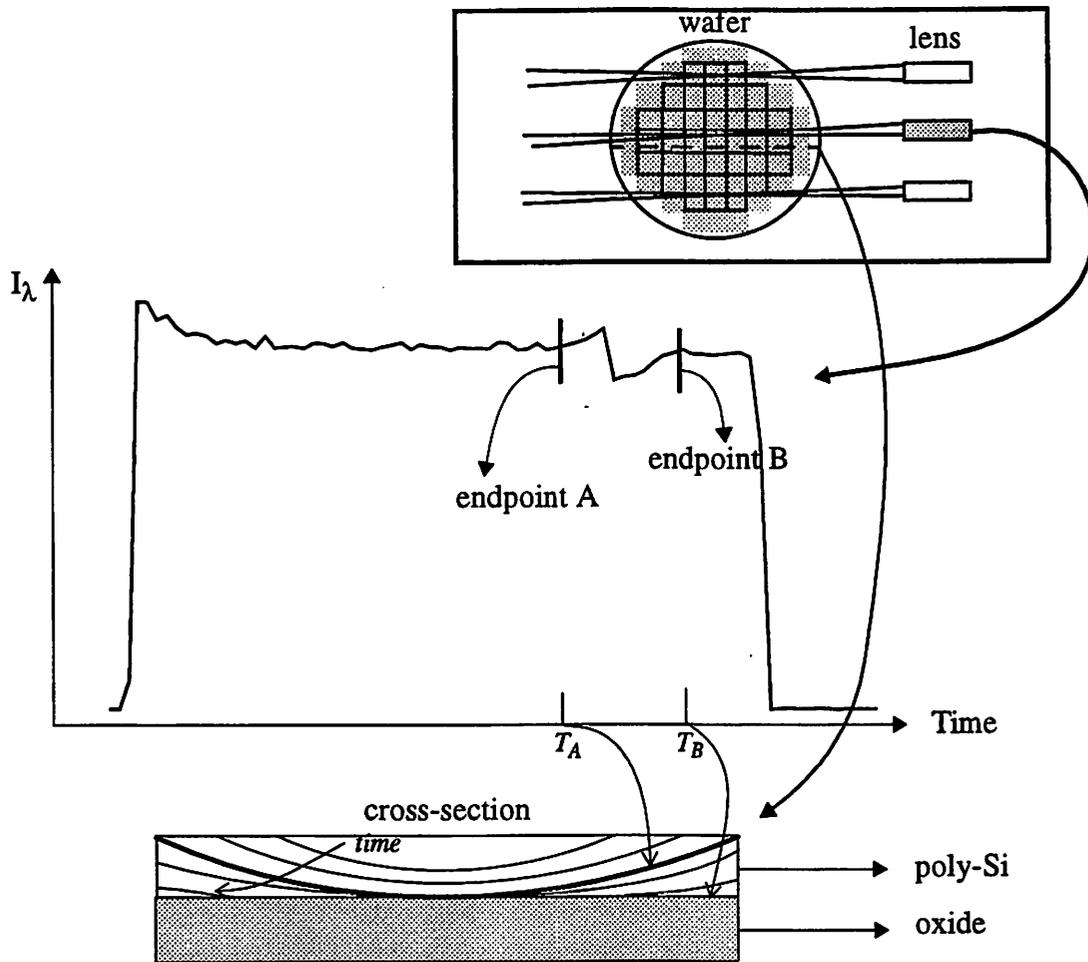


Figure 5.9 Poly-Si etched spatial profile vs. OES trajectory.

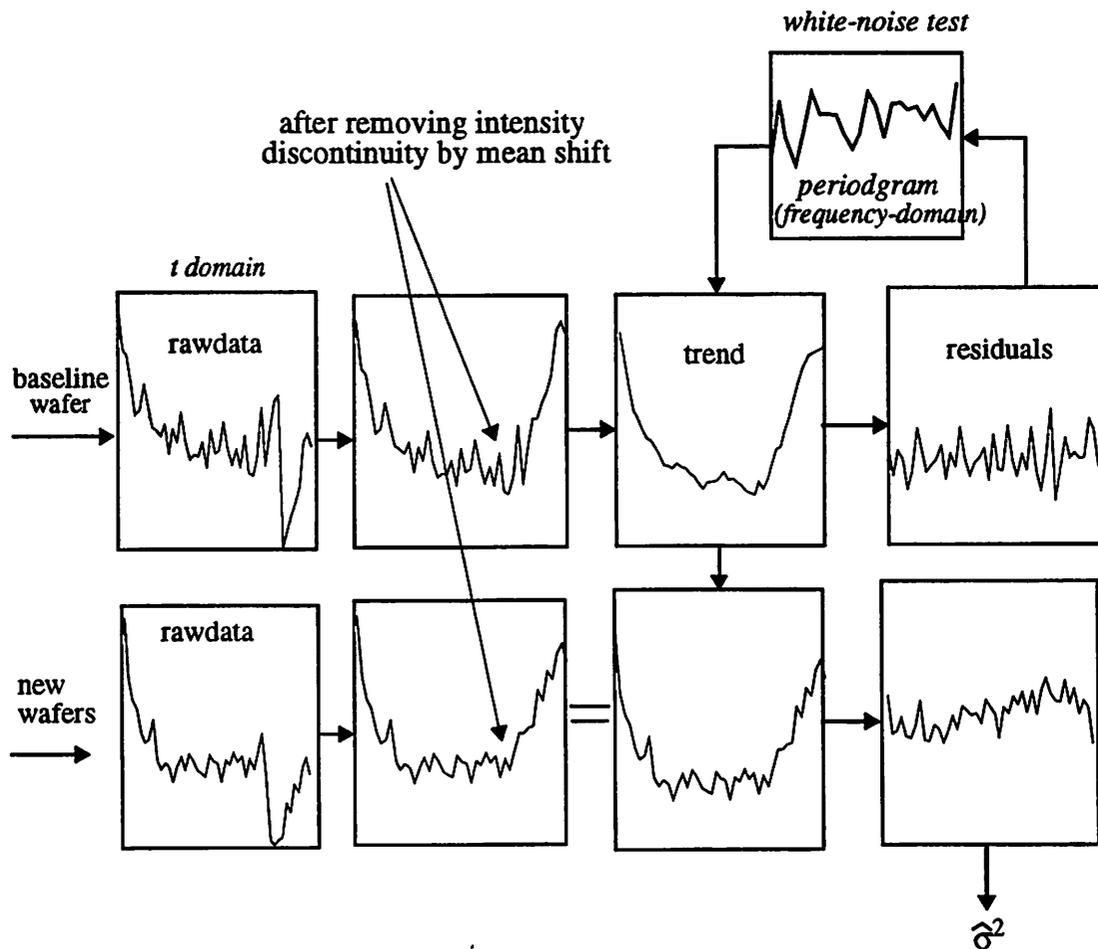


Figure 5.10 Temporal variance extraction from the OES trajectories.

One mechanism that might relate the spatial CD signature to the temporal OES signature is given below: As illustrated in Figure 5.9, point A in the emission trajectory signifies the time when the underneath oxide starts to expose. Its corresponding poly-Si thickness profile is shown in Figure 5.9. Point B in the same emission trajectory implies that the poly-Si layer is fully cleared at T_B . The transition between point A and B indicates the spatial uniformity of the etch process. Generally, a shorter transition means a better etch uniformity. Figure 5.11 shows that the A-B transition time of a normal wafer is shorter than that of an abnormal wafer.

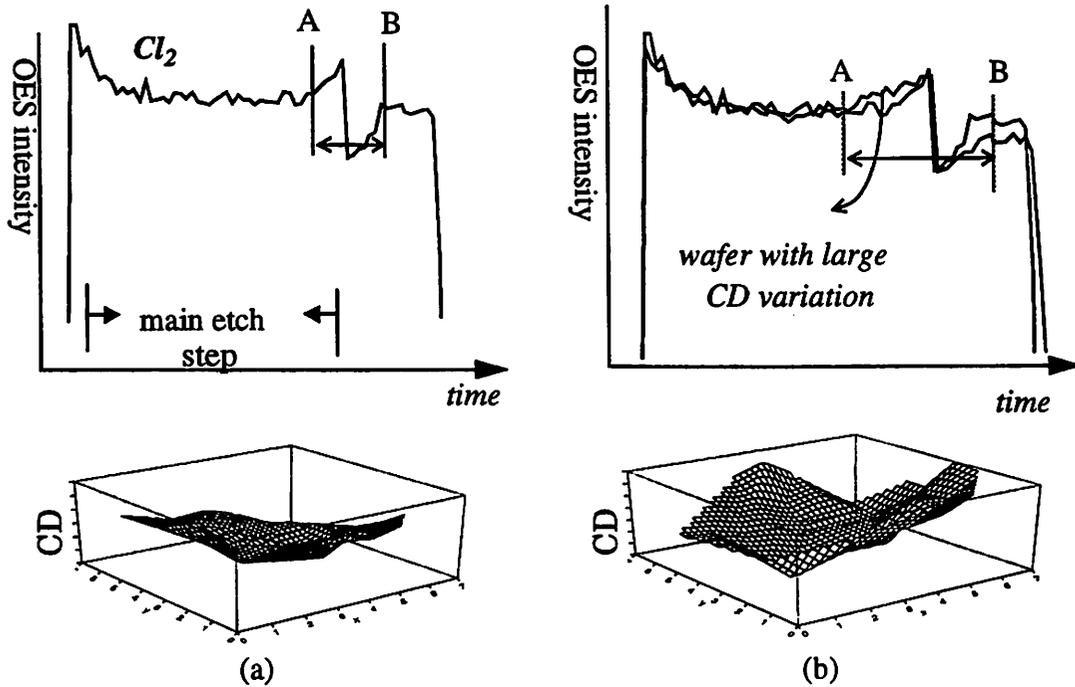


Figure 5.11 The OES trajectories and CD plots for (a) wafer #2; and (b) wafer #30

5.4.2. Extraction of spatial variation of OES signals for a three-beam sensor

The spatial variation of OES signals is often considered as the intensity differences among the three beams at a selected spectral feature. However, it is important to note that the different intensities among the three beams are mostly caused by the sensor collection efficiency and the unequal plasma volume viewed, rather than the plasma density spatial non-uniformity. This problem is usually compensated by actinometry, a technique described in Section 2 [5.9][5.10]. In this context, actinometry is used to determine the *absolute* concentration of chemical species of interest only if the chemicals are suitable and the ratio of P_M and P_Z shown in Eq. (2.6) is a constant. In the Cl_2/He plasma, it is not ideal to compare the absolute concentrations in order to determine their spatial distribution. Nonetheless, since the *relative* variation of the plasma density is more important than its absolute value in the application of process control, we propose an alternative approach to

describe the *relative* wafer-to-wafer variations of spatial distribution of plasma density and energy without determining their absolute concentrations.

Consider the emission ratio of the species of interest Z to that of the second inert species M:

$$\left(\frac{I_Z}{I_M}\right)_i = \left(\frac{N_Z P_Z Q(\lambda_Z) A_Z K}{N_M P_M Q(\lambda_M) A_M K}\right)_i \equiv C \cdot \left(\frac{N_Z P_Z}{N_M P_M}\right)_i \quad (5.6)$$

where N , P , $Q(\lambda)$, A , and K are defined in Eq. (2.1), and i is the beam index ranging from 1 to 3. For each beam, K is identical and thus can be cancelled. The ratio A_Z/A_M is independent of the beam used and can be considered as a constant C . The apparatus function Q also cancels out if the wavelengths of the two lines are similar (i.e., within a few nm to each other). In our case, the ratio of P_{He} and P_{Cl} is a complicated function of plasma variables (see Eq. (2.2)), and hence can not be simplified as a constant. Therefore, the spatial variation of emission is related not only to the ground density ratio of species Cl and He , but also to the ratio of their electron impact excitation functions, and can be written as:

$$spatial\ variation \sim \left(\frac{N_{Cl} P_{Cl}}{N_{He} P_{He}}\right)_i / \left(\frac{N_{Cl} P_{Cl}}{N_{He} P_{He}}\right)_j \quad (5.7)$$

where $i \neq j$. To simplify this analysis, we assume $N_{He} P_{He}$ is uniform in the chamber since He is non-reactive throughout the etch process. As a result, Eq. (5.7) can be rewritten as

$$spatial\ variance \sim \frac{(N_{Cl} P_{Cl})_i}{(N_{Cl} P_{Cl})_j} \quad (5.8)$$

However, the remaining challenge for this techniques is how to decouple N and P from Eq. (5.7) so that the spatial information of ion density, electron density, and electron

temperature can be extracted to map CD spatial variances for this Cl_2 -He plasma. This difficulty limits the application of actinometry in this study.

Alternatively, an empirical approach is used to extract the spatial variation of plasma emission. Similar to the method we used to extract the temporal variation from OES signals, the baseline spatial behavior of selected λ (in this case, we use the Cl_2 line) is also established from the baseline wafer used in Section 4.1. The “spatial trend” consists of only three points normalized between 0 and 1. Then the spatial curves extracted from other wafers are subtracted from this trend curve. The deviation from the trend suggests a spatial fluctuation of the plasma. We use a similar expression as shown in Eq. (5.4) to describe the “spatial variance” of plasma (in this case, n is equal to 3). The physical normalization approach, in which the emission intensity is converted to plasma density, thus becomes unnecessary. An overview of this approach is illustrated in Figure 5.12. Note that the temporal variance of emission intensities determined in Section 5.4.1 is an indicator of the temporal variation of chemical reaction, while the “spatial variance” of emission intensities determined in this section is more of an indicator of the behavior of the plasma spatial uniformity.

5.4.3. Modeling Results

The extracted spatial and temporal variances are used to estimate the spatial variance of the CD measurements. The results show that most of the estimated CD spatial variance is explained by the OES temporal variance, rather than the spatial variance. Specifically, a least-squares regression fit for wafer-to-wafer CD spatial variance shows that an adj. R^2 of 0.82 is achieved by the OES temporal variance alone, while an adj. R^2 of 0.22 is achieved by the OES spatial variance alone. In this section, we only present the

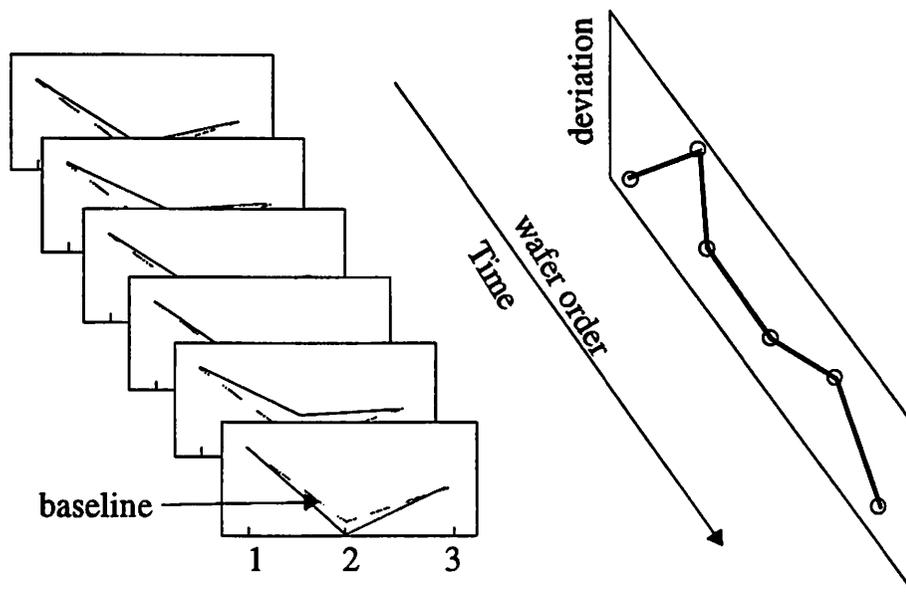


Figure 5.12 Extraction of OES spatial variance.

results derived from the Cl_2 line ($\lambda=260\text{nm}$). Nonetheless, this approach can be easily extended to other wavelengths as well.

It is hypothesized that the spatial variation of the normalized OES readings is caused by plasma non-uniformity during the etch process, while the temporal variation of OES readings is either due to the plasma disturbance, the photoresist profile variation, or poly-Si film thickness variation, as tabulated in Table 5.4. Because the OES *spatial* variation has only a small influence on the resulting CD spatial variation in terms of their low correlation, this indicates that the plasma non-uniformity may not be the major cause of CD spatial variation. Instead, the variation resulting from the previous lithography/film deposition steps, which can be detected by the OES trajectory discontinuity, is more responsible for the post-etch CD spatial variation.

Table 5.4
Possible physical causes of the decomposed OES variation.

| Variation component | Possible Physical Causes |
|---------------------|---|
| Temporal | plasma energy spatial non-uniformity, plasma density spatial non-uniformity, species concentration spatial non-uniformity, endpoint detector malfunctions, poly-Si film/PR thickness non-uniformity |
| Spatial | plasma energy spatial non-uniformity, plasma density spatial non-uniformity, species concentration spatial fluctuation |

5.4.4. Monitoring of CD Spatial Variation

The previous technique provides a promising way to monitor the wafer-to-wafer CD spatial variance. Because the extracted components of OES variances described in Section 5.4.1 are filtered from the normalized baseline trend models, the deviations from these trends can be also used to establish conventional control charts, based on the assumption that these deviations are normally distributed random variables. This assumption is later confirmed by the Fourier analysis [5.4], as seen in the periodogram in the top of Figure 5.10. Moreover, because the data decomposition allows us to model many different types of variations, the resulting fault detection tends to be more complete compared with the time-series ARIMA approach [5.7] which only allows us to monitor temporal variation.

To demonstrate this technique's fault detection capability, we monitor the OES temporal variance by a χ^2 control chart⁶. These extracted variations generally exhibit the wafer-

6. χ^2 control chart is based on a hypothesis test that the variance of a normal distribution equals to the baseline variance. The test statistics for this hypothesis follows a chi-square distribution [5.2]. This test is very useful in many quality-control applications [5.2]. For example, consider a normal random variable with variance σ^2 . If σ^2 is less than or equal to some value, say the baseline σ_0^2 , then the natural inherent scatter of the process will be well within the design requirements. However, if σ^2 exceeds σ_0^2 , then the natural scatter in the process will exceed the specification limits. Note that this test is based on the assumption that variance of the baseline process is known, and that the variables are *i.i.d.* normal.

to-wafer variation arising from such things as deviations of the emission intensity of Cl_2 line from their specified trajectories. A χ^2 control chart is used for each extracted variance with a control limit of

$$\chi^2_{\alpha/2, n-1} \cdot \frac{\sigma_o^2}{n-1} \quad (5.9)$$

where $\chi^2_{\alpha/2, n-1}$ is the $\alpha/2$ percentage point of the chi-distribution with $n-1$ degrees of freedom, σ_o^2 is the sample variance of baseline wafer, and n is the sample number for each wafer (n is typically larger than 50 in this case).

The results show that the OES sensors can detect 5 out of 6 wafers having an abnormal spatial CD uniformity, as shown in Figure 5.13. The top chart in Figure 5.13 shows the actual measured spatial CD variance over wafers, while the bottom chart in Figure 5.13 shows its corresponding OES temporal variance.

5.5. Discussion and Summary

In this chapter, we demonstrate that run-to-run wafer state model using an in-line sensor is effective in capturing the process variation and explaining 63%~77% of the esti-

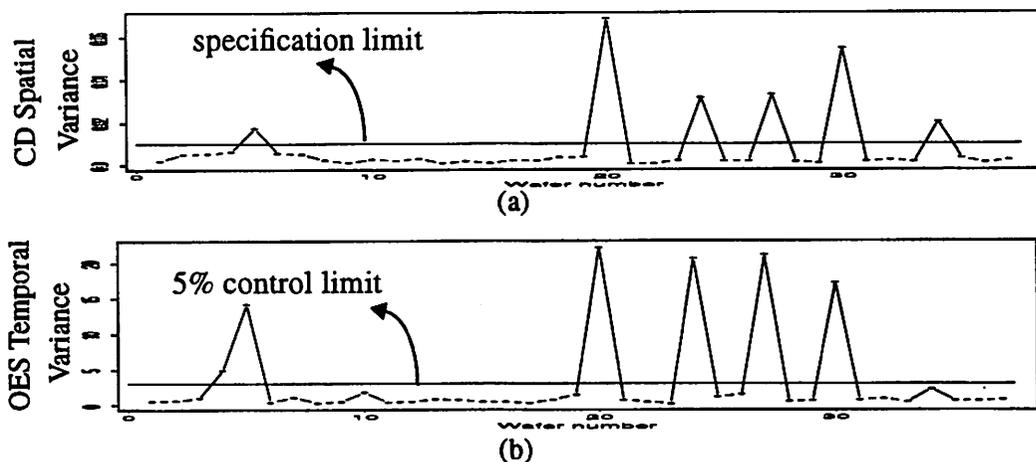


Figure 5.13 The SPC chart of (a) CD spatial variance; (b) OES temporal variance

estimated CD variations. Two different modeling techniques, principal component regression and partial-least square regression are used to predict the CD reductions of semiconductor wafers using OES sensor signals. Neither modeling method is overwhelmingly better than the other. PCR and PLSR show a similar modeling capability. 67% and 80% of the CD variations can be explained by PCR and PLSR, respectively. Since there is no other set of data to verify the prediction accuracy for each model, a jackknife method is used to estimate the prediction error sum of squares. The results indicate that both modeling approaches have a similar SEP.

Although the results of the wafer-to-wafer poly-Si CD models are acceptable in terms of their resulting R^2 values, the goodness-of-fit tests and the high SEP values suggest that these models are not good. This may be attributed to two factors. First, as suggested by White *et al.* [4.3], the linear modeling approach is not sufficient to explain the CD variation. They suggest that other non-linear approaches such as non-linear partial least squares analysis and neural network models appear to show better modeling results. Second, as described in Chapter 4, the link between CD reduction and plasma emission intensity may not be significant so that the models based on the OES signals can not fully explain the CD variations. It is also important to note that the logarithmic emission intensity has also been investigated as the input space in our modeling schemes, but the results show no improvement.

Nevertheless, we have demonstrated that CD within-wafer spatial variations, rather than the CD within-wafer averages, can be more readily detected using OES sensor signals. It has been shown that the single-wavelength approach applied to the OES sensor data can be used to provide new insights with existing process knowledge. The results indicate that the majority of the CD spatial variance can be explained by OES temporal profile behavior. However, since CD spatial variation is not only related to the etch process but also to the

previous process step as well, OES sensors and conventional statistical techniques are less able to reveal enough information, unless the temporal variation of OES data is also included in the modeling.

The next challenge is to implement this technique into manufacturing sites with tighter process windows. Since the wafers in the IC Fabs are mostly processed under the same “recipe”, the variations of CD and OES are less significant, compared with the case in our designed experiment. This may lead to some degree of difficulty in implementation. Additionally, the three-beam approach is often insufficient to describe the spatial distribution of plasma. Therefore, we propose an experiment using an alternative scanning SROES technique in order to establish the observability of spatial OES non-uniformity to plasma and wafer characteristics with less calibration problems and more spatial resolution. This technique will be introduced in the next chapter.

References for Chapter 5

- [5.1] T. Boynton, W. Yu, and J. Pak, "Gate CD Control for a 0.35mm Logic Technology", 1997 IEEE International Symposium on Semiconductor Manufacturing Conference Proceedings, San Francisco, California, Oct. 6-8, 1997, p. F-9.
- [5.2] D.C. Montgomery, *Introduction to Statistical Quality Control*, 2nd ed., John Wiley and Sons, 1991.
- [5.3] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, New York, 1993.
- [5.4] W. N. Venables and B. D. Ripley, *Modern Applied Statistics with S-plus*, Springer-Verlag, New York, 1994.
- [5.5] D. Manos and D. Flamm, *Plasma Etching*, Academic Press, London, 1989.
- [5.6] S. F. Lee and C. J. Spanos, "Prediction of Wafer State After Plasma Processing Using Real-time Tool Data", *IEEE Transaction on Semiconductor Manufacturing*, vol. 8, no. 3, pp. 252-261, 1995.
- [5.7] S. F. Lee et al., "RTSPC - A Software Utility for Real-Time SPC and Tool Data Analysis", *IEEE Transaction on Semiconductor Manufacturing*, vol. 8, no. 1, pp. 17-25, 1995.
- [5.8] J. S. Simonoff, *Smoothing Methods in Statistics*, New York, Springer-Verlag, 1996.
- [5.9] M. Lieberman et al., *Principals of Plasma Discharges and Materials Processing*, John Wiley and Sons, New York, 1994.
- [5.10] J. Coburn and M. Chen, *J. Vac. Sci. Technol.*, vol. 18, 1981, p.683.

CHAPTER 6

Scanning Spatially Resolved OES

6.1. Introduction

While conventional endpoint detection techniques often provide temporally-resolved information about the etch process, they do not provide any spatial resolution. With the emergence of 300mm wafer production, the utility of a spatially-resolved endpoint detector is expected to increase greatly. While multi-beam SROES is a very promising method for detecting spatial plasma uniformity, its promise has not yet been fulfilled. There are several problems that need to be addressed. First, the intensity differences between OES beams due to the asymmetry of grating efficiency can not be easily normalized. Secondly, when only few beams are used, they are often insufficient to describe the spatial distribution of plasma. Therefore, we propose an experiment using a *scanning* SROES technique in order to observe plasma non-uniformity. We expect to face fewer calibration problems and enjoy better spatial resolution, using more than 20 viewing “chords”.

The optical emission signatures acquired from scanning SROES will be decomposed and normalized so as to map them onto the etch characteristics. We are also compar-

ing the spatially-resolved emission intensity profiles with the machine endpoint traces collected by the etcher. Experiments have been carried out in the Berkeley Microfabrication Laboratory in cooperation with Lam Research. The objective of this work is to use these techniques for providing direct etch spatial uniformity information, as well as for spatially-resolved endpoint detection.

6.2. Previous Work

Several research groups have developed optical diagnostic sensors and algorithms for measuring 2-dimensional (2-D) or 3-dimensional (3-D) images of plasma, and measuring the spatial distribution of reactants in a plasma reactor. In this section, we briefly introduce their work and sensor designs.

Hareland and Buss [6.1] at Sandia National Laboratories have developed a sophisticated imaging instrument, consisting of a monochromator with bilaterally adjustable entrance and exit slits, video camera with video cassette recorder and image processor, as shown in Figure 6.1(a). The lateral image collected from the plasma etch chamber was transformed to 3-D spatial maps of emitting species of the Cl_2/He plasma by the Abel inversion (described in Chapter 3). A patterned 6" Si wafer with poly-Si lines was etched so as to investigate the relation between local etch rate and radial emission intensity profiles of excited atomic Cl. The results show that there is a direct correlation between the radial emission intensity of atomic Cl and the etch rate for polycrystalline silicon.

Pender *et al.* at the University of Michigan [6.2] used the Abel inversion to determine the radial optical emission profile in a parallel plate radio frequency system known as the GEC Reference Cell. The optical system imaged the parallel light collected from the emission of Ar discharges at a given height between the electrodes. This image was rotated by 90 degrees to project parallel light rays on to a photodiode array for various pressure

and power settings, as illustrated in Figure 6.1(a). This was repeated for a number of points across the electrode at various heights. Etch uniformity was also examined in this study. The results indicated that the non-uniformity of this discharge was directly related to the etch rates. Buie *et al.* [6.3][6.4] have conducted a similar experiment to construct the optical emission radial image at various pressure and power settings for 1024 points across the electrode, at a given height above the bottom powered electrode for the Ar line at 750.4nm.

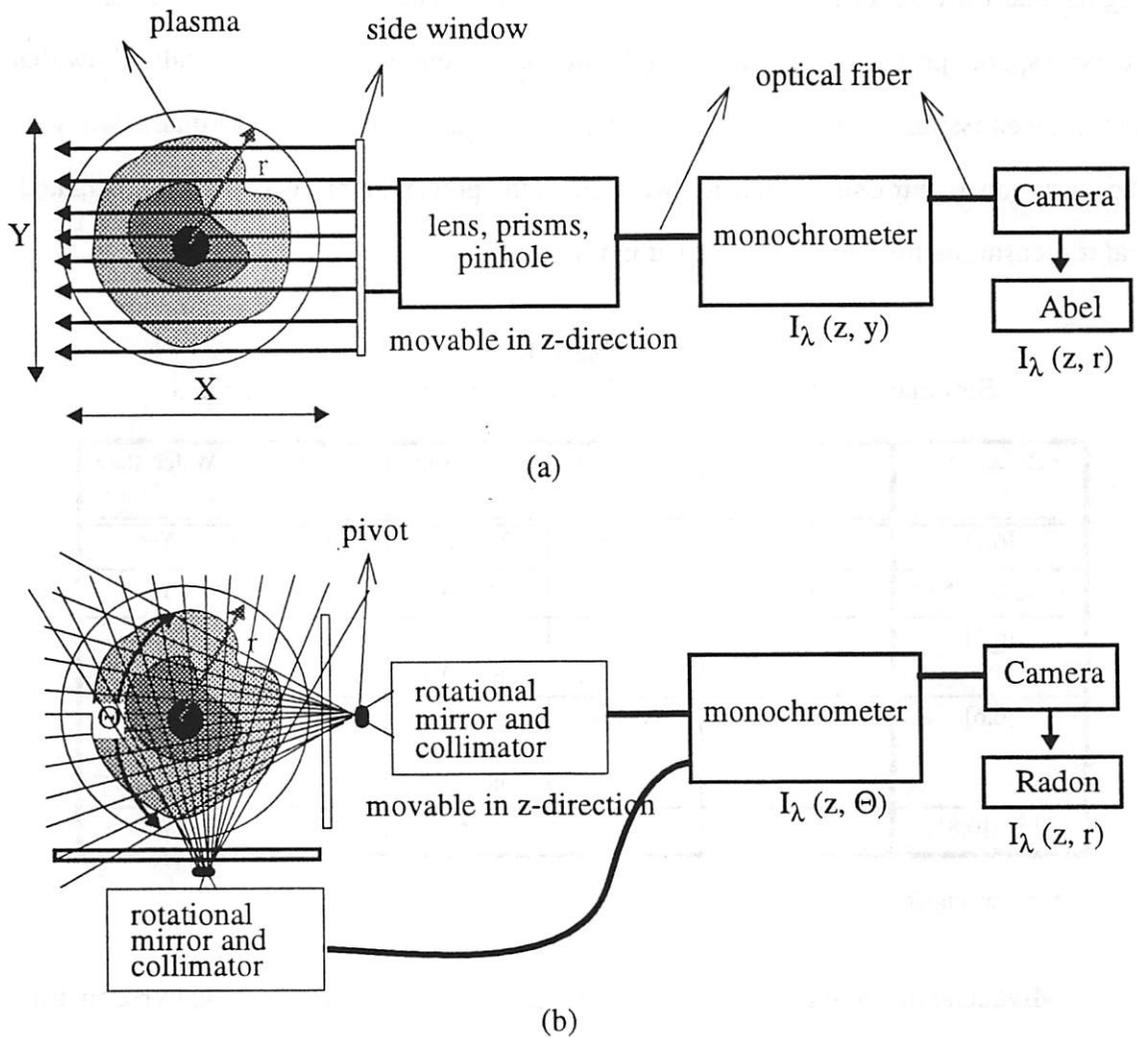


Figure 6.1 Schematic illustration showing the top view of the discharge chamber for (a) references [6.1][6.2][6.3][6.4][6.5]; (b) reference [6.6].

The plasma was monitored for the entire 30 minute etch. The results indicated that the plasma emissivity and the etch depth are related.

Beale *et al.* [6.5] at the University of Wisconsin-Madison also used the Abel inversion of line-integrated intensities of Ar emission to characterize the 2-D maps of emission in a planar, inductively coupled discharge (ICP) of argon. The collection optics were mounted on a movable stand outside a 6-inch diameter viewport on the side of the cylindrical chamber. The line of sight of the collection optics was oriented to image chords in planes perpendicular to the chamber axis. Stand positions were selected to include three heights, and for each of these, several radial positions. The contour plots of the emission intensity spatial profile for two different Ar pressures were obtained. The results show that maximum emission intensity in a location which is off-axis and close to the quartz window. The variation of emission intensity with different power levels was also investigated. Wafer measurements were not included in this study.

Table 6.1
Summary of the previous work related to spatially resolved OES

| Reference | Plasma | Inversion | Lambda(nm) | Published Year | Wafer-state correlation |
|-----------------|--|-----------|------------------------------------|----------------|-------------------------|
| [6.1] | Cl ₂ /He | Abel | 726 (Cl) | 1993 | Yes |
| [6.2][6.3][6.4] | Ar; CF ₄ /Ar/O ₂ | Abel | 750 (Ar) | 1993,95,96 | Yes |
| [6.5] | Ar | Abel | 912.5 (Ar), 867 (Ar) | 1994 | No |
| [6.6] | Ar | Radon | 419 (Ar), 434 (Ar), 396 (Ar) | 1991 | No |
| [6.7][6.8] | Ar | * | * | 1996,97 | Yes |

*: not available

Miyake *et al.* [6.6] have applied a more general tomographic method to reconstruct the optical emission profile of several species in a ring-shaped DC magnetron discharge. Instead of imaging the parallel light emission. Miyake's group collected the optical emis-

sion profiles from two ports at 90 degrees to each other by two identical scanning mirror systems. In each case, the direction of the slice of collected emission was pivoted about the point and controlled by the position of a mirror (along x or y, see Figure 6.1(b)) in order to obtain the radial profile. A series of 100 fan-beam scans in the x-y plane were conducted and the respective 2-D radial profiles were reconstructed by using the inverse Radon transform [6.6].

Shannon *et al.* at the University of Michigan [6.7][6.8] also examined a similar light collection system based on a rotating point sensor at the surface of the view window. The radial emissivity profile was reconstructed by a regularized inversion algorithm. The correlation between etch rate and radial emissivity profile was observed. A summary of their works is tabulated in Table 6.1.

All these previous works use the sensor systems providing either low spatial or temporal resolution. Moreover, only single wavelength experiments were presented. Here we report on a more complete scanning OES system with good spatial, temporal, and *spectral* resolution for the purpose of real-time monitoring of plasma etch processing.

6.3. Instrument Set-up at the Berkeley Microfabrication Laboratory

A Verity Instruments scanning mirror system with an Oriental Motor's RFK 5-phase stepper motor is mounted on the viewport of Lam 4400 plasma etcher in the Berkeley Microfabrication Laboratory. A set of line integral measurements for a selected spectral line is collected by scanning the mirror on the x-y plane (see Figure 6.2) with an increment of $\Delta\theta^\circ$ over the range of $-\theta^\circ \sim +\theta^\circ$, dependent on the required spatial resolution. The typical data acquisition rate is 20~30Hz. At this data acquisition rate, the plasma can be scanned in 1~3 seconds. Each scan cycle consists of line discharge intensity measurements of many different angular positions in the sweep. Fan-shaped regions of discharge emission are thus collected by this computer controlled scanning mirror system, which optimizes two important variables: scanning speed and scanning steps (that is, the number of stops for each scan

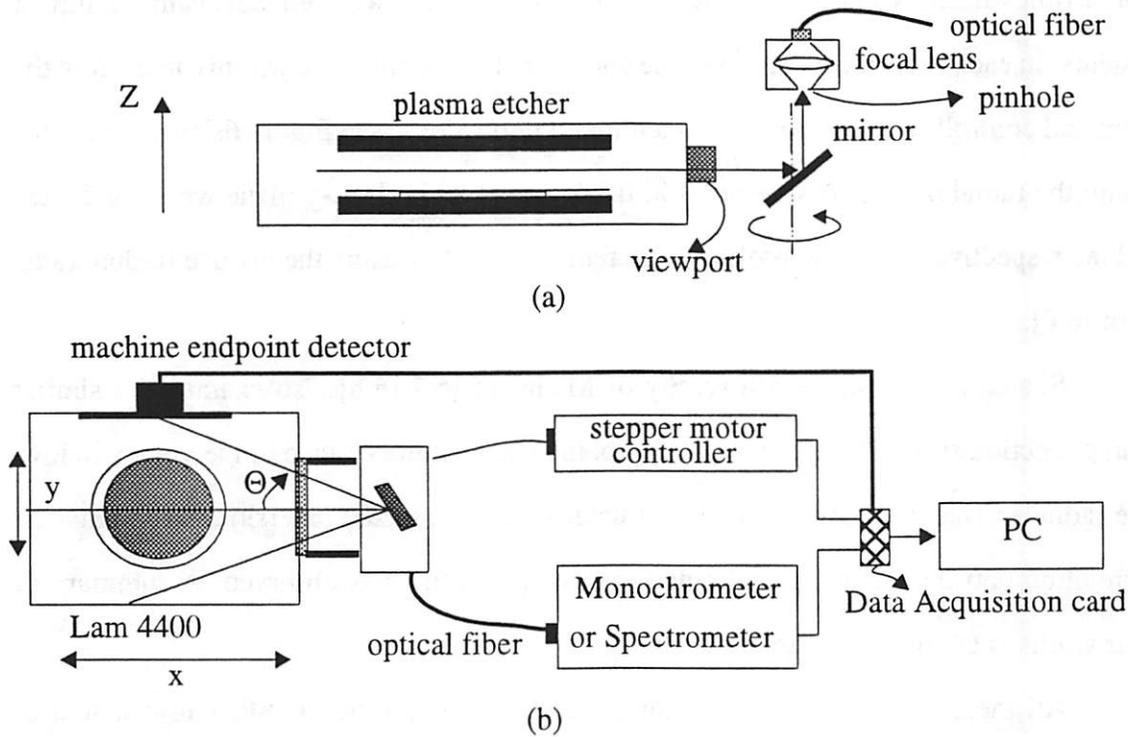


Figure 6.2 Schematic illustration of the scanning SROES set-up and discharge chamber (a) side view; (b) top view.

cycle), while the scanning angle is given. The angle of each scanning step of the stepper motor is controlled by the pulse train clock rate sent by a 166 MHz PentiumTM PC. The driver of the stepper motor is written using the National Instrument's LabViewTM software package.

A slice of the light emission signals is deflected 90° from x-y plane to z-axis (see Figure 6.2 and Figure 6.3) by a mirror, and collected by means of a series of UV-grade fused-silica lens and optical fiber. A 1mm pinhole is placed at the focal point of the focal lens in order to block out the surrounding background light. The collected light emission signals are then transferred to an SC Technology DES-310 monochromator, and thereafter dispersed by a 600 groove/mm grating system with a spectral resolution of about 1nm. The wavelength can be chosen with the help of a 3-digit LED display which is built into the monochromator. Light intensity at this wavelength is then detected by a photomultiplier

tube (PMT). The snapshots of the emission are triggered by the acknowledge signals of the stepper motor to ensure the data acquisition takes place only when the stepper motor is stepped. The data acquisition rate is therefore equal to the number of scanning steps per second.

A DC voltage between 0 and 10 volts is applied to adjust the gain of the PMT. The resulting analog output of the monochromator (i.e., the light emission counts at a selected wavelength) is sent into National Instrument's AT-MIO-64E-3 data acquisition card via a BNC cable. The lateral image of plasma is recorded for each scan. As a result, a high spatial resolution (e.g., 20 steps per scan) can be achieved while the temporal resolution (e.g., scan cycle time ~ 1 second) of plasma imaging is still acceptable. It is important to note that the viewing angular range of this scanning OES is determined by the geometry of the etching chamber. Figure 6.3 depicts the geometry of the etch chamber and the scanning OES system.

Even though our objective is to collect full spectral data (as reported later in this thesis), we initially collected single spectral lines. In this context, it is crucial to choose a

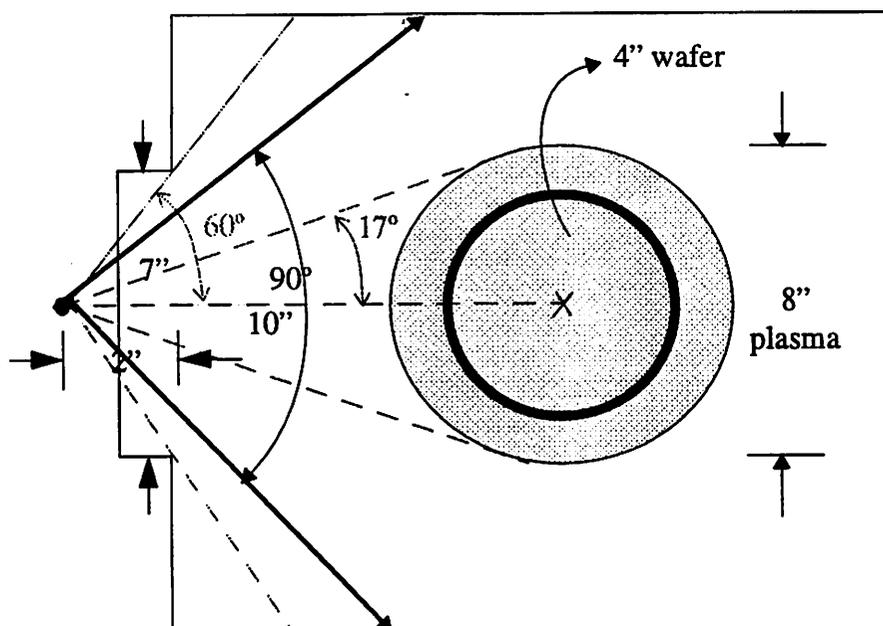


Figure 6.3 Geometrical illustration of the etch chamber (Lam 4400) and optical paths of the scanning mirror

proper monitored optical emission line which has high signal-to-noise (SN) ratio, and most importantly, relates well to the wafer-state parameter of interest such as the etch rate. In the case of polysilicon etch, we can choose the endpoint wavelength of 262 nm which corresponds to the Cl_2 concentration, or 405nm which is related to SiCl . We chose the 262nm in this experiment because of its high S/N ratio. The machine endpoint signals are also simultaneously acquired from the SECII port on the machine and sent to the acquisition card. Their optical emission traces are displayed together on the computer screen for comparison.

In an additional experiment, we have also collected full spectral-range data using an SC Technology SentryTM 400I spectrograph with an image intensifier. The scanning spatial images of 512 different wavelengths can be simultaneously collected throughout the etch process. Because the time for acquiring a spectrum is longer than that for the single-wavelength case, the typical acquisition rate is about 20 steps per cycle, resulting in a scan cycle time of 2.6 seconds. Table 6.2 lists the scanning rates and cycle times for single-wavelength and full-spectrum SROES, respectively.

Table 6.2
The optimized scanning SROES system settings

| Scanning SROES System | monochrometer | spectrometer |
|------------------------|---------------|--------------|
| Data acquisition rate | 25Hz | 10Hz |
| Scan steps per cycle) | 25 | 20 |
| Cycle time | 1 second | 2.6 second |
| viewing angle (degree) | ~35 | ~43 |

One challenge in designing this scanning SROES system is to optimize the scanning steps and acquisition rates while keeping the cycle time as short as possible, in order to maximize the temporal and spatial resolution. Figure 6.4 depicts the “resolution” trade-

offs. Case 1 illustrates an OES system with a high spatial resolution (e.g., small step angle) but low temporal resolution (e.g., long cycle time). Case 2 illustrates an OES with a high temporal resolution (e.g., high scanning rate, long cycle time) while the spatial resolution is low. Case 3 illustrates an ideal system in which both temporal and spatial resolutions are high. One obvious solution to achieve high temporal and spatial solutions is to increase the speeds of microprocessor/microcontroller, data acquisition routines and motor motion.

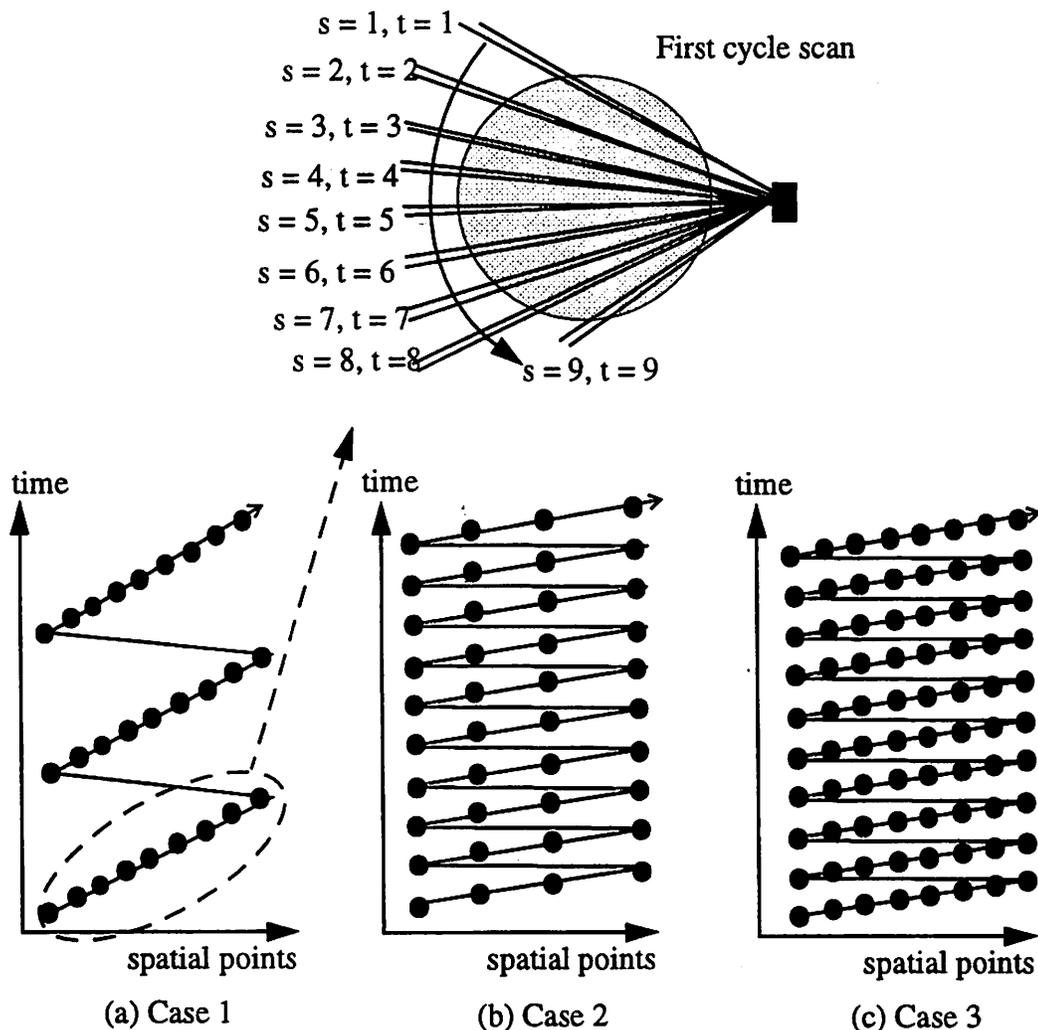


Figure 6.4 Schematic “resolution” plots for three different SROES systems with (a) high spatial resolution and low temporal resolution; (b) low spatial resolution and high temporal resolution; (c) high spatial resolution and high temporal resolution. (• denotes data point at time t and spatial point s)

6.4. Experimental Design Used to Collect Scanning OES Data

The experiment includes two phases. In the first phase, a simple experiment is conducted in order to examine the spatially-resolved endpoint detection capability using this scanning SROES system. The experiment includes seven 4" wafers which are covered with 30nm gate oxide and 400nm poly-Si films without patterns on the wafers. All wafers are processed under the same baseline recipe (listed in Chapter 5) on a Lam 4400 poly-Si etcher. The detector used in this experiment is a monochromator turned at 262nm.

Table 6.3
Machine settings for a 2^4 experimental design

| Input settings | Low | Center-point | High |
|--------------------------------|-----|--------------|------|
| RF Top (Watts) | 270 | 300 | 330 |
| RF Bottom (Watts) | 135 | 150 | 165 |
| Pressure (mtorr) | 11 | 12 | 13 |
| HBr/Cl ₂ flow ratio | 2 | 3 | 4 |

In the second phase, a two-level, four-factor, full factorial experiment with four center points is conducted on a Lam 9400 high-density plasma etcher with an etchant of Cl₂/HBr. The experimental variables are pressure, RF top electrode power, RF bottom electrode power, and HBr/Cl₂ gas flow ratio (see Table 6.2). The OES used in this phase is a SC Technology Sentry™ 400I full spectral-range spectrometer. A typical optical emission spectrum collected from Lam 9400 is shown in Figure 6.5, which is quite different from the spectrum collected from Lam 4400 (etchant is only Cl₂) as shown in Figure 5.1. The spatial profiles are collected at 512 wavelengths ranging from 200nm to 1024nm with a resolution of 2~3 nm/pixel. The data are collected throughout a 60-second timed main etch plus a 20-second timed over-etch under various etching conditions. All poly-Si film is removed completely after the etch process. Off-line oxide thickness measurements are per-

formed both pre- and post-etch on the Nanospec. Thirteen points are recorded for each wafer.

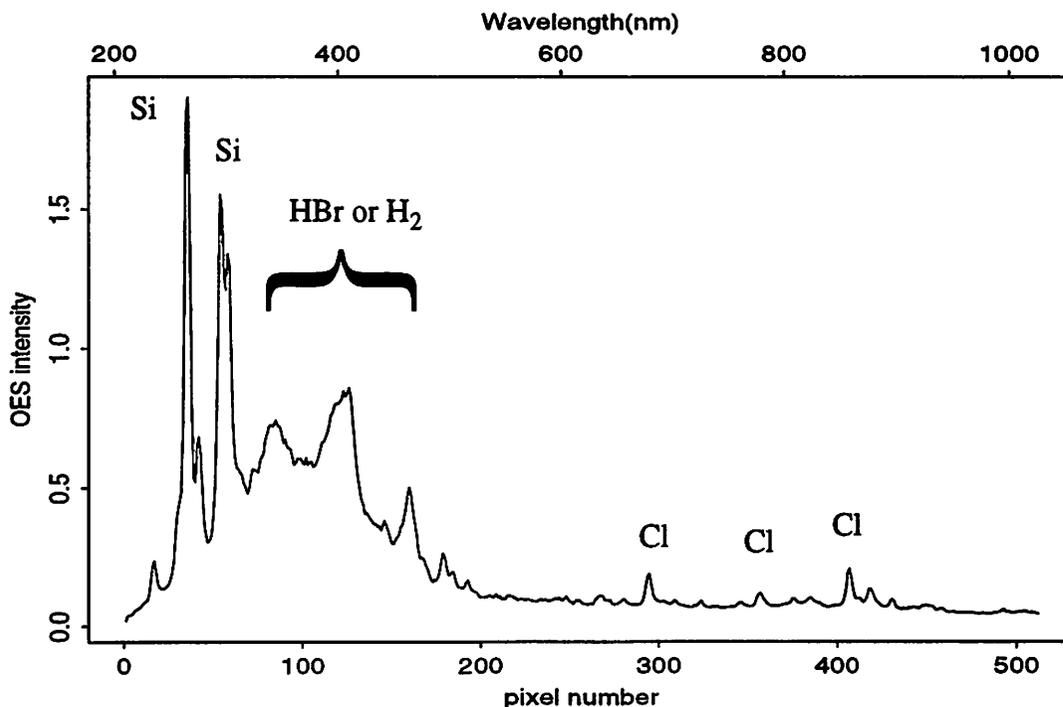


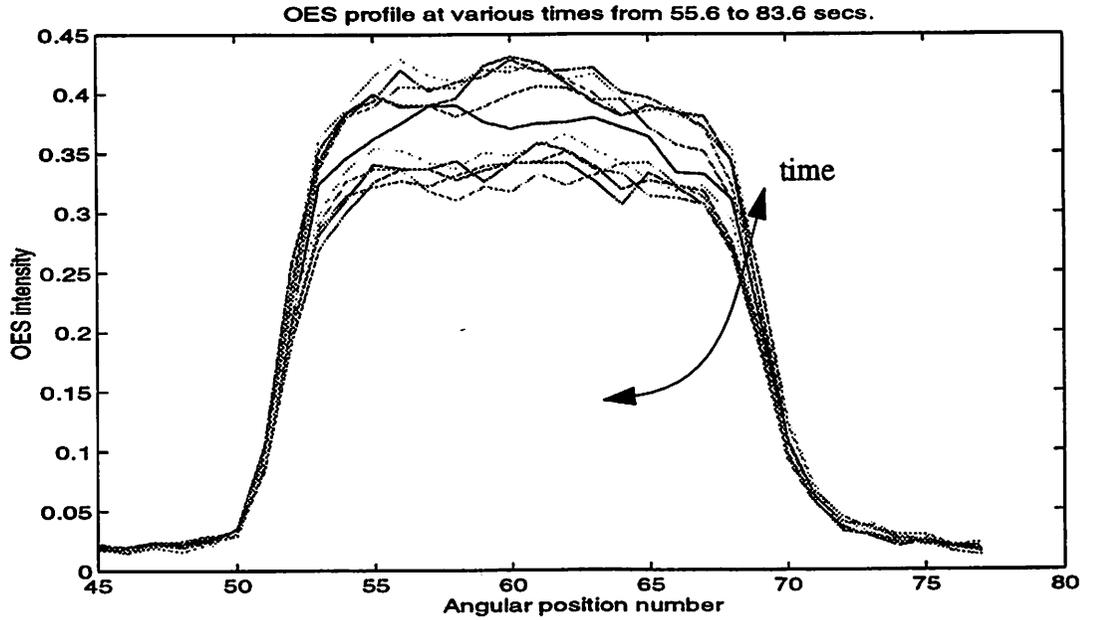
Figure 6.5 A typical spectrum collected from the Lam 9400 during main etching (etchant is Cl_2/HBr) with spectral lines labeled with their corresponding chemical species.

6.5. Data Analysis

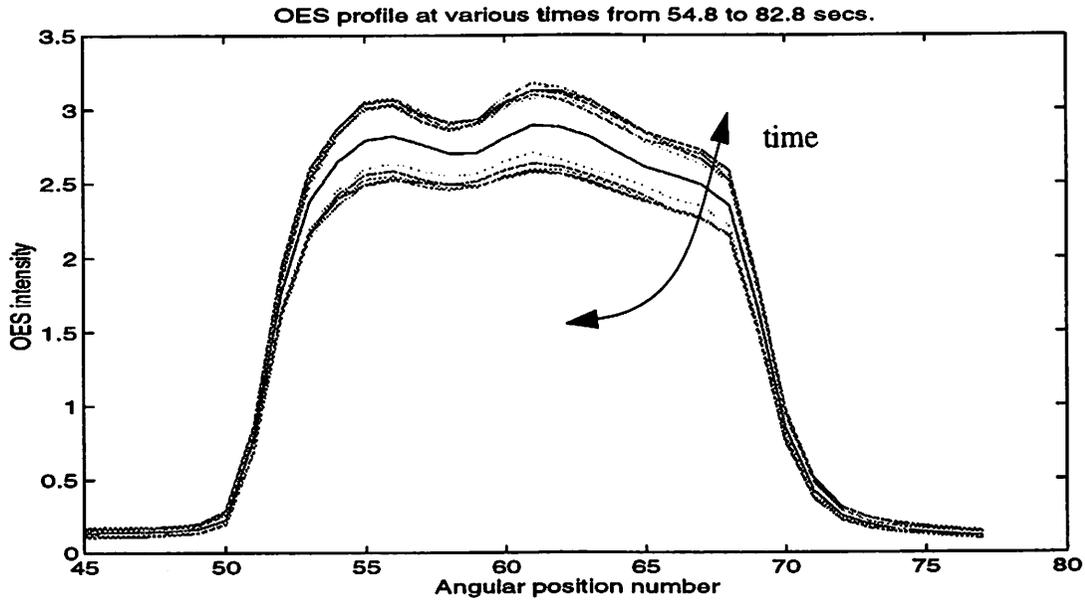
6.5.1. Phase 1: Single Wavelength

A typical two-dimensional spatial profile collected from this scanning SROES system with monochromator is shown in Figure 6.6. The most interesting feature of these spatial images is their temporal dependence. In order to distinguish the plasma-induced emission variability from the background noise, one needs a high S/N ratio of the emission signals at a selected wavelength. Therefore, the SROES is operated at 262nm (the strongest Cl_2 line) in this experiment. Figure 6.7 depicts a typical evolution plot of plasma spatial profile at the 262nm spectral line. One can clearly identify the endpoint from this plot. In

order to further eliminate the noise induced by the non-plasma factors, careful data smoothing, both along time axis as well as along Θ axis, is sometimes necessary. In those cases, a locally-weighted smoothing technique is used.



(a)



(b)

Figure 6.6 Typical plasma lateral profiles collected by scanning SROES with monochromator turned at (a) 405nm; (b) 262nm

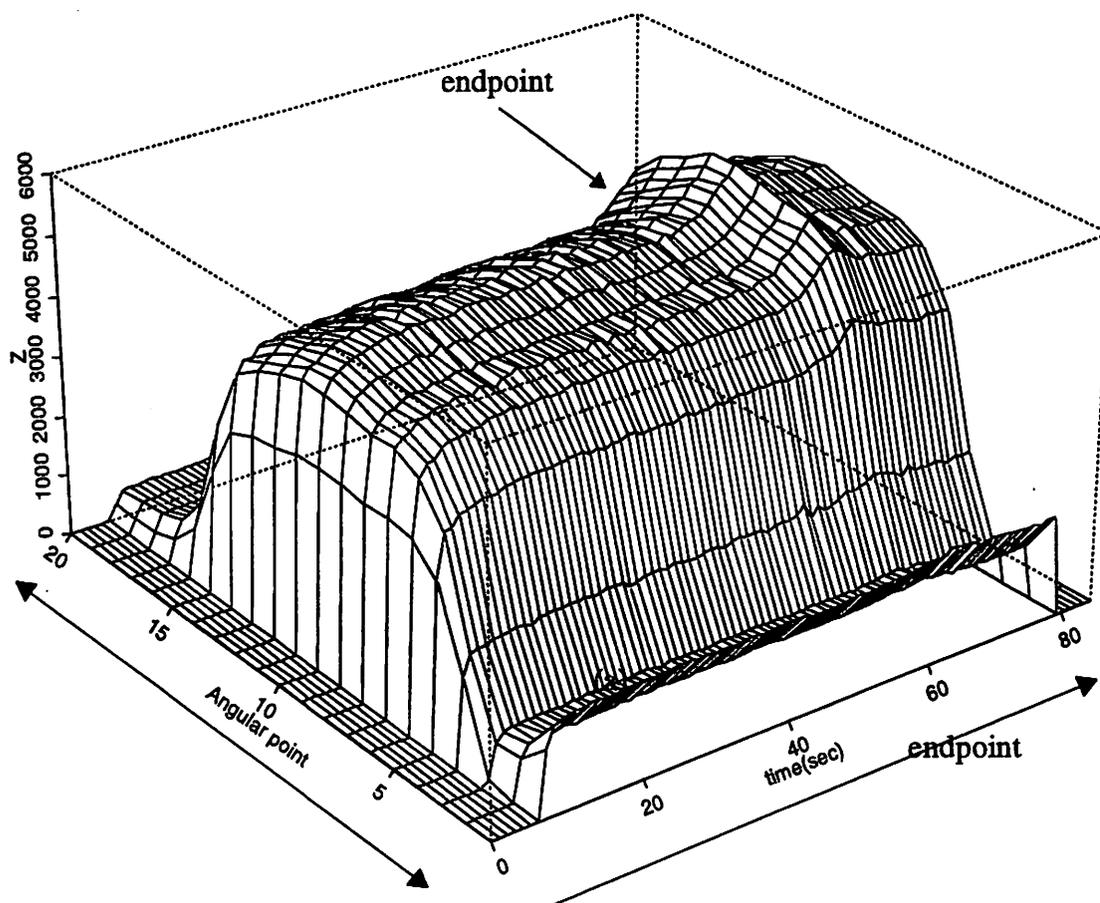


Figure 6.7 The temporal evolution of spatial plasma images (at 262nm) before smoothing.

The next challenge is to extract a spatially-resolved endpoint from this plot. Because the trajectory of plasma emission might evolve differently, depending on the angular point, an empirical approach is presented here to detect the spatial uniformity of endpoints: first, we take the ratios of the OE intensities of the adjacent scanning cycles, that is, we determine the value of

$$R_{i,j} = I_{i+1,j} / I_{i,j} \quad (6.1)$$

where $R_{i,j}$ is the intensity ratio, and $I_{i,j}$ is the smoothed OE intensity at cycle i and angular position j . In an uniform etch process, the ratio should be independent of the angular posi-

tion and is approximately a constant for each scanning cycle. By reconstructing a 3-D plot based on these ratios $R_{i,j}$ as a function of time and angular position (see Figure 6.8(a)), we can observe that the ratio profiles are mostly independent of angle during the main-etch and over-etch period, but become dependent on angle while the wafer is undergoing endpoint. In other words, this “spatially-resolved time dependence” means that the spatial profile of OE evolves at different rates while the etch process is at endpoint. This information can be further used to monitor the etch uniformity.

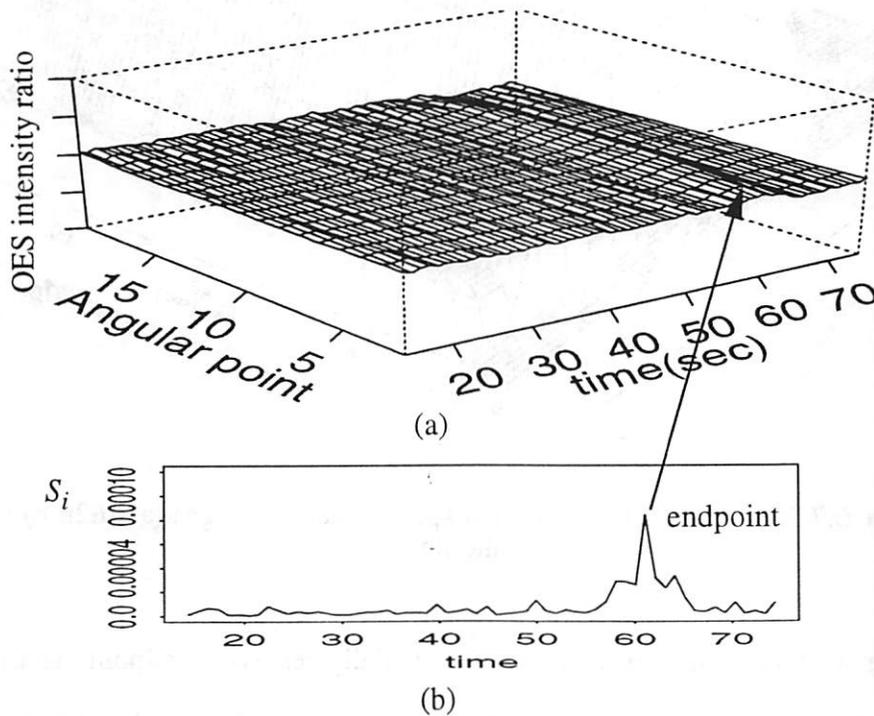


Figure 6.8 (a) OE intensity ratios v.s. time, and (b) its corresponding S_i plot

A statistical score S_i is defined by the standard deviations of $R_{i,j}$ dividing by the mean of $R_{i,j}$ at cycle i , that is,

$$S_i = \frac{\left(\left(\sum_{j=1}^n (R_{i,j} - \overline{R_{i,j}})^2 \right) / (n-1) \right)^{0.5}}{\left(\sum_{j=1}^j R_{i,j} \right) / n} \quad (6.2)$$

where n is the total number of angular points. This score is an indirect measure of the changes of the spatial profile at cycle i as shown in Figure 6.8(b), and can be used to predict the etch uniformity.

It is important to note that no direct information of etch uniformity is available in this experiment because the poly-Si film is removed completely during the etch process. Nonetheless, etch uniformity can be indirectly extracted by measuring the post-etch oxide thickness after the over-etch process, based on the assumption that etch rate non-uniformity mainly contributes to plasma non-uniformity rather than poly-Si film thickness non-uniformity. This data set can be used to verify the fault detection capability of this scanning SROES. Figure 6.9 shows the charts based on the spatial variance S_i and the corresponding over-etch uniformity. One can see that some wafers with large etch non-uniformity can be captured by the S_i chart.

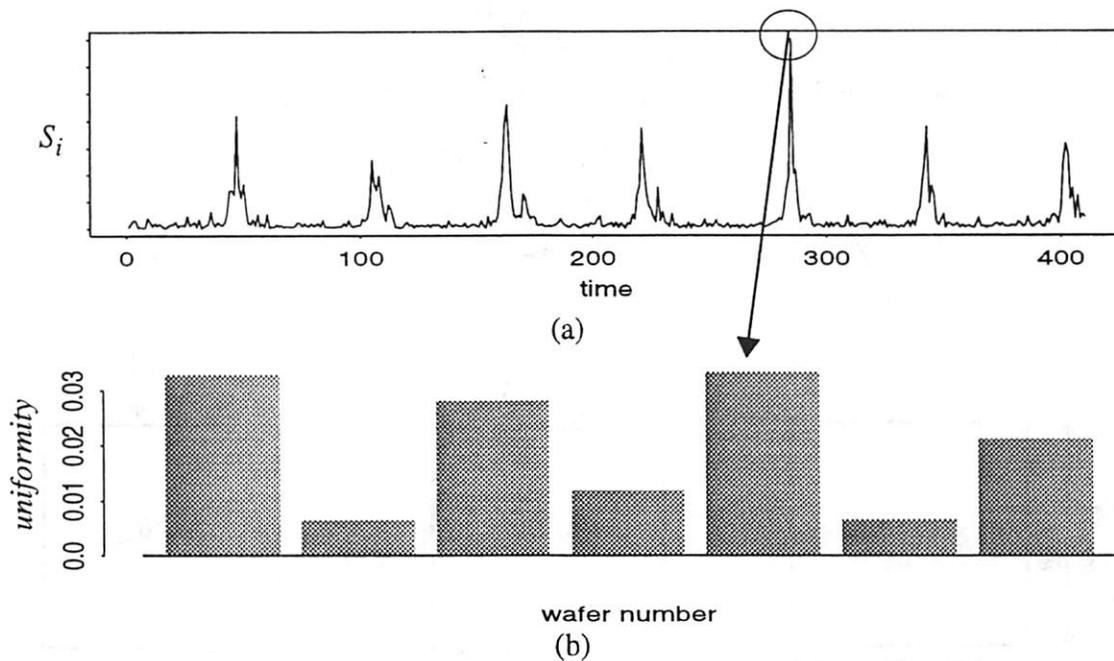


Figure 6.9 Control chart based on (a) statistical score S_i , and (b) over-etch uniformity

6.5.2. Phase 2: Multiple Wavelengths

While single wavelength spatial profiles, collected at high speed, conserve the purpose of process monitoring and control, a full spectral fingerprint of plasma emission, even with a relatively low scanning rate, can be a comprehensive means of process diagnosis. In this section, our scanning system incorporates full-range spectroscopy to acquire the scanning 3-D images of plasma at multiple wavelengths, so as to characterize the plasma etch

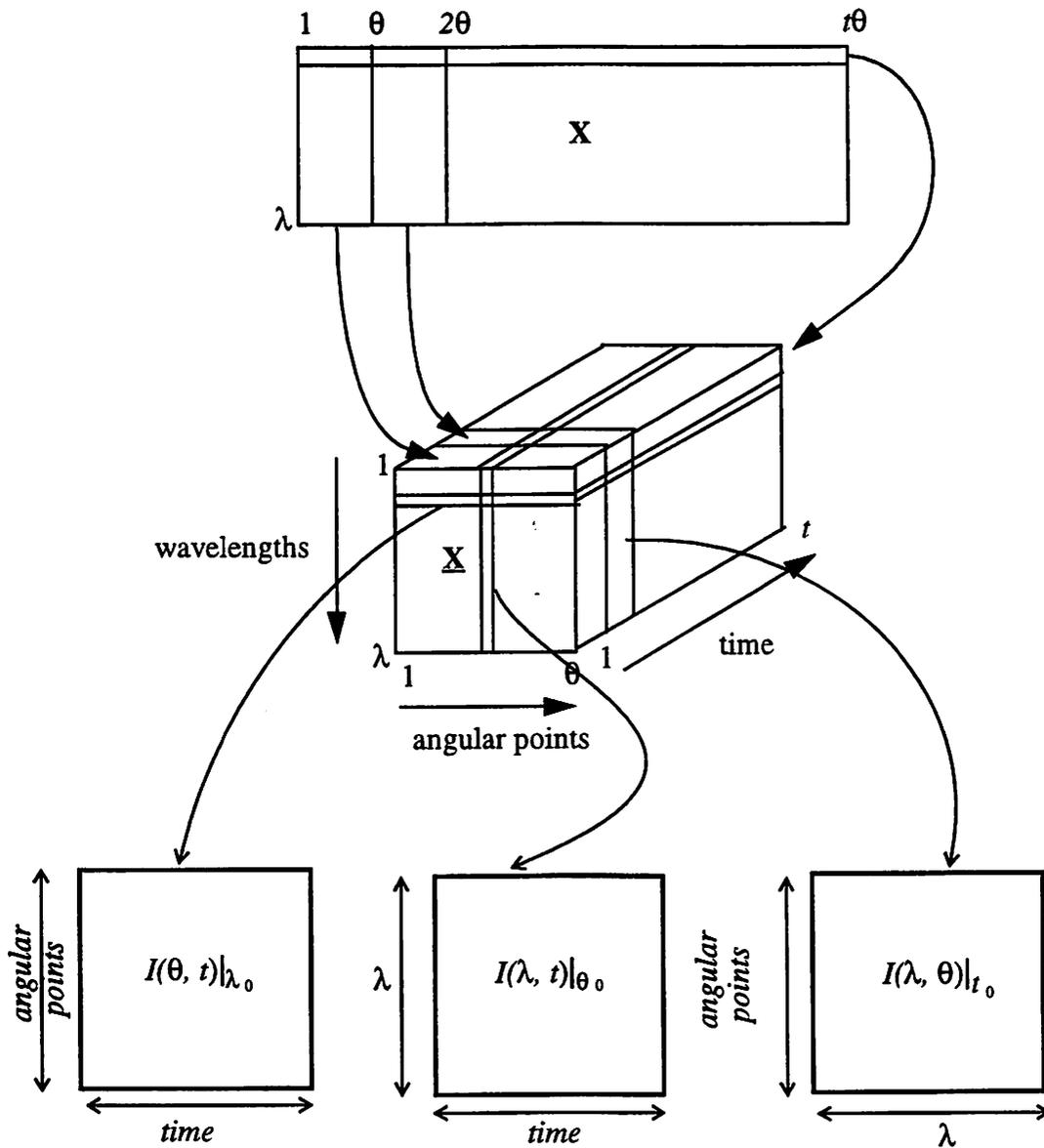


Figure 6.10 Arrangement and decomposition of a three-way array of scanning OES data set

process. The scanning rate is set at 20 steps per cycle, which corresponds to a cycle time of 2.6 seconds and a scanning angle that spans about 43 degrees. The data provided by this full-spectral scanning SROES require additional computation. In this section, we propose three different approaches to deal with these data sets.

6.5.2.1. Signature Analysis

To investigate the plasma images for one wafer, the measurement data is organized into a three-dimensional array $\underline{\mathbf{X}} (t \times \lambda \times \theta)$, in which θ is the angular point ranging from 1 to 20, λ is the wavelength, and t is the cycle number. The different wavelengths are organized along the vertical side, the measurement angular points are organized along the horizontal side, and their time evolution is revealed along the third dimension, as illustrated in Figure 6.10. Each horizontal slice through this array (see Figure 6.10) is a $\theta \times t$ data matrix representing the time trajectories of emission intensities for all the angular points at a single wavelength λ . Each vertical slice (see Figure 6.10) is a $\lambda \times t$ matrix representing the evolution of spectra at a specified angular point S . Each slice on a $\lambda \times \theta$ plane represents the spectral evolution along different angular points at a specified measurement cycle t .

Figure 6.11 shows typical $I(\theta, t)$ plots at 252nm (Si) and 838nm(Cl) lines, respectively. In this example, this specific wafer has undergone a long, three-step process: step 1 is the main etch for the poly-Si film, step 2 is the over-etch process for the underlying thin gate oxide, and step 3 is the continuous over-etch process for the underlying Si substrate. Si emission intensity appears to be higher during the Si etch steps than that during the oxide over-etch. On the contrary, Cl intensity is lower during the Si etch step compared with that during the oxide over-etch process. Figure 6.12 shows a typical $I(\lambda, t)$ plot at $\theta = 10$ (central angular point), and Figure 6.13 shows a typical $I(\lambda, \theta)$ plot during the 10th scanning cycle (midway through the main etch process).

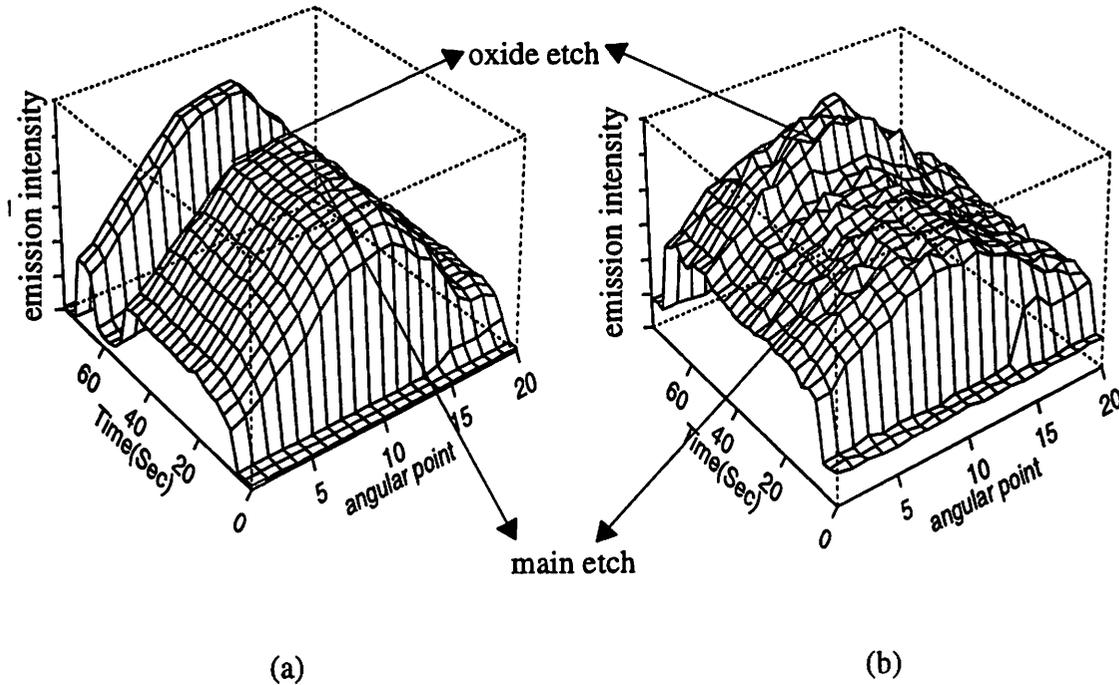


Figure 6.11 Typical $I(\theta, t)$ trajectory at (a) 252nm; (b) 838nm.

Figure 6.14 shows these emission signatures at different top-electrode power levels. It reveals that power has a stronger effect on the optical emission intensity levels, rather than that on the spatial or temporal profiles. Similarly, these emission profiles show little dependence on the bottom-electrode power, pressure, and HBr/Cl_2 gas flow ratios. In the future, with higher scanning speed and spectral resolution, these plots may be used as the signatures of the plasma chamber conditions. One possible application of the 3-D emission signature database is that it can be used to characterize and monitor the plasma and machine conditions with “inert” plasma without processing expensive test wafers.

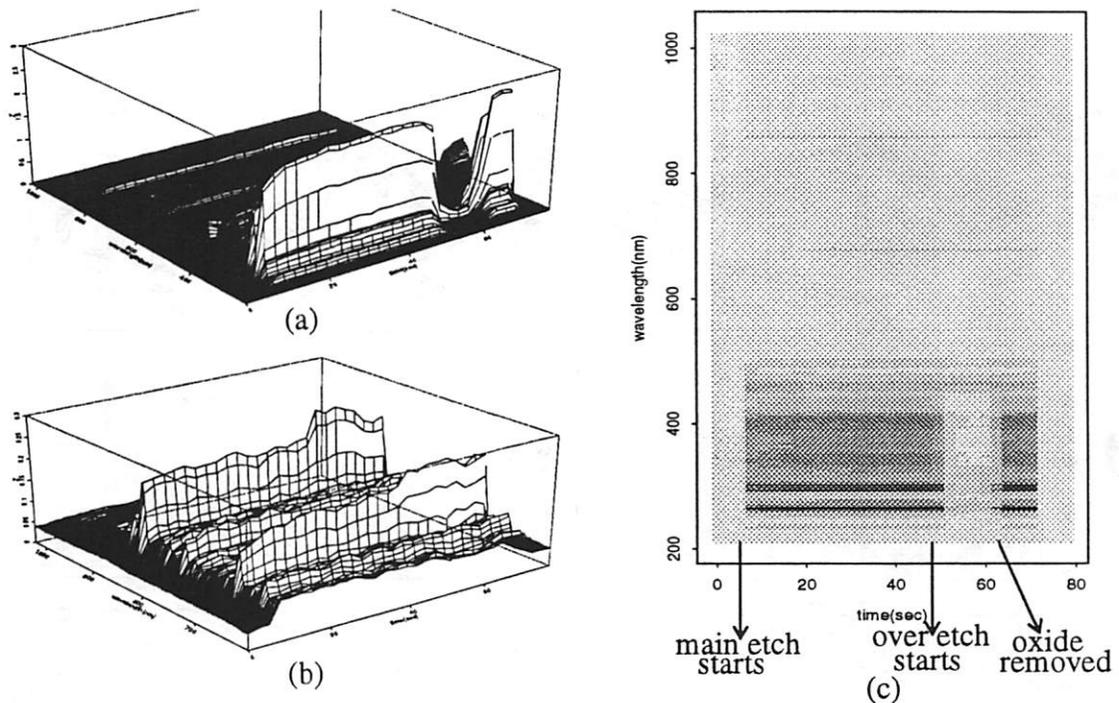


Figure 6.12 Typical $I(\lambda, t)$ plots at the central angular point (a) full spectral range plot, (b) 500-1000nm spectral range plot, (c) full spectral range image plot (NOTE: for image plot, black marks maximum intensity and white marks minimum intensity)

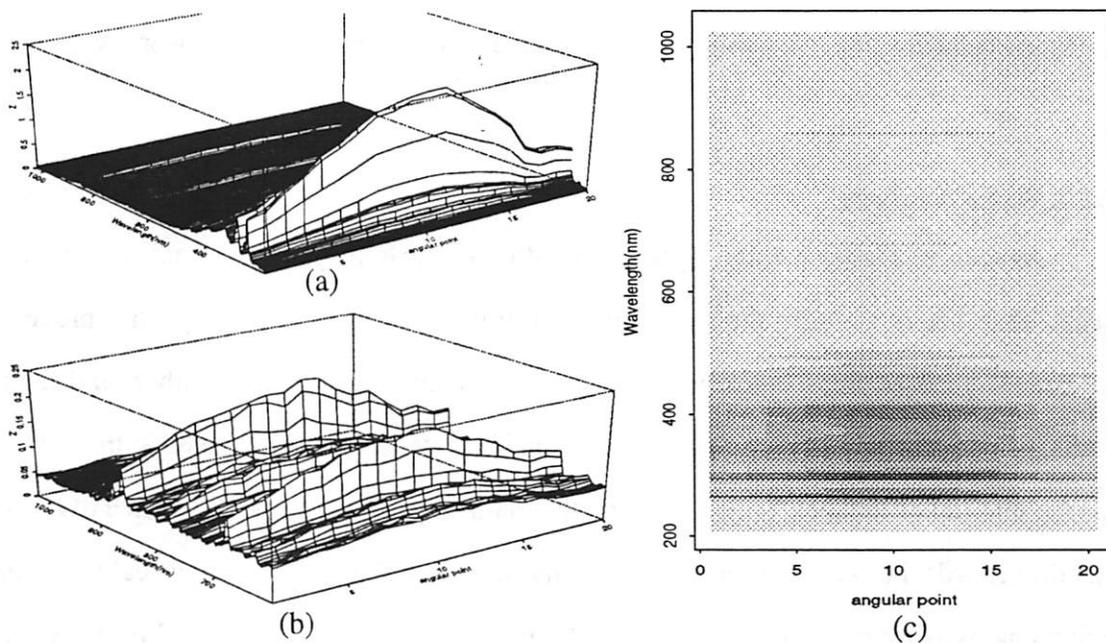


Figure 6.13 Typical $I(\lambda, \theta)$ plots at the 10th scanning cycle (a) full spectral range plot, (b) 500-1000nm spectral range plot, (c) full spectral range image plot

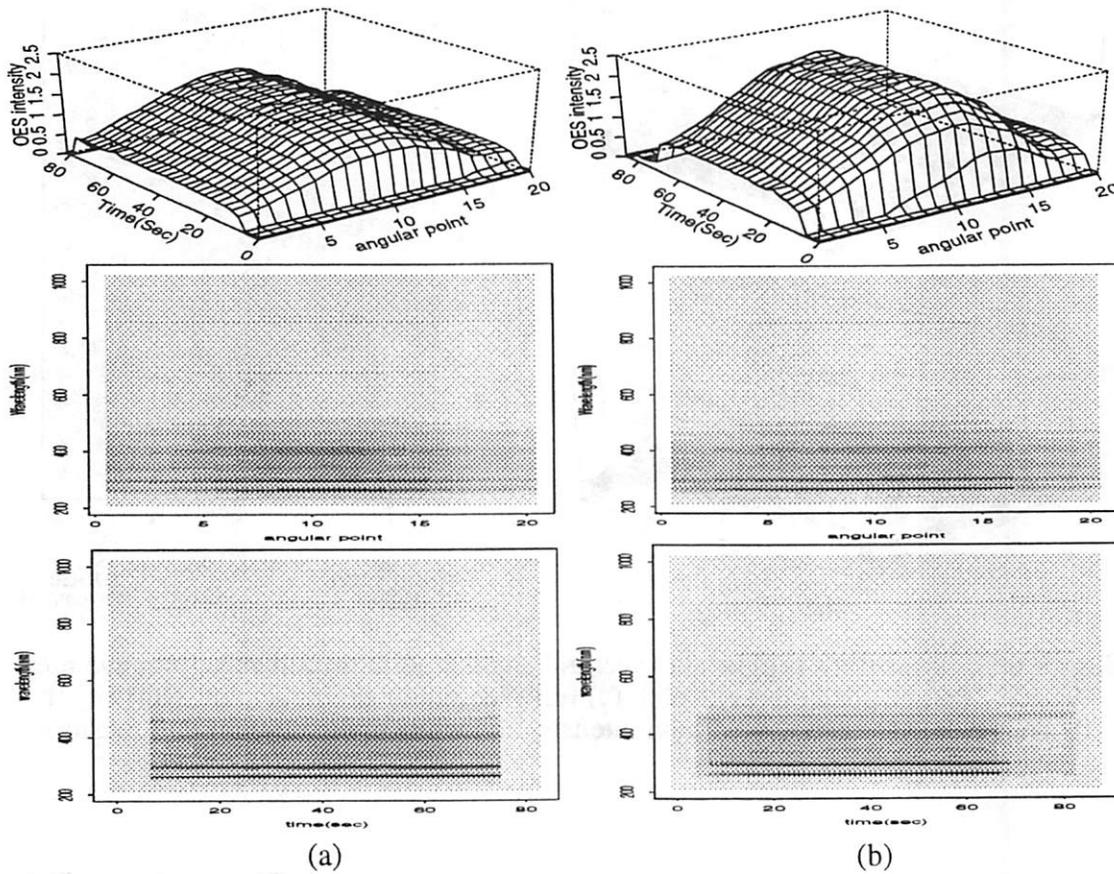


Figure 6.14 $I(\lambda, \theta)$, $I(\lambda, t)$, and $I(\theta, t)$ plots at different top electrode power levels (a) 270 W, (b) 330 W.

6.5.2.2. Ratio Analysis

As seen in Figure 6.4, the light collected by the scanning OES system outside the process chamber is actually the total emission along a linear path through the process chamber. The linear paths correspond to chords through the circular chamber, and their lengths differ according to their angular position. The total emission received at the detector is the integral of the local emission along that path length. Due to the difference of these collection chords, bell-shaped curves are usually observed. Since this geometrical factor in the Si signal is also present in the Cl signal at the same acquisition, the spatial profiles can be corrected by normalizing the Si emission intensity by that of Cl, in order to visualize the true spatial profiles. This is seen in Figure 6.15, in which we take the ratio of the normal-

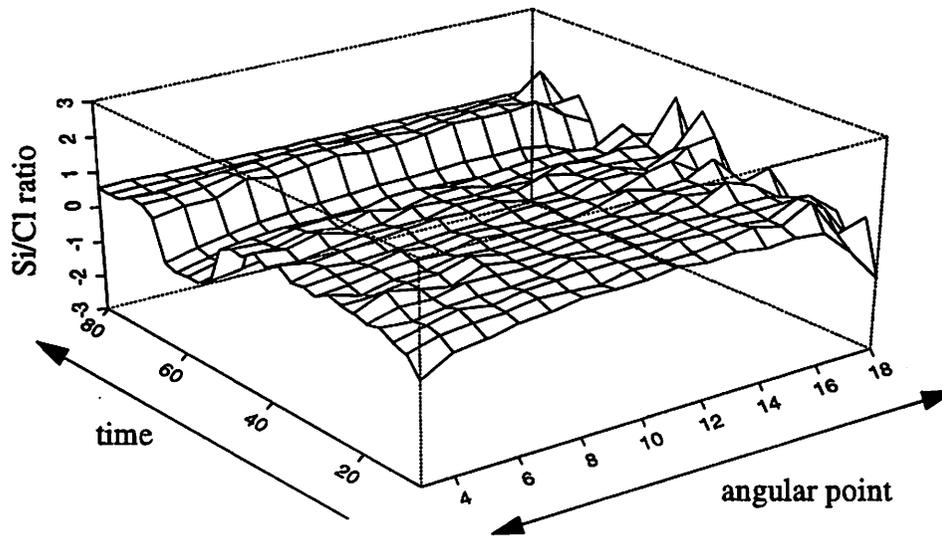


Figure 6.15 A typical ratio plot of normalized spatial-temporal profiles of $I(252\text{nm})/I(838\text{nm})$

ized ¹spatial-temporal profiles of two different wavelengths. Note that the temporal information is still retained in this plot, while the spatial profile is normalized. This information can be used to indicate the spatial uniformity of the plasma emission, allowing an indirect observation of spatial etch variability.

6.5.2.3. Principal Component Analysis

A statistical data compression technique, Principal Component Analysis (PCA), can be also implemented to compress the spectral information into few principal components. In this way, the plots depicted in Figure 6.12 and Figure 6.13 can be reduced to those shown in Figure 6.16 and Figure 6.17, respectively. In addition, by extending the PCA compression shown on Figure 6.16 to all angular points, we can construct the spatial-temporal profiles for the first four principal components (PCs), as depicted in Figure 6.18.

In comparison to the similar profiles sampled at the selected wavelength (see Figure 6.11), the profiles based on the compressed PCs are more difficult to interpret phys-

1. the spatial-temporal profiles are normalized for a total range for 0 to 1 before taking the ratio.

ically. However, there are several advantages in using the PCA approach. First, the spectral information is statistically filtered to only a few components so that no chemical identification is necessary. Secondly, because the geometrical factor associated with the spatial profiles is mostly included in the lower order PCs, as described in Section 6.5.2.2, higher order PCs often contain more information of plasma emission disturbance than that of lower order PCs (however, in this case, the PCs higher than 4th order just contain white noise). For instance, the 2nd PC shown on Figure 6.18, reveals the etch transitions at different angular points without the typical bell-shaped spatial profiles. Additionally, the 3rd and 4th PCs shown on Figure 6.18 reveal the emission disturbances during etch transitions. By enhancing the scanning speed, this PCA approach can be incorporated with the framework proposed in Section 6.5.1 in order to extract the spatial etch uniformity. Thirdly, PCA is essentially a data smoother which can automatically move the data noise to the higher order PCs, while smoothing the profiles of the 1st and 2nd PCs.

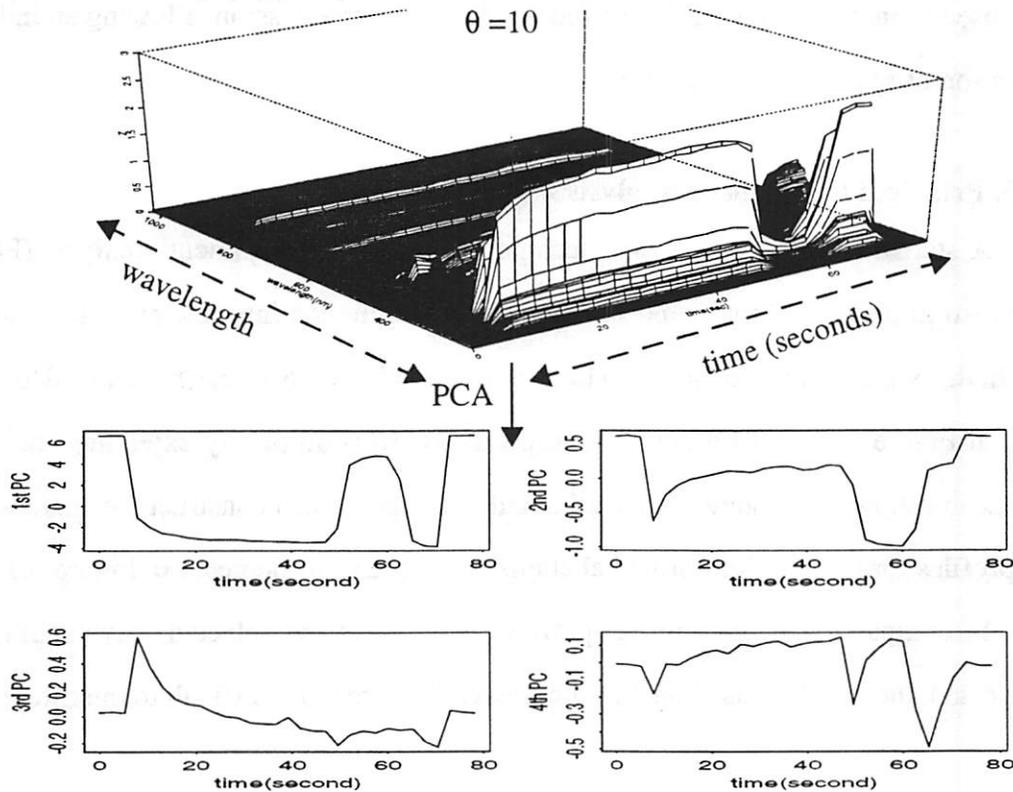


Figure 6.16 Temporal profiles at the four principal components after PCA compression (the first four PCs explain 94% of the variability of optical emission)

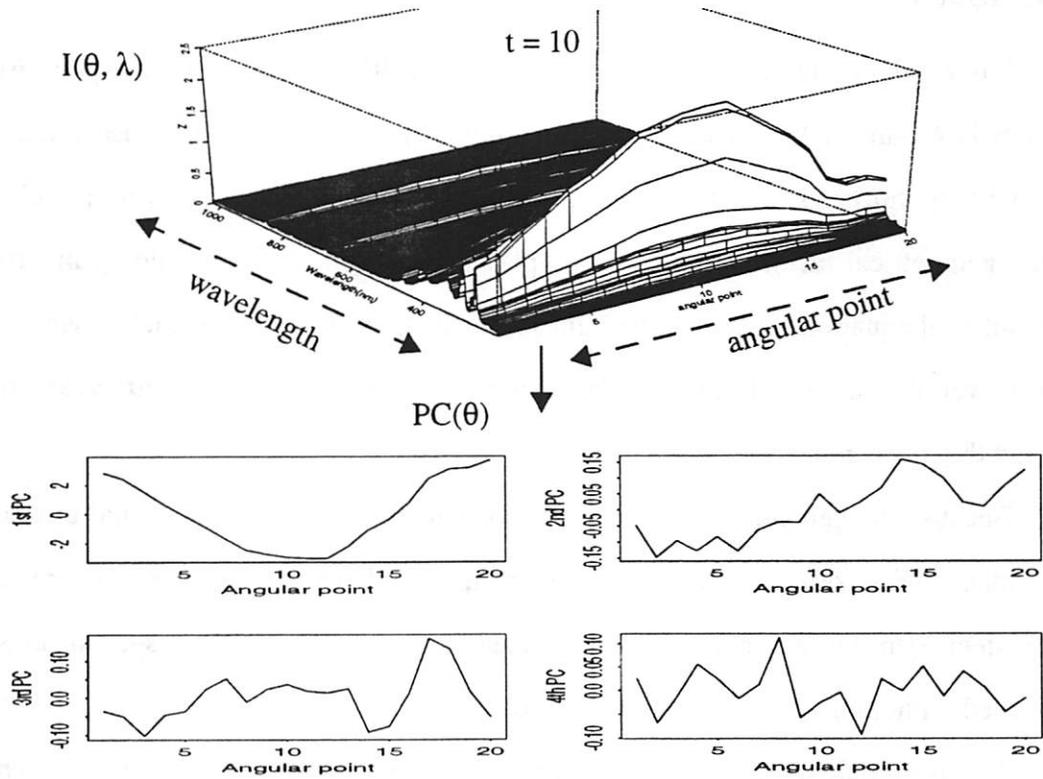


Figure 6.17 Spatial profiles at the four principal components after PCA compression (the first four PCs explain 98% of the variability of optical emission)

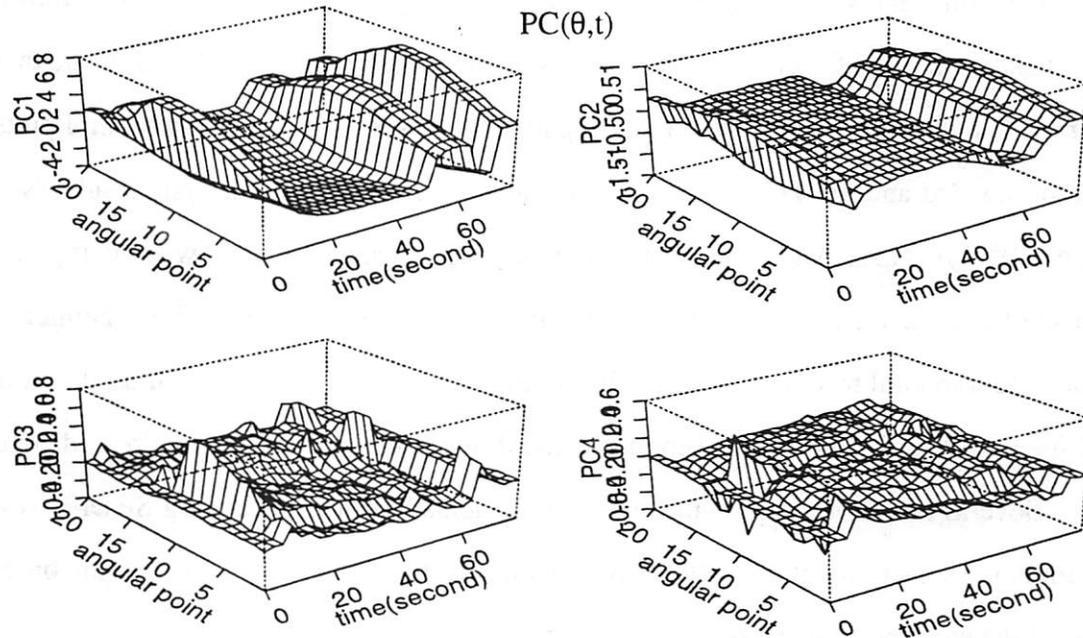


Figure 6.18 $PC(\theta, t)$ plots of the spatial profiles of the first four principal components

6.6. Discussion

It is worth noting that the Abel inversion [6.9][6.10] is not used in this study. Abel inversion is a pure mathematical technique applied in the study of extended radiation source with cylindrical symmetry. It was proposed [6.1][6.2][6.3][6.4][6.5][6.7] to eliminate the geometrical factor of the collected plasma emission spatial profile by inverting Abel's integral equation, assuming the light source is optically thin and axially symmetrical. However, the success of this technique is heavily relied on many geometrical assumptions and the low signal noise [6.9][6.10].

Because the geometrical factor of the collected plasma emission spatial profile is independent of its corresponding measurement time and wavelength, the temporal and spectral dependence of the spatial image collected by our scanning SROES system can still be revealed without utilizing the Abel inversion.

The post-etch remaining oxide thicknesses are measured for all wafers. Their within-wafer variations are also determined and plotted in Figure 6.19(a). Note that the oxide layers on some wafers are etched completely during the etch process, and therefore no oxide layer remains for these wafers. The results from the rest of the wafers indicate that wafer # 4 shows large etch rate variation. Figure 6.19(b) shows the spatial-temporal plots based on the 3rd and 4th PCs, as well as the plot based on its ratio analysis results (see Figure 6.19 (c)). These patterns do not exhibit any significant difference when compared to those of normal wafers. This might be due to the fact that the spectral (i.e., chemical), spatial, and temporal resolution of the full spectrum scanning SROES system used in this experiment is insufficient to reveal enough etch information, which will be addressed next.

Several issues are important for the implementation of the scanning SROES system. First, the sensitivity of the SROES system must be higher than the noise level in order to detect the etch-induced emission variability.

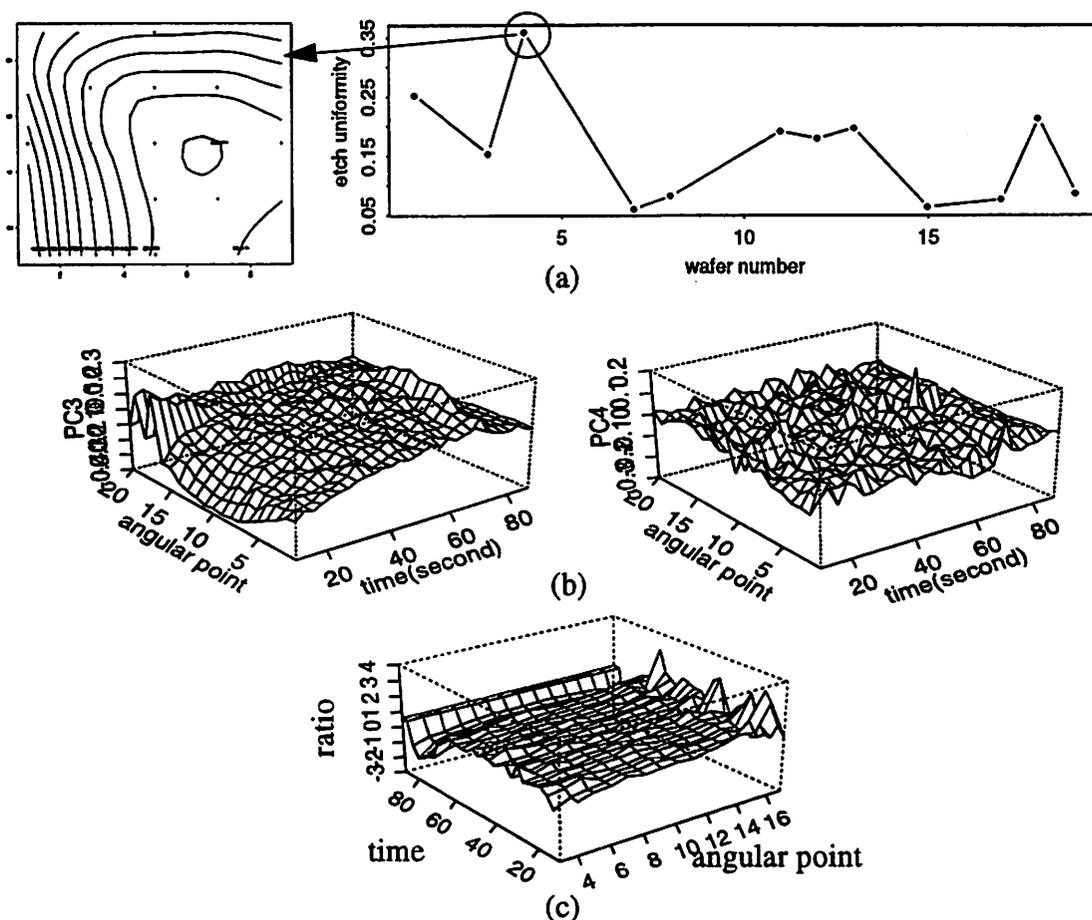


Figure 6.19 (a) oxide etch rate uniformity of wafer #4 and its contour plot, (b) its corresponding spatial-temporal plot based on 3rd and 4th PCs, and (c) its corresponding ratio plot of wavelength 252nm and 838nm.

Secondly, the scanning SROES is only able to detect the plasma-induced etch variability, but not the etch variability due to other factors such as wafer temperature, lithography non-uniformity, film defects and polymer deposition.

Thirdly, the scanning rate of this system has to be very high in order to remove the confounding effect between temporal resolution and spatial resolution, as described in Section 6.3. Since the typical endpoint transition lasts 1~3 seconds, a fast scanning SROES with short scan cycle time is required to detect the spatially-resolved endpoints.

Lastly, high spectral-resolution spectroscopy is important for detecting the minor species such as etch by-products. However, high resolution spectroscopy requires a detector with many pixels (e.g., 1024 or 2048 pixels). The acquisition time associated with this

high-resolution system is usually longer, thereby degrading the scanning rates and hence increasing the scanning cycles. Thus, optimizing the temporal, spatial, and spectral resolutions of the scanning SROES system is crucial for a potential application. For the purpose of process control and etch uniformity monitoring, a high scanning rate system with a monochromator is acceptable, as long as the key spectral line is determined *apriori*. For the purpose of process diagnosis, full spectral-range spectroscopy with high spectral resolution is necessary in order to investigate the emission signatures from many different chemical species simultaneously, despite the fact that the scanning rate of this system may be significantly lower than the monochromator case. Presently, an integrated system with both high scanning rate and high spectral resolution is under construction.

6.7. Summary

In this chapter we have demonstrated a novel approach to monitor etch spatial uniformity. Much effort has been devoted to build a scanning SROES system, which enables us to monitor the plasma variability with high spatial and temporal resolutions. We first attempt to use this system with a monochromator as a “spatially-resolved endpoint detector”. The results show that the time dependence of the spatial profiles is indeed more important than the profiles themselves. This framework has also been extended to a scanning SROES with a full spectrometer. The obvious advantage in using the spectrometer is that we can monitor a large number of different chemicals, allowing us to characterize the plasma even at the cost of losing the temporal resolution (because of the longer scanning cycle time). We have proposed three different approaches to analyze these data sets. It appears that none of these approaches is able to relate these scanning profiles to wafer states. The PCA is a more statistically robust approach compared to other empirical approaches, but may be too difficult to interpret practically, and may also need additional computational resources.

In order to maintain a high acquisition rate while still having high spectral resolution, an improved scanning OES system with improved software and better hardware design is being built. This new system can acquire signals with a potential maximum acquisition rate of 100Hz with an adjustable spectral range. The signals collected by this improved system can be used as fingerprints of the plasma emission conditions. One future direction is to analyze these fingerprints with the help of the signal processing algorithm such as Fourier Transformation [6.11] to extract the etch-induced emission disturbance. Moreover, this system possibly can be used to investigate the Cl₂ plasma density spatial profiles using Ar as an actinometer [6.12]. That is, the spatial profile of the ratios of the selected Cl and Ar lines can be used to indicate the spatial distribution of the plasma density *in-situ* and in real-time. The result also can be confirmed and further investigated by the traditional Langmuir probe [2.8].

References for Chapter 6

- [6.1] W. A. Hareland and R. J. Buss, "Optical Diagnostic Instrument for Monitoring Etch Uniformity During Plasma Etching of Polysilicon in Chlorine-Helium Plasma", Sandia Report SAND93-0455, 1993
- [6.2] J. Pender, et. al., "Radial Optical Emission Profiles of Radio Frequency Glow Discharges" *J. Appl. Phys*, vol. 74, no. 5, Sep. 1993, pp. 3590-3592.
- [6.3] M. J. Buie, et. al., "In Situ Diagnostic for Etch Uniformity", *J. Vac. Sci. Technol. A*. vol. 13, no. 4, Jul/Aug, 1995, pp. 1930-1934.
- [6.4] M. J. Buie, et. al., "In Situ Sensor of Spatially Resolved Optical Emission", *IEEE Transactions on Plasma Science*, vol. 24, no. 1, Feb. 1996, pp. 111-112.
- [6.5] Beale D. F., et al., "Spatially Resolved Optical Emission for Characterization of a Planar Radio Frequency Inductively Coupled Discharge", *Vac. Sci. Technol. A*. vol. 12, no. 5, Sep/Oct, 1994, pp. 2775-2779.
- [6.6] S. Miyake, et. al., "Diagnostics of Direct-current-magnetron Discharges by the Emission-Selected Computer-Tomography Technique", *Vac. Sci. Technol. A*. vol. 10, no. 4, Jul/Aug, 1992, pp. 1135-1139.
- [6.7] S. C. Shannon, et. al., "Plasma Tomography Systems for Industrial Plasma Tools", *Proceedings in the International Conference of Plasma Science*, 1996, p.282.
- [6.8] S. C. Shannon, et. al., "An Optical RIE Process Uniformity Control Sensor", *Proceedings in the International Conference of Plasma Science*, 1997, p.233.
- [6.9] C. Fleurier and J. Chapelle, "Inversion of Abel's Integral Equation - Application to Plasma Spectroscopy", *Computer Physics Communication*, Vol. 7, 1974, pp. 200-206.
- [6.10] M. J. Buie, et al., "Abel's Inversion Applied to Experimental Spectroscopic Data With Off-axis Peaks", *J. Quant. Spectrosc. Radiat. Transfer.*, Vol. 55, No.2, 1996, pp.231-243.
- [6.11] A. Oppenheim and A. Willsky, *Signals and System*, 2nd edition, Prentice Hall, 1997 New Jersey.
- [6.12] J. Coburn and M. Chen, *J. Vac. Sci. Technol.*, vol. 18, 1981, p.683.

CHAPTER 7

Conclusion and Future Work

7.1. Thesis Summary

This thesis presents a methodology for estimating etching performances using the information provided by real-time OES sensors. The sensor readings were first collected throughout the oxide plasma etching processes and then used with existing statistical techniques to model etch rate, etch rate spatial uniformity, aspect-ratio dependent etching. Various modeling techniques such as Principle Component Analysis (PCA), and Partial Least Squares (PLS) have been employed to relate the various OES signatures to etching performance. The results show that 87% of oxide etch rate variation and more than 95% of the variation in their within-wafer uniformity can be explained by these models, although the OES signals can only explain 65% of the variation in aspect-ratio dependent etch (ARDE).

We then examined the optical emission spectra collected from a commercial high-density plasma metal etcher under various machine input settings. Similarly, these optical emission spectra collected during aluminum etch and over etch (i.e., oxide etch) steps can

be employed in combination with several statistical and physical techniques to model the wafer states such as over-etch rate and aluminum critical dimension (CD), and machine states such as TCP coil power and chamber pressure. This study is based on the optical emission spectroscopy (OES) system installed on a Lam 9600 TCP high-density plasma (HDP) metal etcher at Texas Instruments (TI).

Our study shows that OES is only superior in capturing the variations of few machine variables such as TCP power. Specifically, our results indicate that more than 80% of the wafer-to-wafer plasma light-emission variation is explained by TCP power alone, while chamber pressure and bias power can only explain a small portion of variation. Additionally, we observe that OES is more sensitive to the variations in the etching process driven by ion-bombardment than to the variations driven by the inhibitor-driven mechanism. In other words, the implementation of OES sensor for monitoring the etch variability is dependent on the etch mechanisms. Our results suggest that OES sensors are better able to capture oxide etch rate variation (more than 94% of wafer-to-wafer variation is explained), while only less than 44% of CD variation is modeled by OES under the same experimental conditions.

Although the OES sensor is capable of continuously collecting spectra in real-time, this study only looked at a single spectral snapshot per wafer. Hence, we develop methods for characterizing the time behavior of the plasma. These methods allow us to incorporate real-time data into our models, thus increasing the accuracy of our predictions. Most of this effort has been devoted to decompose the variations of signals acquired from a spatially-resolved optical emission spectroscopy into several components, which provide valuable information of etch spatial uniformity. We have proposed an empirical approach to extract the variability of OES signals. It is conducted by the temporal and spatial signature analysis of a selected wavelength. The results show

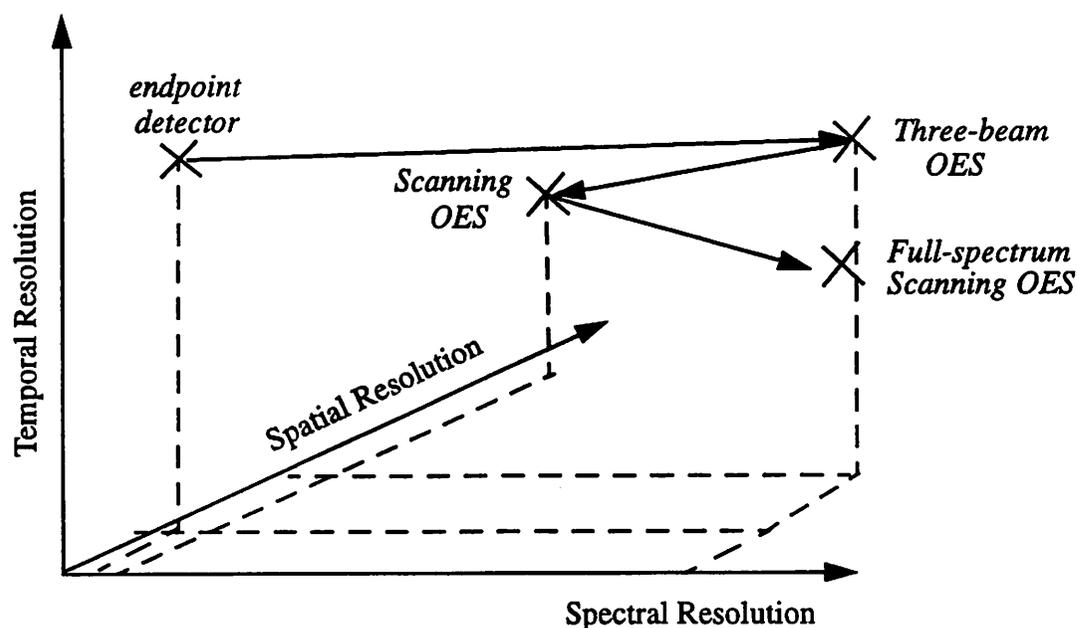


Figure 7.1 OES “resolutions”

that the CD spatial variation is mostly explained by the temporal, rather than their spatial component of sensor signals. The monitoring of the spatial CD variation *in-situ* is also accomplished by a statistic describing the variation of the real-time OES signals. This SPC scheme is employed to indicate the nature of the CD spatial non-uniformity.

The scanning SROES provides another way to monitor the spatial variability of etching by enhancing the spatial resolution of the SROES sensor, as illustrated in Figure 7.1. The time dependence of the spatial emission profiles is used to resolve the spatial endpoints. The information extracted from the 3-D spatial profiles is further used to monitor the spatial uniformity of etch process. This project is also extended to a full-spectrum scanning OES system, in which we can achieve high resolution of spatial, and spectral information (see Figure 7.1), however, at the expense of temporal resolution. Potentially, these color 3-D signatures can be used to characterize the plasma conditions.

In summary, we believe that the information provided by the OES sensors can be categorized into three parts: First, the emission intensity level at selected wavelengths

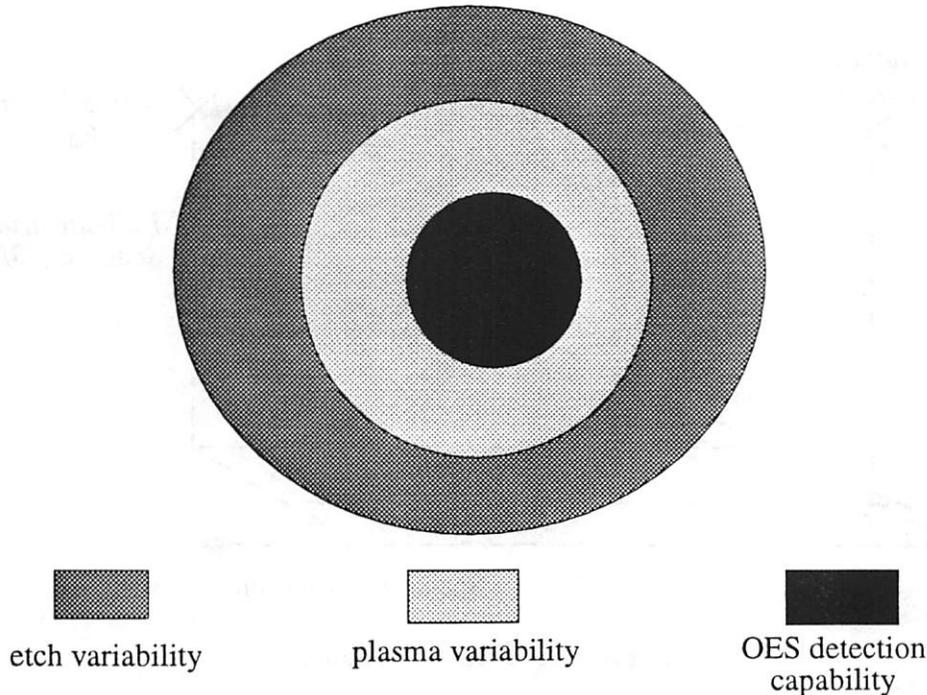


Figure 7.2 A schematic illustration of the relations among etch variability, plasma variability, and OES detection capability (not in scale).

can be used to characterize the density of excited species. This might relate to etch activity that depends on the etch mechanisms; Secondly, the ratio of the intensities of two selected wavelengths can be used to normalize the intensity drift and relate to the ion density (e.g., actinometry), for selected wavelengths; Thirdly, the emission profiles, including both temporal and spatial profiles, can be used to indicate plasma and etch uniformity.

Nevertheless, there are several issues related to the implementation of OES sensors. First, an OES sensor can only partially explain the causes of etch process variability. Because etch process involves a wide variety of reactions among molecules, atoms, ions, and electrons, an OES sensor is only capable of capturing the density variability of the chemical species visible on the collected spectrum. It is still questionable whether the OES sensor is able to detect other non-plasma factors such as wafer temperature, chamber wall temperature, polymer deposition, thin film uniformity and quality, and

lithography variations, which also have significant impact on the wafer. As a result, the emission signals collected from the OES sensor relate more to etch rates, rather than selectivity, etch-stop, anisotropy, plasma damage, and microloading effects. Figure 7.2 illustrates the observed relations among the etch process variability, plasma variability, and OES detection capability.

7.2. Future Work

Although the methods developed in this study are based on OES data collected from plasma etchers, the methodology presented is general and can be applied to other types of equipment and sensor readings. For example, multivariate data collected from RF monitors, or residual gas analysis (RGA) sensors can be used in the same manner. Moreover, this methodology can also be applied to other semiconductor equipment that can be monitored by a multivariate sensor.

By comparison with other *in-situ* OES sensors, our unique full-spectral range scanning SROES sensor, providing a high resolution spatial and spectral information, is promising in the application of *in-situ* process control and monitoring. The temporal resolution (i.e., acquisition rates) can be enhanced by a faster microprocessor/microcontroller. SROES can be further extended to monitor the spatially-resolved endpoints and characterize the plasma conditions. For example, the plasma etch equipment can be conditioned by the scanning OES signatures, instead of using monitor wafers. Statistical data filters such as PCA may be utilized to compress the spectral information so that a compressed 2-D spatial plot based on principal components is extracted for the characterization of the plasma conditions. The etch transition detection method based on PCA and jump linear filtering [7.1] might be another solution that will determine the spatially-resolved endpoints.

References for Chapter 7

- [7.1] S. Rangan, *et al.*, "Modeling and Filtering of Optical Emission Spectroscopy Data for Plasma Etching Systems", 1997 IEEE International Symposium on Semiconductor Manufacturing Conference, Oct. 6-8, 1997, San Francisco, California, p. B-41

Appendix A

S-plus Code

A.1 The Splus routine for quick principal component analysis

```

> PCA
function(X)
(
  nr <- nrow(X)
  nc <- ncol(X)
  Xbar <- apply(X, 2, mean)      # mean of columns (length=nc)
  Xout <- outer(rep(1, nr), Xbar) # rows of Xbar repeated nr times
  Xc <- X - Xout # Centered X
  XXt <- (Xc %*% t(Xc))/nr      # XXt instead of XtX
  o <- eigen(XXt, symmetric = T) # do eigendecomp. of XXt
  V <- matrix(0, nr, nc)
  V <- t(X) %*% o$vectors # get e-vectors of XtX
  V <- t(t(V)/apply(V, 2, vecnorm)) # make all e-vectors unit
length
  list(scores = ((X - Xout) %*% V)[, 1:nr - 1], loadings = V[, 1:nr -
1],
        center = Xbar, var = (o$values)[1:nr - 1])
)

```

The bootstrap routine for cross-validation

```

> PCACV
function(X, y, maxFactors)
(
  N <- nrow(X) # total number of samples (wafers)
  resid <- matrix(0, N, maxFactors) # residual (prediction error)
matrix
  PRESS <- numeric(maxFactors) # prediction error sum of squares
  for(i in 1:N) (

```

Appendix A

```
# Do PCA on all samples except one
pcobj <- MYPC(X[ - i, ])
for(j in 1:maxFactors) {
# Fit y to linear model of the first j scores (with intercept)
lsobj <- lsfit(pcobj$Scores[, 1:j], y[ - i], intercept
= T)
# Predict the excluded sample and calculate residual (pred. error)
Xc <- X[i, ] - pcobj$center # center X[i,]
sco <- t(as.matrix(Xc)) %*% pcobj$loadings[, 1:j]
# scores
pred <- sum(as.numeric(sco) * lsobj$coef[-1]) + lsobj$
coef[1]
resid[i, j] <- pred - y[i]
PRESS[j] <- PRESS[j] + (resid[i, j])^2
}
cat("i=", i, "\n")
}
list(resid = resid, PRESS = PRESS)
}
```

A.2 The Splus routine for partial least-square analysis

```
> PLS1.pred
function(X, y, maxa)
(
Xc <- scale(X, center = T, scale = F) # Center each column
yc <- scale(y, center = T, scale = F)
nr <- nrow(Xc)
nc <- ncol(Xc)
w <- matrix(0, nc, maxa) ## Correlation between inputs/response
tt <- matrix(0, nr, maxa) ## Scores
p <- matrix(0, nc, maxa) ## Input loadings
qq <- as.numeric(0) ## Response loading (length=maxa)
# for prediction
aveX <- apply(X, 2, mean) ## avg. of columns (a vector, length=nc)
aveY <- mean(y) ## (a scalar)
b <- matrix(0, nc, maxa) ## Beta for prediction
b0 <- matrix(0, 1, maxa) ## Offset (constant term)
a <- 1
while(T) {
v <- as.numeric(t(Xc) %*% yc)## (yc assumed to be a column
matrix)
w[, a] <- v/sqrt(sum(v * v))
## loading weights (unit length)
tt[, a] <- Xc %*% w[, a] ## input scores
v <- as.numeric(tt[, a])
inv <- 1/sum(v * v)
p[, a] <- t(Xc) %*% tt[, a] * inv ## input loadings
qq[a] <- sum(yc * v) * inv ## response loading
Xc <- Xc - as.matrix(tt[, a]) %*% t(p[, a])
yc <- yc - tt[, a] * qq[a]
## Coefficients for prediction. Model: y = b0 + Xb
b[, a] <- w[, 1:a] %*% solve(t(p[, 1:a]) %*% w[, 1:a]) %*% qq[1:a]
b0[, a] <- aveY - sum(aveX * as.numeric(b[, a])) # (scalar)
## convergence criteria
if(a == maxa)
break
}
```

```

        if(a == nr - 1) break
      #   if (sum(diag(t(Xc) %*% Xc)) <= 5) break
#   if (sum(abs(yc)) <= 1e-10) break
      a <- a + 1
    }
# return values and names
  list(weights = w, scores = tt, loadings = p, q = qq, E = Xc, f = yc, b
        = b, b0 = b0, aveX = aveX, aveY = aveY)
)

```

A.3 The Jackknife routine for cross-validation

```

> PLS1CV
function(X, y, maxFactors)
(
  N <- nrow(X)      # total number of samples (wafers)
  resid <- matrix(0, N, maxFactors)      # residual (prediction error)
matrix
  PRESS <- numeric(maxFactors)      # prediction error sum of squares
  for(i in 1:N) (
# Do PLS on all samples except one
    obj <- PLS1.pred(X[ - i, ], y[ - i], maxFactors)
    # Predict the excluded sample and calculate residual (prediction
error)
# for each number of PLS factors
    for(j in 1:maxFactors) (
      resid[i, j] <- (obj$b0[, j] + sum(X[i, ] * obj$b[, j])
) - y[i]
      PRESS[j] <- PRESS[j] + (resid[i, j])^2
    )
    cat("i=", i, "\n")
  )
  list(resid = resid, PRESS = PRESS)
)

```

A.4 Routine for determining OES temporal variance

```

> OES.temp
#temporal plots at selected wavelength for each wafer for exp2
#Mc: pixel number. (Mc=34 => 262nm)
function(Waf1=2,Waf2, Mc=34)
(
  OES1 <- numeric()
  OES2 <- numeric()
  DEV <- numeric()
  VAR <- numeric()
  wafNameList <- character()
  par(mfrow=c(3,1))
# determine the trend model

  trend <- matrix(scan("oes_st2/waf16"), ncol=1024, byrow=T)
  leng <- nrow(trend)
  for(q in 1:30){
    if(trend[q+2,Mc] > 3*trend[q,Mc]){
      m <- (q+3)
      break
    }else{
      m <- -1
    }
  }
)

```

Appendix A

```

    )
    for(j in 40:(leng-1)) {
      if(trend[j+1,Mc] < 0.5*trend[j,Mc]){
        n <- j-1
        break
      }else {n<-leng}
    }
    trend <- trend[m:n,Mc]
    trend0_trend
    LL <- length(trend)
    for (k in 20:LL){
      , if(trend[k]==min(trend[20:LL])){
        n0 <- k
        break
      }
      else{ }
    }
  )
  shift<- trend[n0-1]-trend[n0+2]

  trend <- c(trend[1:(n0-1)], (trend[(n0+2):LL]+shift))
  trend <- (trend-min(trend))/(max(trend)-min(trend))
  trend <- supsmu(1:length(trend),trend,span=1/40,bass=10)$y

# read data
for(i in Waf1:Waf2) {
  if(i < 10) {
    wafPrefix <- "waf0"
  } else {
    wafPrefix <- "waf"
  }
  wafName1 <- paste("oes_st1/",wafPrefix, i, ".st1", sep = "")
  wafName2 <- paste("oes_st2/",wafPrefix, i, sep = "")
  wafName3 <- paste("oes_st3/",wafPrefix, i, ".st3", sep = "")

  W1 <- matrix(scan(wafName1,numeric()),ncol=1024,byrow=T)
  W2 <- matrix(scan(wafName2,numeric()),ncol=1024,byrow=T)
  W3 <- matrix(scan(wafName3,numeric()),ncol=1024,byrow=T)
  t <- nrow(W2)
  title <- paste("wafer ",i,"\n")
  #In this matrix, col is wavelength, row is time;

  if (i!=34 & i!=4 & i!=5 & i!=20 & i!=24 & i!=27 & i!=30){
    # Remove the zero points
    if (Mc==34){
      for(q in 1:30){
        if(W2[q+2,Mc] > 3*W2[q,Mc]){
          m <- (q+3)
          break
        }else{
          m <-2
        }
      }
    }
    for(j in 10:t-1) {
      if(W2[j+1,Mc] < 0.5*W2[j,Mc]){
        n <- j-2
        break
      }else {n<-(t-1)}
    }
    cat(i,m,n,"\n")
  }
}

```

```

W1 <- W1[m:n,Mc]
W2 <- W2[m:n,Mc]
W3 <- W3[m:n,Mc]
L0 <- length(W2)

D1 <- supsmu(1:L0,W1,span=1/40,bass=10)$y
D2 <- supsmu(1:L0,W2,span=1/40,bass=10)$y
D3 <- supsmu(1:L0,W3,span=1/40,bass=10)$y

diffW1_diff(D1)
diffW2_diff(D2)
diffW3_diff(D3)

diffW1_supsmu(1:length(diffW1),diffW1,
              span=1/10,bass=20)$y
diffW2_supsmu(1:length(diffW2),diffW2,
              span=1/10,bass=20)$y
diffW3_supsmu(1:length(diffW3),diffW3,
              span=1/10,bass=20)$y

for (k in 20:length(diffW1)-3){
  if (diffW1[k]==min(diffW1[20:length(diffW1)-3])){
    n1 <- k
    break
  }
  else{ }
}

cat(i,m,n1,n,"\\n")
Title <- paste("wafer#",i,sep=" ")
shift<- W1[n1]-W1[n1+2]
new <- c(W1[1:n1],(W1[(n1+2):L0]+shift))
normal <- (new-min(new))/(max(new)-min(new))

trend1 <- approx(trend,n=length(normal))$y
dev <- normal-trend1
plot(W1,ylab="Cl_st1",xlab="etch time",
     type="b",main=Title)
plot(new,ylab="Cl_st1",xlab="etch time",
     type="b",main=Title)
plot(dev,ylab="deviation",xlab="etch time",
     type="l",main=Title)
OES1 <- c(OES1,W1)
OES2 <- c(OES2,new)
DEV <- c(DEV,dev)

var <- cbind(i,sum(dev^2)/(length(dev)-1))
VAR <- rbind(VAR,var)
}
)

par(mfrow=c(2,1))list(trend=cbind(1:length(trend1),trend1),
OES1=cbind(1:length(OES1),OES1),
OES2=cbind(1:length(OES2),OES2), DEV=cbind(1:length(DEV),DEV))
plot(OES1,xlab="time",ylab="OES Intensity(C12)",type="l")
plot(OES2,xlab="time",ylab="OES Intensity(C12)",type="l")
plot(DEV,xlab="time",ylab="Deviation",type="l")
plot(VAR,xlab="wafer order",ylab="Temporal Variance",type="b")
)

```

Appendix A

A.5 Routine for read and analysis of the scanning SROES data (monochromater case)

```
function(waf1=6,waf2=12,smooth=T)
(
  par(mfrow=c(4,2))
  timeall <- numeric()
  unif.all <- numeric()
  #default parameter

  # s is angular point
  s <- 10
  # t is cycle number
  t <- 15

  for (i in waf1:waf2){

    OES <- numeric()
    OESnew <- numeric()
    cycleNameList <- character()
    cat(paste("waf",i," is loading",sep=""),"\n")
    wafername <- paste("wafer ", i, sep="")
    cycl <- 0
    cyc2.matrix <- cbind(c(1:15),c(89,36,19,95,0,84,79,84,85,81,
      82,80,38,61,21)-1)
    if (i==cyc2.matrix[i,1]){cyc2 <- cyc2.matrix[i,2]}

    # read data from ASCII format

    if(i < 10) {
      waferPrefix <- "wafer00"
    } else {
      waferPrefix <- "wafer0"
    }
    wafname <- paste("data/",waferPrefix,i,"/",sep="")

    for(j in cycl:cyc2) (
      if(j < 10) {
        cyclePrefix <- "cycle00"
      } else {
        cyclePrefix <- "cycle0"
      }

      cycleName1 <- paste(wafname,cyclePrefix, j, ".dat", sep = "")
      cat(paste("cycle",j," is loading",sep=""),"\n")
      W <- matrix(scan(cycleName1,numeric()),ncol=2,byrow=T)

      OES <- rbind(OES,W)
    )

    time <- matrix(OES[,1],ncol=20,byrow=T)
    cycletime<- time[3,1]-time[2,1]
    cat(paste("Cycle time for wafer#",i, " is ", cycletime," seconds",
      sep=""),"\n")
  }
)
```

```

OES <- OES[,-1]
# arrange the data set into a 3-d array format
OES <- matrix(OES,ncol=20,byrow=T)
# row:time; col:spatial points ==> OES[t,s]

# Plots of one-dimensional signatures

#      par(mfrow=c(2,1))
#      Sp <- c(1:20)
#      Time <- c(cyc1:cyc2)*cycletime

#      # plot spatial profile

#      plot(Sp,OES[t,],xlab="angular points",ylab="OES intensity",type="l")
# plot temporal profile

#      plot(Time,OES[,s],xlab="cycle number",ylab="OES intensity",type="l",
#           main=wafername)

# Plots of two-dimensional signatures

#      par(mfrow=c(1,1))
#      persp(Time,Sp,OES,xlab="time(sec)",ylab="Angular
#           point",zlab="OE intensity",cex=0.7)
#           zlab="OE intensity",cex=0.7)
#           if (i <10){
#               printname1a <- paste("plot.waf0", i, ".ps", sep="")
#           }
#           else{
#               printname1a <- paste("plot.waf", i, ".ps", sep="")
#           }
#           printgraph(file=printname1a)

# Extract OES time dependence
if (smooth==T){
  # smoothing
  data1.1 <- numeric()
  data <- numeric()
  #cycle
  m <- nrow(OES)
  #spatial point
  n <- ncol(OES)

  # smooth x-axis
  for (i in 1:m){
    D1.1 <- supsmu(1:n,OES[i,],span=1/40,bass=10)$y
    data1.1 <- rbind(data1.1,D1.1)
  }

  # smooth y-axis
  for (j in 1:n){
    D1 <- supsmu(1:m,data1.1[,j],span=1/40,bass=10)$y
    data <- cbind(data,D1)
  }
}
else{
  data <- OES
}

totaltime <- nrow(data)

```

Appendix A

```
time <- numeric()
uniformity <- numeric()
ia <- 2
ib <- 19
offset <- 10
par(mfrow=c(2,2))
if (i == 12){offset <- 7}

for (k in 15:(totaltime-offset)){
  OESratio <- data[k+1,ia:ib]/data[k,ia:ib]
  OESnew <- rbind(OESnew,OESratio)
  time <- rbind(time,Time[k])

  unif <- var(OESratio[])/mean(OESratio[])
  uniformity <- rbind(uniformity,unif)
}

persp(time, Sp[ia:ib], OESnew, xlab="time(sec)", ylab="Angular point",
       zlim=c(0.8,1.2), zlab="It/It-1")
if (i <10){
  printname1a <- paste("3Dratio.waf0", i, ".ps", sep="")
} else{
  printname1a <- paste("3Dratio.waf", i, ".ps", sep="")
}
printgraph(file=printname1a)

# determine the spatial uniformity information
par(mfrow=c(2,2))
plot(x=time,y=uniformity,type="l",xlab="time",ylab="",
     ylim=c(0,0.00012))
if (i <10){
  printname1a <- paste("unif.waf0", i, ".ps", sep="")
} else{
  printname1a <- paste("unif.waf", i, ".ps", sep="")
}
printgraph(file=printname1a)
unif.all <- rbind(unif.all, uniformity)
cat("waiting.....","\n")
)
par(mfrow=c(2,1))
plot(unif.all,xlab="time",ylab="",type="l")
)
```

A.6 Routine for read and analysis of the scanning SROES data (spectroscopy case)

```
function(waf1,waf2, pca=T, nor=F)
{
  par(mfrow=c(2,2))

  #default parameter

  #s is angular point
  s <- 10
  #t is cycle number
  t <- 10

  # tranfer pixel number to wavelength
  pixel <- matrix(c(1,36,54,512),ncol=1)
  lambda<- matrix(c(200,262,309,1024),ncol=1)
```

```

a <- coef.lm(lm((lambda-pixel)))[1]
b <- coef.lm(lm((lambda-pixel)))[2]

x <- c(1:512)
wavelength <- a + b*x

for (i in waf1:waf2) {
  OES <- numeric()
  cycleNameList <- character()
  cat(paste("waf",i," is loading",sep=""),"\n")
  wafername <- paste("wafer ", i, sep="")
  cyc1 <- 4
  cyc2.matrix <- cbind(c(1:20),c(29,32,32,34,28,31,27,28,29,30,29
,28,31,25,29,28,28,27,29,29)-1)
  if (i==cyc2.matrix[i,1]){cyc2 <- cyc2.matrix[i,2]}

  # read data from ASCII format
  if(i < 10) {
    waferPrefix <- "waf0"
  } else {
    waferPrefix <- "waf"
  }
  wafname <- paste("data/",waferPrefix,i,"/",sep="")

  for(j in cyc1:cyc2) {
    if(j < 10) {
      cyclePrefix <- "CYCLE00"
    } else {
      cyclePrefix <- "CYCLE0"
    }

    cycleName1 <- paste(wafname,cyclePrefix, j,".DAT.d", sep = "")
    W <- matrix(scan(cycleName1,numeric()),ncol=513,byrow=T)
    OES <- rbind(OES,W)
  }

  time <- matrix(OES[,1],ncol=20,byrow=T)
  cycletime <- time[3,1]-time[2,1]
  cat(paste("Cycle time for wafer#",i, " is ", cycletime," sec-
onds",sep=""),"\n")
  OES <- OES[,-1]
  # arrange the data set into a 3-d array format
  OES <- array(OES,dim=c(20,(cyc2-cyc1+1),512))

  # Plots of one-dimensional signatures

  # plot spectrum
  plot(OES[s,t,],xlab="pixel number",ylab="OES intensity",type="l")
  par(new=T,yaxs="d")
  plot(x=wave-
length,y=OES[s,t,],axes=F,type="l",lty=2,xlab="",ylab="",mex=0.1)
  axis(side=3)
  mtext(side=3,line=3,"Wavelength(nm)")
  par(yxas="r")

  write(t(OES[s,t,]),ncol=1,file="spectrum")

```

Appendix A

```
# plot spatial profile
# pixel 36 is the main Br peak, Pixel 407 is Cl line
Mc <- 36
Mc2 <- 407
plot(OES[,t,Mc],xlab="angular points",ylab="OES intensity",type="l")
plot(OES[,t,Mc2],xlab="angular points",ylab="OES intensity",type="l")
plot(OES[,t,Mc]/OES[,t,Mc2],xlab="angular points",ylab="OES inten-
sity",type="l")

# plot temporal profile
plot(OES[s,,Mc],xlab="cycle number",ylab="OES inten-
sity",type="l",main=wafername)
Time <- c(cyc1:cyc2)*cycletime
Sp <- c(1:20)

if (pca == T){
  par(mfrow=c(3,3))
  OESPCA1 <- numeric()
  OESPCA2 <- numeric()
  OESPCA3 <- numeric()
  OESPCA4 <- numeric()

  for (k in 1:20){

    OES.PCA <- PCA(OES[k,,])$scores[,1:4]
    OESPCA1 <- cbind(OESPCA1,OES.PCA[,1])
    OESPCA2 <- cbind(OESPCA2,OES.PCA[,2])
    OESPCA3 <- cbind(OESPCA3,OES.PCA[,3])
    OESPCA4 <- cbind(OESPCA4,OES.PCA[,4])
  }
  cat(dim(OESPCA1),"\n")
  persp(Time,Sp,OESPCA1,ylab="angular point",xlab="time(sec
ond)",zlab="PC1")
  persp(Time,Sp,OESPCA2,ylab="angular point",xlab="time(sec
ond)",zlab="PC2")
  persp(Time,Sp,OESPCA3,ylab="angular point",xlab="time(sec
ond)",zlab=paste("PC3",sep=""))
  persp(Time,Sp,OESPCA4,ylab="angular point",xlab="time(sec
ond)",zlab="PC4")
  if (i <10){
    printname <- paste("T_W_PC4.waf0", i, ".ps", sep="")
  }
  else{
    printname <- paste("T_W_PC4.waf", i, ".ps", sep="")
  }
  printgraph(file=printname)
}

if (pca == F){
# Plots of two-dimensional signatures
  par(mfrow=c(1,1))
  persp(Time,wavelength,OES[s,,],xlab="time(sec)",ylab="wave-
length(nm)",cex=0.5)
  if (i <10){
    printname1a <- paste("T_W.waf0", i, ".ps", sep="")
  }
  else{
    printname1a <- paste("T_W.waf", i, ".ps", sep="")
  }
  printgraph(file=printname1a)

  par(mfrow=c(1,2))
  image(Time,wavelength,OES[s,,],xlab="time(sec)",ylab="wave
```

```

    length(nm)", cex=0.7)
if (i <10){
  printname1b <- paste("T_W_image.waf0", i, ".ps", sep="")
} else{
  printname1b <- paste("T_W_image.waf", i, ".ps", sep="")
}
printgraph(file=printname1b)

par(mfrow=c(1,1))
persp(Sp,wavelength[],OES[,t,],xlab="angular point",ylab="Wave
length(nm)", cex=0.6)
if (i <10){
  printname2a <- paste("S_W.waf0", i, ".ps", sep="")
} else{
  printname2a <- paste("S_W.waf", i, ".ps", sep="")
}
printgraph(file=printname2a)
par(mfrow=c(1,2))
image(Sp,wavelength,OES[,t,],xlab="angular point",ylab="Wave
length(nm)", cex=0.7)
if (i <10){
  printname2b <- paste("S_W_image.waf0", i, ".ps", sep="")
} else{
  printname2b <- paste("S_W_image.waf", i, ".ps", sep="")
}
printgraph(file=printname2b)

# normalize the temporal and sptail profiles to between 0 and 1
if (nor == T){
  par(mfrow=c(2,2))
  t0 <- 10
  ia <- 3
  ib <- 17
  OES1 <- (OES[ia:ib, ,Mc]-min(OES[ia:ib,t0,Mc]))/(max(OES[ia:ib,t0,Mc])
-min(OES[ia:ib,t0,Mc]))
  OES2 <- (OES[ia:ib, ,Mc2]-min(OES[ia:ib,t0,Mc2]))/
(max(OES[ia:ib,t0,Mc2])-min(OES[ia:ib,t0,Mc2]))
  persp(Sp[ia:ib],Time,OES1/OES2,xlab="",ylab="",cex=0.7,zlab="")
  persp(Sp,Time,OES[, ,Mc2],xlab="angular point",ylab="Time(Sec)",cex=0.7)
  persp(Sp,Time,OES[, ,Mc]/OES[, ,Mc2],xlab="angular point",
ylab="Time(Sec)",cex=0.7,zlab="Br/Cl ratio")
}
else {
  par(mfrow=c(1,1))
  t0 <- 5
  persp(Sp,Time,OES[, ,Mc]/OES[, t0,Mc],xlab="angular point",
ylab="Time(Sec)",cex=0.7,zlim=c(0,2.5),zlab="OES intensity")
  persp(Sp,Time,OES[, ,Mc2],xlab="angular point",ylab="Time(Sec)",cex=0.7)
  persp(Sp,Time,OES[, ,Mc]/OES[, ,Mc2],xlab="angular point",
ylab="Time(Sec)",cex=0.7,zlab="Br/Cl ratio")
  image(Sp,Time,OES[, ,Mc],xlab="angular point",ylab="Time(Sec)")
  if (i <10){
    printname3 <- paste("S_T.waf0", i, ".ps", sep="")
  } else{
    printname3 <- paste("S_T.waf", i, ".ps", sep="")
  }
  printgraph(file=printname3)
}
}
cat("waiting.....", "\n")
)

```

Appendix A

A.7 Routine for wafer-state contour plot

```
function(waf1, waf2, loess = T){
  uniformity <- numeric()
  # col 1 in map is the die number, col2 and 3 are the site address
  map <- matrix(scan("map"), ncol = 2, byrow = T)
  num_nrow(map)

  par(mfrow = c(1, 1))
  del <-c(2,5,6,9,10,14,16,20)
  for(i in waf1:waf2){
    if (i!= 2 && i!=5 && i!=6 &&i!=9 && i!=10 && i!=14 && i!=16&& i!=20){
      cat(paste("waf",i," is loading",sep=""),"\n")
      wafername <- paste("wafer",i,sep="")

  ## read data from ASCII format
      if(i < 10) {
        waferPrefix <- "waf0"
      } else {
        waferPrefix <- "waf"
      }
      wafname.pre <- paste("pre/",waferPrefix,i,sep="")
      wafname.post <- paste("post/",waferPrefix,i,sep="")

      oxide.pre <- matrix(scan(wafname.pre,numeric()),ncol=1,byrow=T)
      oxide.post <- matrix(scan(wafname.post,numeric()),ncol=1,byrow=T)
      oxide <- (oxide.pre - oxide.post)
      unif <- sqrt(var(oxide))/mean(oxide)
      unif <- cbind(i,unif)
      uniformity <- rbind(uniformity, unif)

      # fomattting data
      topo <- cbind(map[, 1], map[, 2], oxide)
      dimnames(topo) <- list(1:num, c("x", "y", "z"))
      topo <- data.frame(topo)
      # plot

      if(loess) {
        #1. contour plot
        topo.loess <- loess(z ~ x * y, topo, degree = 1, span= 0.50)
        resid_topo.loess$residuals

        topo.mar <- list(x = seq(1, 9, 0.5),y = seq(1, 9, 0.5))
        topo.lop <- predict(topo.loess, expand.grid(topo.mar))

        contour(topo.mar$x,topo.mar$y,topo.lop,xlab="",ylab="",
          levels=seq(min(topo.lop),max(topo.lop),30),labex=0.5)
        points(topo)
      }
    }
  }
  par(mfrow=c(2,1))
  plot(x=uniformity[,1],y=uniformity[,2],xlab="wafer number",
    ylab="etch uniformity",type="b")
  printgraph(file="uniformity_plot.ps")
}
```

Appendix B

MATLAB Code

B.1 Scanning OES Profile

```

% GRAPH_EX
% This programs reads the scanning OES data from the file and creates
% a 3D plot.  If desired, will create ps files of plots:
%
%       oes2dxx.ps      2d profile
%       oes3dxx.ps      3d profile
%       oes_enptxx.ps   machine endpoint
%
% You may wish to change axes and labels on graphs.

% Constants
% -----
fnum = 20;           % File number
print_fig = 1;      % 1 = print figures, 0 = no printing
print_parse = 1;    % 1 = create parse data file, 0 = no file

tstep = 0.025;      % Sample period (in secs)
SICL = 405;         % Spectral lines
CL2 = 211;
line_num = SICL;    % Line wavelength, used in determining
                   % constants for parsing data

% Data parsing contants
pkwid = 8;          % max. width of the beginning peak
pksep = 90;         % min. separation between peaks
if (line_num == SICL)
    pkht = 0.1;     % min. height of the peak
else                % for Cl line
    pkht = 1.0;
end

```

Appendix B

```

% Read OES data from file
% -----
% x0    Scanning OES data
% fix   Fixed OES data
% t0    Time
disp('Reading OES data from file...');
[x0,fix,t0] = read_soes(fnum);
l0 = length(x0);
tint = [t0(1):tstep:t0(l0)-tstep]';

% Parse data into cycles
% -----
% Data (t0,x0) is interpolated to create (tint,xint) where the
% the time spacing is uniform. The interpolated data is then
% partitioned into cycles in the matrix Z
% Z(i,j) = OES value of i^th cycle, j^th position
% Tm(i) = time of the beginning of the i^th cycle

% Parse data routine
tint = [t0(1):tstep:t0(l0)-tstep]';
disp('Parsing and interpolating OES data...');
[xint,pkloc,Tm,Z] = parse_soes(t0,x0,tint,pkwid,pksep,pkht);
[n,per] = size(Z);

% Output matrix Z to file
if (print_parse)
    disp('Creating parse data file...');
    fn = sprintf('sroes%02d_parsed', fnum);
    fid = fopen(fn,'wt');
    if (fid <= 0)
        error('Cannot open parse data file');
    end
    for i = 1:n
        fprintf(fid,'%12.4e ', Z(i,:));
        fprintf(fid,'\n');
    end
    fclose(fid);
end

% Plot data
% -----

% Figure 1: OES profile at various times
ta = 20; tb = 30;          % min. and max. cycle numbers to plot
ia = 45; ib = 77;        % min. and max. angular position numbers to plot
plot([ia:ib]', Z(ta:tb,ia:ib)');
xlabel('Angular position number');
ylabel('OES intensity');
title_str = sprintf('OES profile at various times from %4.1f to %4.1f
secs.',... Tm(ta),Tm(tb));
title(title_str);
if (print_fig)
    set(gcf,'InvertHardcopy','on');
    prtstr = sprintf('print oes2d%02d -dpsc', fnum);
    eval(prtstr);
end

```

```

end
disp('Hit any key to continue...');
pause;

% Figure 2: 3-D plot of OES data

ta = 1; tb = length(Tm);
mesh([ia:ib]', Tm(ta:tb) , Z(ta:tb,ia:ib));
view([1 1 1]);
Zmax = max(max(Z));
axis([ia ib Tm(ta) Tm(tb) 0 Zmax]);
xlabel('Angular position number');
ylabel('Time (secs.)');
zlabel('OES intensity');
title('Scanning OES profile');

if (print_fig)
    set(gcf,'InvertHardcopy','on');
    prtstr = sprintf('print oes3d%02d -dpsc', fnum);
    eval(prtstr);
end

disp('Hit any key to continue...');
pause;

% Figure 3: 3-D plot of OES smoothed data/ or Ratios
% A is the smoothed Z data file via splus

ta = 5; tb = length(Tm)-7;
mesh([ia+3:ib-3]', Tm(ta:tb) , A(ta:tb,ia+3:ib-3));
view([1 1 1]);
%Zmax = max(max(A));

axis([ia ib Tm(ta) Tm(tb) 0.7 1.3]);
xlabel('Angular position number');
ylabel('Time (secs.)');
zlabel('OES intensity ratio (smoothed)');
%title('Scanning OES profile');

if (print_fig)
    set(gcf,'InvertHardcopy','on');
    prtstr = sprintf('print oes3d%02d_smoothed -dpsc', fnum);
    eval(prtstr);
end

disp('Hit any key to continue...');
pause;

% Figure 4: Machine endpoint
% machine endp
ta = 800;
tb = 3995;
plot(t0(ta:tb),fix(ta:tb));
axis([t0(ta) t0(tb) 0.7 1.2]);
title('Machine endpoint');
xlabel('Time (secs)');
ylabel('OES intensity');
if (print_fig)

```

Appendix B

```

    set(gcf, 'InvertHardcopy', 'on');
    prtstr = sprintf('print oes_enpt%02d -dpsc', fnum);
    eval(prtstr);
end

*****
function [xint,pkloc,Tm,Z] = parse_oes(t0,x0,tint,pkwid,pksep,pkht)
% PARSE_SOES  Parses scanning OES data into cycles
%
%   [xint,pkloc,Tm,Z] = parse_oes(t0,x0,tint,pkwid,pksep,pkht);
%
% Parsing:
% -----
%   1. Raw data (t0,x0) is first interpolated onto time vector
%      tint to produce xint.
%   2. The beginnings of the sweeps are identified by a small
%      peaks. The small peaks are defined by the parameters pkht,
%      pkwid and pksep. Using these parameters, the peak locations
%      are determined and stored in the vectors pkloc, and Tm.
%   3. Data is organized into by cycles in the matrix Z.
%
% Arguments
% -----
% t0      Time values for OES data.
% x0      OES data at time values t0.
% tint    Time values to interpolate data.
% pkwid   Maximum width in data points of peaks
% pksep   Minimum separation between peaks in num. data points
% pkht    Minimum height of peak
%
% xint    Interpolated data: xint(i) = x at time = tint(i)
% pkloc   data point number of peak locations
% Tm      time of peak locations
% Z       Parsed OES data Z(i,:) = data from i th sweep

% Check time vectors
l0 = length(x0);
l1 = length(t0);
l = length(tint);
if (l1 ~= l0)
    error('vectors x0 and t0 must be the same length');
elseif (any(t0(2:l0) <= t0(1:l0-1))) | (any(tint(2:l) <= tint(1:l-1)))
    error('time values must be in ascending order');
elseif (tint(1) < t0(1)) | (tint(l) >= t0(l0))
    error('time values must be in ascending order');
end

% Interpolate data
xint = zeros(l,1);
k = 1;
for i = 1:l
    while (t0(k+1) <= tint(i))
        k = k+1;
    end
    p = (tint(i) - t0(k))/(t0(k+1)-t0(k));
    xint(i) = p*x0(k+1) + (1-p)*x0(k);
end

% Determine starting points of scan by identifying peaks

```

```

k = pkwid+1; n = 0; done = 0;
pkloc = [];
while (~done) & (k <= l-pkwid)
    while (xint(k) <= pkht) | (xint(k+pkwid) > pkht) ...
        | (xint(k - pkwid) > pkht)
            k = k + 1;
            if (k > l-pkwid)
                done = 1;
                break;
            end
        end
    end
    if (~done)
        [pkmax,j] = max(xint(k-pkwid:k+pkwid));
        n = n + 1;
        pkloc = [pkloc; k-pkwid+j-1];
        k = pkloc(n) + pksep;
    end
end

% Parse matrix into data Z.
per = round(mean(pkloc(2:n) - pkloc(1:n-1))); % Compute avg. period
pkloc = [flipud([pkloc(1)-per:-per:1]'); pkloc; ...
        [pkloc(n)+per:per:length(xint)-per+1]'];
% Add cycles at beginning and end of data to include
% power up and down periods
n = length(pkloc);
Z = zeros(n,per);
Tm = zeros(n,1);
for i = 1:n
    Z(i,:) = xint(pkloc(i):pkloc(i)+per-1');
    Tm(i) = tint(pkloc(i));
end

*****

% this function is for data coverstion from ASCII to matlab file
fid = fopen('wafer20_ratio')
A = fscanf(fid,'%f',[112 inf])
A = A'
fclose(fid)
rhodes.eecs 135# more read_soes.m
function [scan_dat,fix_dat,tm] = read_soes(fnum)
% READ_SOES Reads scanning OES data from file
%
% [scan_dat,fix_dat,tm] = read_soes(fnum);
%
% fnum File number (file is named 'datxx')
% scan_dat Scanning OES data
% fix_dat Fixed position OES data
% tm Time stamp

% Open data file
DAT_DIR = '/home/users/spanos/rwchen/crete/ScanOES/test';
%DAT_DIR = '/a/crete/crete1/rwchen/ScanOES/test';
fn = sprintf('%s/DAT%02d', DAT_DIR, fnum);
fid = fopen(fn,'rt');
if (fid <= 0)
    error('Cannot open data file');
end

```

Appendix B

```
% Read data
z = fscanf(fid, '%f', [3 inf]);
tm = z(1,:);
scan_dat = z(2,:);
fix_dat = z(3,:);

% Close file
fclose(fid);
```

Appendix C

Experimental Design

C.1 AMD's AMAT 5300 experiment

| Run # | Source (Watts) | Bias (Watts) | C ₂ F ₆ flow rate (sccm) | Roof Temp (°C) | Throttle (Percent) |
|-------|----------------|--------------|--|----------------|--------------------|
| 1 | 2300 | 1800 | 20 | 260 | 30 |
| 2 | 2900 | 1800 | 40 | 260 | 30 |
| 3 | 2300 | 1400 | 20 | 260 | 100 |
| 4 | 2300 | 1400 | 40 | 260 | 30 |
| 5 | 2900 | 1400 | 20 | 260 | 30 |
| 6 | 2900 | 1400 | 40 | 260 | 100 |
| 7 | 2300 | 1800 | 40 | 260 | 100 |
| 8 | 2600 | 1600 | 30 | 260 | 65 |
| 9 | 2900 | 1800 | 20 | 260 | 100 |
| 10 | 2600 | 1600 | 30 | 270 | 30 |
| 11* | 2600 | 1600 | 30 | 270 | 65 |
| 12* | 2600 | 1600 | 30 | 270 | 65 |
| 13* | 2600 | 1600 | 30 | 270 | 65 |
| 14 | 2600 | 1400 | 30 | 270 | 65 |

Appendix C

| Run # | Source (Watts) | Bias (Watts) | C ₂ F ₆ flow rate (sccm) | Roof Temp (°C) | Throttle (Percent) |
|-------|----------------|--------------|--|----------------|--------------------|
| 15 | 2900 | 1600 | 30 | 270 | 65 |
| 16 | 2600 | 1600 | 40 | 270 | 65 |
| 17 | 2300 | 1600 | 30 | 270 | 65 |
| 18 | 2600 | 1600 | 30 | 270 | 100 |
| 19 | 2600 | 1600 | 20 | 270 | 65 |
| 20 | 2600 | 1800 | 30 | 270 | 65 |
| 21* | 2600 | 1600 | 30 | 270 | 100 |
| 22 | 2300 | 1400 | 40 | 280 | 30 |
| 23 | 2900 | 1800 | 20 | 280 | 30 |
| 24 | 2300 | 1400 | 20 | 280 | 30 |
| 25 | 2900 | 1400 | 20 | 280 | 100 |
| 26 | 2900 | 1800 | 40 | 280 | 100 |
| 27 | 2900 | 1400 | 40 | 280 | 30 |
| 28 | 2300 | 1800 | 20 | 280 | 100 |
| 29 | 2600 | 1600 | 30 | 280 | 65 |
| 30 | 2300 | 1800 | 40 | 280 | 30 |

*: center points

C.2 Berkeley's Lam 4400 experiment

| Run | Pressure | RF Power | Gap | Cl ₂ /He Ratio | Cl ₂ | He |
|-----|----------|----------|-----|---------------------------|-----------------|-----|
| 1 | 370 | 235 | 0.9 | 0.42 | 183 | 437 |
| 2 | 370 | 275 | 0.8 | 0.45 | 180 | 400 |
| 3* | 425 | 235 | 0.8 | 0.45 | 180 | 400 |
| 4 | 370 | 235 | 0.7 | 0.48 | 200 | 420 |
| 5 | 425 | 275 | 0.7 | 0.45 | 180 | 400 |
| 6 | 425 | 275 | 0.8 | 0.42 | 172 | 408 |
| 7 | 425 | 235 | 0.8 | 0.45 | 180 | 400 |

| Run | Pressure | RF Power | Gap | Cl ₂ /He Ratio | Cl ₂ | He |
|-----|----------|----------|-----|---------------------------|-----------------|-----|
| 8 | 480 | 315 | 0.9 | 0.42 | 160 | 380 |
| 9 | 425 | 275 | 0.8 | 0.45 | 180 | 400 |
| 10* | 425 | 275 | 0.8 | 0.45 | 180 | 400 |
| 11 | 425 | 275 | 0.8 | 0.45 | 168 | 372 |
| 12 | 480 | 315 | 0.9 | 0.48 | 200 | 420 |
| 13 | 425 | 275 | 0.8 | 0.45 | 180 | 400 |
| 14 | 480 | 315 | 0.9 | 0.48 | 200 | 420 |
| 15 | 370 | 315 | 0.9 | 0.42 | 160 | 380 |
| 16* | 425 | 275 | 0.8 | 0.45 | 180 | 400 |
| 17 | 425 | 315 | 0.8 | 0.45 | 180 | 400 |
| 18 | 370 | 235 | 0.9 | 0.48 | 175 | 365 |
| 19* | 425 | 275 | 0.8 | 0.45 | 180 | 400 |
| 20 | 370 | 315 | 0.7 | 0.48 | 175 | 365 |
| 21 | 480 | 275 | 0.8 | 0.45 | 180 | 400 |
| 22 | 480 | 235 | 0.9 | 0.48 | 175 | 365 |
| 23* | 425 | 275 | 0.8 | 0.45 | 180 | 400 |
| 24 | 480 | 315 | 0.7 | 0.48 | 175 | 365 |
| 25* | 425 | 275 | 0.8 | 0.45 | 180 | 400 |
| 26 | 480 | 235 | 0.9 | 0.42 | 183 | 437 |
| 27 | 370 | 315 | 0.7 | 0.48 | 175 | 365 |
| 28 | 425 | 275 | 0.8 | 0.48 | 188 | 392 |
| 29 | 425 | 275 | 0.9 | 0.45 | 180 | 400 |
| 30 | 480 | 315 | 0.7 | 0.42 | 183 | 437 |
| 31 | 480 | 235 | 0.7 | 0.48 | 200 | 420 |
| 32 | 425 | 275 | 0.8 | 0.45 | 192 | 428 |
| 33* | 425 | 275 | 0.8 | 0.45 | 180 | 400 |
| 34 | 370 | 315 | 0.7 | 0.42 | 183 | 437 |

Appendix C

| Run | Pressure | RF Power | Gap | Cl ₂ /He Ratio | Cl ₂ | He |
|-----|----------|----------|-----|---------------------------|-----------------|-----|
| 35* | 425 | 275 | 0.8 | 0.45 | 180 | 400 |
| 36* | 425 | 275 | 0.8 | 0.45 | 180 | 400 |
| 37 | 370 | 315 | 0.9 | 0.48 | 200 | 420 |

*: center points

C.3 Berkeley's Lam 9400 experiment

| Run | Pressure (mtorr) | Top power (Watts) | Bottom power (Watts) | Cl ₂ (sccm) | HBr (sccm) | HBr/Cl ₂ | time (secs) |
|-----|------------------|-------------------|----------------------|------------------------|------------|---------------------|-------------|
| 1* | 12 | 300 | 150 | 50 | 150 | 3 | 60 |
| 2 | 14 | 270 | 165 | 40 | 160 | 4 | 60 |
| 3 | 10 | 270 | 135 | 65 | 130 | 2 | 60 |
| 4 | 14 | 330 | 165 | 40 | 160 | 4 | 68 |
| 5 | 10 | 330 | 135 | 65 | 130 | 2 | 60 |
| 6 | 10 | 330 | 165 | 65 | 130 | 2 | 67 |
| 7* | 12 | 300 | 150 | 50 | 150 | 3 | 60 |
| 8 | 14 | 270 | 135 | 65 | 130 | 2 | 60 |
| 9 | 14 | 330 | 165 | 65 | 130 | 2 | 65 |
| 10 | 14 | 330 | 135 | 65 | 130 | 2 | 65 |
| 11 | 14 | 270 | 165 | 65 | 130 | 2 | 65 |
| 12 | 10 | 270 | 135 | 40 | 160 | 4 | 65 |
| 13 | 14 | 330 | 135 | 40 | 160 | 4 | 60 |
| 14 | 10 | 330 | 135 | 40 | 160 | 4 | 63 |
| 15 | 14 | 270 | 135 | 40 | 160 | 4 | 63 |
| 16 | 10 | 330 | 165 | 40 | 160 | 4 | 60 |
| 17* | 12 | 300 | 150 | 50 | 150 | 3 | 60 |
| 18 | 10 | 270 | 165 | 65 | 130 | 2 | 60 |

| Run | Pressure (mtorr) | Top power (Watts) | Bottom power (Watts) | Cl ₂ (sccm) | HBr (sccm) | HBr/Cl ₂ | time (secs) |
|-----|------------------|-------------------|----------------------|------------------------|------------|---------------------|-------------|
| 19 | 10 | 270 | 165 | 40 | 160 | 4 | 63 |
| 20* | 12 | 300 | 150 | 50 | 150 | 3 | 60 |

*: center points

Appendix D

Locally Weighted Regression Smoothing

In locally weighted regression smoothing, we build the smooth function $s(x)$ point-wise as follows:

1. Take a point, say x_0 . Find the k nearest neighbors of x_0 , which constitute a neighborhood $N(x_0)$. The number of neighbors k is specified as a percentage of the total number of points. This percentage is called the span.
2. Calculate the largest distance between x_0 and another point in the neighborhood:

$$\Delta(x_0) = \max_{N(x_0)} |x_0 - x_i|$$

3. Assign weights to each point in $N(x_0)$ using the tri-cube weight function:

$$W\left(\frac{|x_0 - x_i|}{\Delta(x_0)}\right)$$

where $W(u) = ((1 - u^3)^3,) \quad \text{for } 0 \leq u < 1$

$$W(u) = 0 \quad \text{otherwise}$$

4. Calculate the weighted least squares fit of y on the neighborhood $N(x_0)$. Take the fitted value $\hat{y}_0 = s(x_0)$.
5. Repeat for each predictor value.

Appendix E

Element vs. Wavelength

| element | wavelengths (nm) |
|-----------------|---|
| Al | 308,309,394,396 |
| AlCl | 261,265,268 |
| AlH | 426 |
| Ar | 416,451,485,550,603,697,707,750 |
| B | 250 |
| BCl | 267,272 |
| Br | 290,334,355,470,479,570,576,588,600 |
| C | 248 |
| CCl | 258,307,460 |
| CF | 240,256 |
| CF ₂ | 249,252,246,255,260,263,271,275,288,292,321 |
| CH | 431 |
| CN | 359,386,387,418,420,422,647,693,709,785 |
| CO | 239,246,249,271,283,313,349,370,451,484,520,561,608,662 |
| CO ₂ | 288,290,337 |

Appendix E

| element | wavelengths (nm) |
|------------------|---|
| Cl | 726,741,755,772,775,838 |
| Cl ₂ | 262,309 |
| Cu | 325,327 |
| F | 624 ,635,641,677,683,686,687,690,691,697,704,713,720,733,740, 743,751,755,757 |
| H | 434,486,656 |
| He | 295,319,345,361,363,371,382,387,389,397,403,414,439,444,447, 471,492,502,505,588,668,707,728 |
| Hg | 253.7,365.0,404.7,435.8,546.1 |
| N ₂ | 337,326,331,390,391,428,576,580,655,662,671,688,727,790 |
| NO | 237,245,256,268,272,286 |
| O | 437,497,502,533,544,605,616,646,700,725,777 |
| NH | 336 |
| OH | 307 |
| Si | 252,288 |
| SiF | 640,777 |
| SiCl | 281,282,287,405 |
| SiF ₂ | 390,401 |
| SiN | 441,405,409,413,420,424 |
| SiO | 234,241,249 |
| Ti | 335,365,400 |
| TiCl | 419 |
| TiF | 408 |
| W | 272, 401 ,407,430 |
| WF | 578 |

NOTE: the most prominent wavelength are bold