SUPERVISORY CONTROL SYSTEM FOR A PHOTOLITHOGRAPHIC WORKCELL

by

Sovarong Leang

Memorandum No. UCB/ERL M92/70

14 July 1992

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Acknowledgments

I would like to thank the Semiconductor Research Consortium (SRC) and the National Science Foundation (NSF) (MIP-9014940) for funding this work.

SUPERVISORY CONTROL SYSTEM FOR A PHOTOLITHOGRAPHIC WORKCELL

by
Sovarong Leang

Memorandum No. UCB/ERL M92/70

14 July 1992

SUPERVISORY CONTROL SYSTEM FOR A PHOTOLITHOGRAPHIC WORKCELL

by

Sovarong Leang

Memorandum No. UCB/ERL M92/70

14 July 1992

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Acknowledgments

I would like to thank the Semiconductor Research Consortium (SRC) and the National Science Foundation (NSF) (MIP-9014940) for funding this work.

Table of Contents

Chapter 1 Introduction 1
1.1 Background and Motivation 1
1.2 Thesis Organization 1

Chapter 2 Monitoring System 3
2.1 Introduction 3
2.2 Process Inputs and Outputs 4
2.2.1 Process Inputs 4
2.2.2 Typical Process Outputs 4
2.2.3 Proposed Process Outputs 6
2.3 Metrology 9
2.4 Choosing the Range of the Probing Wavelength 12
2.5 Summary 13

Chapter 3 Model Generation System 14
3.1 Introduction 14
3.2 Experimental Design 14
3.2.1 Empirical Model Generation for the Spin-Coat and Bake Equipment 15
3.2.2 Stepper 20
3.2.3 Developer 22
3.3 Summary 23

Chapter 4 Feedback Control System 24
4.1 Introduction and Background 24
4.2 Detection of Process Disturbances 25
4.2.1 Malfunction Alarms 26
4.2.2 Control Alarms 27
4.3 Algorithm for Adaptively Updating Equipment Models 28
4.3.1 Automated Recipe Generation 30
4.4 Summary 31

Chapter 5 Feed-Forward Control System 32
5.1 Introduction 32
5.2 Feed-Forward Control Alarm Generation 32
5.3 Feed-Forward Control Methodology 33
5.4 Automated Recipe Generation 34
5.5 Summary 35

Chapter 6 Automation of the Supervisory Control System 36

Chapter 7 Experimental Results 37
7.1 Description of the Experimental Set-Up 37
7.1.1 Test Pattern Design and Fabrication 37
7.1.2 The Spin-Coat and Bake Equipment 39
7.1.3 The Stepper 40
7.1.4 The Developer 40

7.1.5 The Photospectrometer for Reflectance and Thickness Measurement 41
7.1.6 The Critical Dimension Measurement Computer 41
7.2 Data Collection and Screening 42
7.3 Experimental Results of the Feedback Control Mechanism 43
7.4 Experimental Results of the Feed-Forward Control Mechanism 49
7.5 Summary 52
Chapter 8 Conclusions 53
REFERENCES 55

Chapter 1 Introduction

1.1 Background and Motivation

To stay competitive, semiconductor industries must develop and efficiently produce new state of the art products. This requires among other things high yield processes. One way to increase the process yield is to reduce process variability, a rather difficult task, since semiconductor processes are not always well understood. Furthermore, processes often drift due to equipment aging, depletion of chemicals, fluctuation in ambient conditions, etc. Also, process parameters often change drastically after machines have been replaced or maintained. The impact of this instability is great because semiconductor process specifications are typically pushed to the limits of equipment capability, which cause the process tolerance to be of the same order as the natural variation of the equipment.

One approach to reduce process variation is to use a supervisory system that controls the process on a real-time basis. This is rendered possible with modern analytical and processing equipment that have the ability to interact with computer driven controllers. These controllers will be able to collect information, manipulate recipes and act immediately to compensate for process drifts. This thesis describes the development and the deployment of such a controller for a photolithographic process sequence (Figure 1).

1.2 Thesis Organization

The control system consists of three parts. The first part, described in chapter 2, is a monitoring system, whose goal is to collect data reliably, and economically. The second part, described in chapter 3, is a model generation system whose goal is to generate and update adaptive equipment models. And finally the third part, described in chapters 4 and 5, is a controller whose goal is to improve the performance of the photolithographic workcell through systematic use of statistical process control (SPC). Chapter 6 discusses the

automation of the supervisory control system. Chapter 7 describes the experimental lithography set-up and presents the experimental results. Conclusions and suggestions for future work are presented in chapter 8.

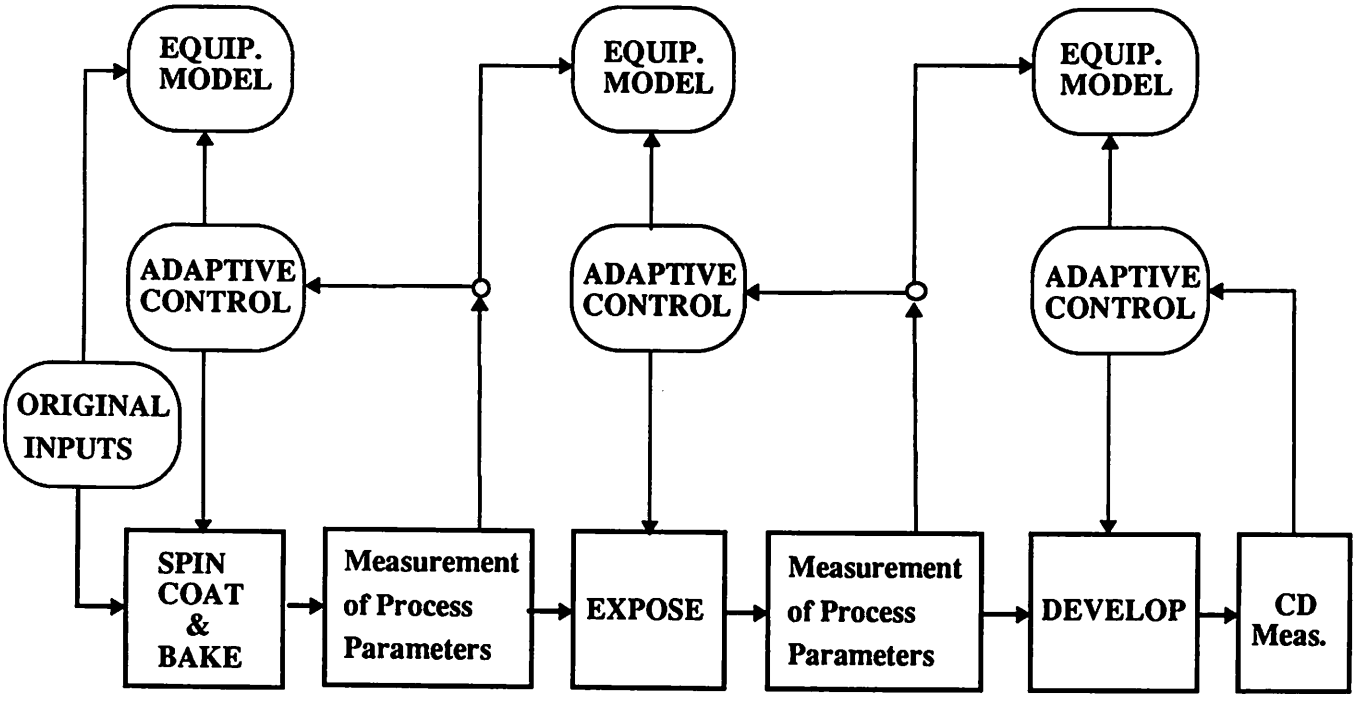


Figure 1 Schematic of a Supervisory Control System for the Lithography Workcell.

Chapter 2 Monitoring System

2.1 Introduction

Tight process specifications can be achieved by exploiting the interdependent parameters of the various photolithography steps. This can be accomplished by applying supervisory control across the equipment responsible for photoresist application, exposure and development (Figure 1) [1]. Traditionally, the photolithographic process has been monitored by measuring the critical dimensions (CDs) of developed photoresist patterns. Although control charts of CDs are very useful in maintaining the overall quality of the photolithographic sequence [2], any observed CD deviations carry the compounded effect of several processing steps, such as spin-coat, pre-bake and post-bake, exposure, and development.

Identification and isolation of the variability of each process step is necessary in order to minimize the variability of the overall pattern transfer operation. This may be accomplished by using SPC charts on the in-line measurements collected after the completion of each step. In this way, several critical equipment responses may be monitored over time. Statistically significant departures from routine operation of early steps may then be corrected by adjusting the process settings of the remaining steps. The settings of each step may also be adjusted to correct for systematic deviations caused by equipment aging, chemical depletion, etc.

However, a practical implementation of such a control scheme requires careful selection of the parameters to be monitored. These parameters should be easy to measure and convey useful information about the status of the wafer and the performance of the equipment. Efficient in-line measurements that will isolate the causes of process variability must therefore be developed for each step.

2.2 Process Inputs and Outputs

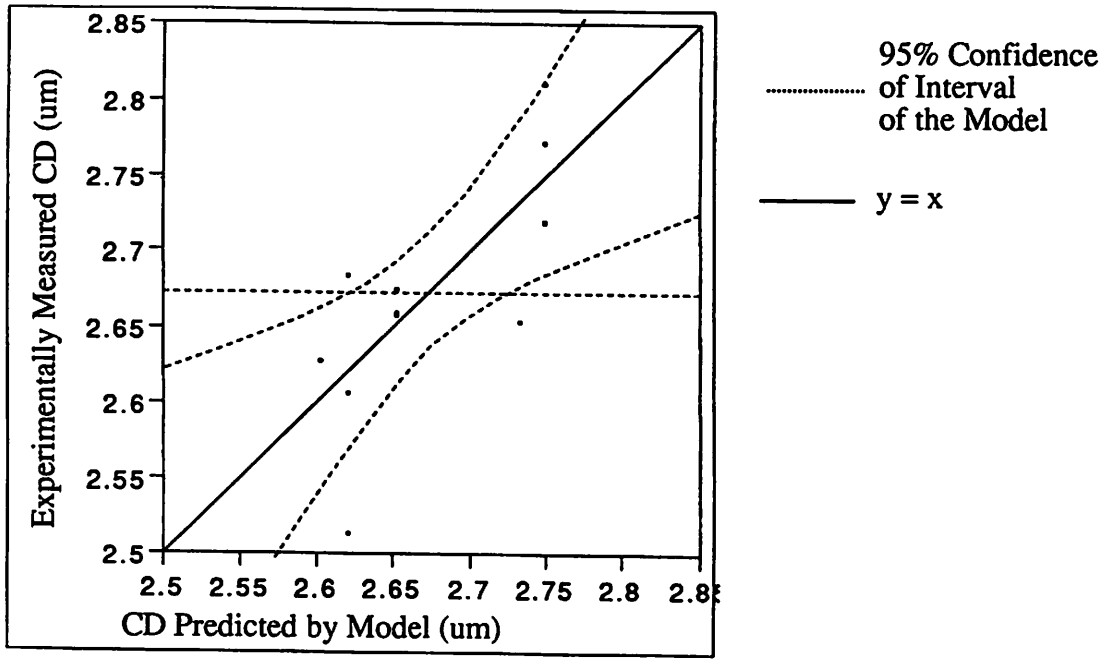
2.2.1 Process Inputs

Process inputs are the variables that affect the outcome of each photolithographic operation. They include easily monitored and controlled equipment settings, such as spin speed, spin time, baking temperature, baking time, exposure time, and development time. In addition to the process inputs, there are also "noise" variables, such as ambient temperature, ambient humidity, photoresist viscosity, and developer concentration that affect the outcome of processes. Since "noise" variables cannot be easily monitored, nor controlled on a run by run basis during production, they are not used as process inputs.

2.2.2 Typical Process Outputs

Historically, photolithography processes have been controlled just by monitoring the CD of the developed patterns. If it was within specifications, the wafer was etched; if not, the resist was stripped and the wafer is processed again. Sometimes, this inefficient CD control procedure is enhanced by monitoring the photoresist thickness. Although this is an improvement over the first CD control algorithm, the photoresist thickness alone gives insufficient information about the status of the photoresist.

To quantify this claim, a statistically designed experiment has been done to characterize the CD. Varying the thickness of the photoresist, and development time, we measured the CD of the resulting photolithography pattern and empirically modelled it using linear regression. The model obtained is shown in Figure 2.



Summary of Fit:

R^2:	0.56
Average Model Prediction Error:	±0.056 um (1σ)
Experimental Error:	±0.025 um (1σ)

Figure 2 Verification of an Empirical CD Model that only Involves Resist Thickness and Development Time

Before looking at the summary of fit, a description of Figure 2 is in order. Figure 2 compares the CD model prediction to the experimentally measured CD. The reason why the model prediction and the experimental results have not been displayed vs. the input parameters is because the model depends on more than two input parameters, and therefore needs more than three dimensions to be represented. Therefore, the type of plot used in Figure 2 is used to verify the validity of a model that depends on more than two terms.

If the model prediction reflects the experimental data well, these will lie on the y=x line. Therefore, since a few experimental data lie far away from the y=x line in this case, the model prediction is inadequate. The two dotted hyperbolae represent the ±2σ confidence interval limits of the model (95% confidence of prediction). Any point between the two hyperbolae represents a response that can be explained by the model. Outliers repre-

sent points whose behavior cannot be explained by the model. Note that the model prediction is more accurate in the center. The further away from the operating point we move, the less accurate the model becomes.

Finally, the *Summary of Fit* shows how well the data fit the model. A perfect fit would result in an R^2 of 1.0, and the closer to 1.0 R^2 is, the better the model. The model prediction error is also important because when compared to the experimental error, it tells us when all the data's variations have been modelled. If the model prediction error is larger than the experimental error, the model would be considered inadequate since it could not account for some experimental results. Only when the model prediction error is equal or smaller than the experimental error can we stop looking for additional terms to put in the model, since adding more terms after reaching that threshold would not give us a more significant model.

Therefore in this case, since 1) several points lie outside the confidence intervals of the model, 2) the R^2 is a mere 0.56, and 3) the experimental error is smaller than the prediction error of the model, the CD model based on resist thickness and development time only, is clearly inadequate. It is therefore necessary to include an additional process response in the CD model: the photoactive compound (PAC) concentration in the photoresist.

2.2.3 Proposed Process Outputs

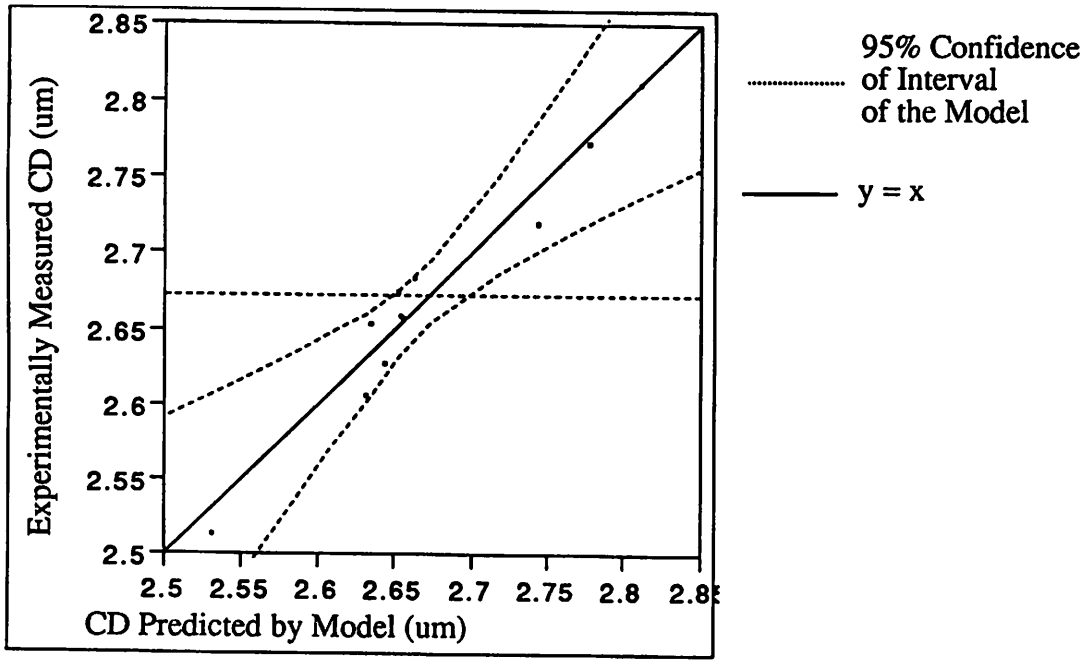
According to Dill [3], the photoactive compound concentration M can effectively model the exposure and development of positive photoresist. Dill's models have been incorporated in many accurate physical simulators [4] [5].

In Dill's exposure model, the absorption of light decreases as M decreases, and in his development model, the reaction can be approximated as a surface-limited etching reaction, whose rate is controlled by the degree of exposure and therefore by M. In summary,

the status of the photoresist can be effectively known by monitoring M. Unfortunately, M cannot be measured directly. Dill's method of extracting M, which consists of extracting the resist parameters A, B, and C, is equally difficult to implement on a production line, since it requires measuring the transmittance of the photoresist, which is done by coating with photoresist a glass wafer that has an index of refraction as close as possible to the index of refraction of the photoresist in order to cancel back surface reflection.

This difficulty in measuring M during production leads to the development of in-line measurements that could be used to instead *infer* the value of M. Watts suggested measuring the absorptivity of light in the photoresist layer [6]. Unfortunately, that procedure is also quite involved and cannot be easily performed during production. Another alternative is to obtain M by measuring the reflectance of the photoresist film, since the reflectance of the photoresist is related to the film absorptivity. However, many problems arise when measuring reflectance: because of the interference patterns of the light intensity within the resist and within the underlying transparent layers, the reflected beam strongly depends on the film thickness. The film thickness tends to vary during processing, since it is related to the photoresist viscosity and to the spin-coat and bake-plate performance. A metrology for reflectance measurements has been developed by Ling [7], that produces readings that are sensitive to changes in M, yet immune to the effects of thickness variation. Although some of the assumptions underlying Ling's metrology are weak, the incorporation of peak reflectance into the equipment models produces a better empirical CD model, as shown in

Figure 3. Therefore, while a better metrology for inferring M is being investigated, Ling's metrology is currently used in the photolithography process controller.



Summary of Fit:

R^2:	0.95
Average Model Prediction Error:	±0.025 um (1 σ)
Experimental Error:	±0.025 um (1 σ)

Figure 3 Verification of an Empirical CD Model that Incorporates Peak Reflectance in addition to Resist Thickness and Development Time

In Ling's metrology, the peak reflectance that occurs between λ = 380 nm and 450 nm is the monitored parameter (Figure 4), since it is relatively insensitive to interference patterns generated by the different indices of refraction of the various films. The peak reflectance depends on three parameters: oxide thickness; resist thickness; and the wavelength at which measurements are made. The weak assumption in Ling's metrology is that it assumes the A and B parameters of the photoresist are relatively constant over various wavelengths, which is not true. In the case of the KTI 820 positive photoresist, B happens to be relatively constant over the range of wavelengths considered, but the A parameter varies significantly with wavelength. Therefore a more correct metrology for measuring

peak reflectance, or more generally, for inferring M will be investigated. Meanwhile however, Ling's metrology will be used.

The peak reflectance is calculated and plotted as a function of photoresist thickness and wavelength λ, while keeping the oxide thickness fixed (Figure 5). From this contour plot, the wavelength at which reflectance peaks, λ_{max}, can determined. And hence peak reflectance is measured.

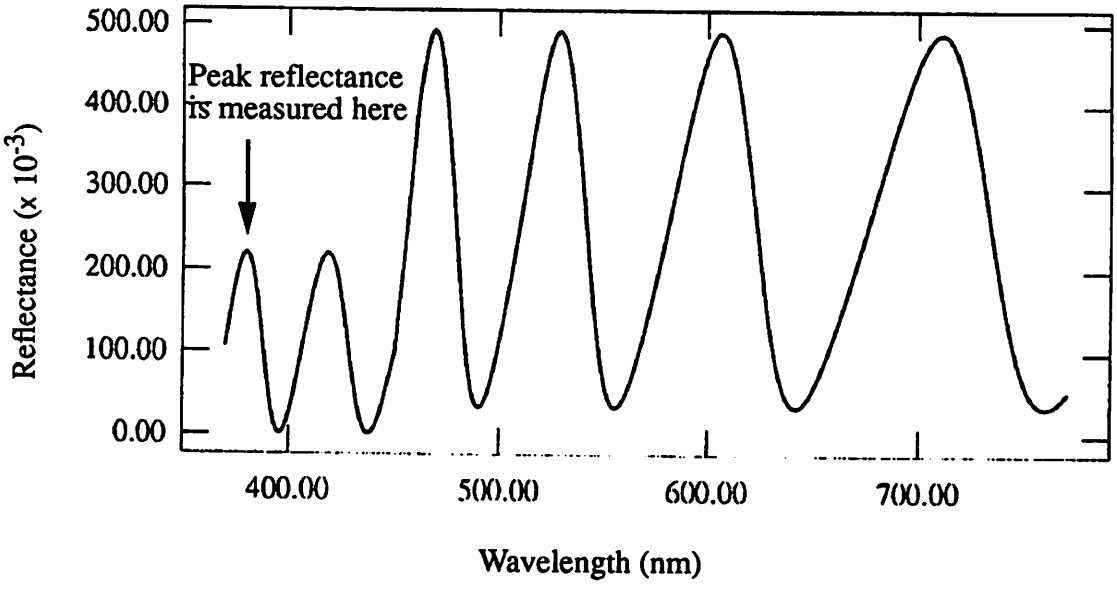


Figure 4 Reflectance of an unexposed Resist-Oxide-Si Wafer vs. Probing Light Wavelength

2.3 Metrology

In summary, the measurement procedure is as follows:

1. Before coating the wafer with photoresist, measure the oxide thickness.
2. Process the wafer and measure the photoresist thickness.
3. Using the contour plot of reflectance vs. λ and resist thickness, that corresponds to the measured oxide thickness, determine λ_{max} and measure the peak reflectance at that wavelength. Then measure the reflectance at two other wavelengths, λ_{max} + 5 nm and λ_{max} - 5 nm.

4. At this point, three reflectance measurements have been made. They correspond to the reflectances measured at the three wavelengths λ_{max} - 5 nm, λ_{max}, λ_{max} + 5 nm. Fit a parabola through those three readings, (λ_{max} - 5 nm, $Refl_1$), (λ_{max}, $Refl_2$), (λ_{max} + 5 nm, $Refl_3$), and determine the maximum reflectance of the air-photoresist-oxide-Si system from the maximum of the parabola.

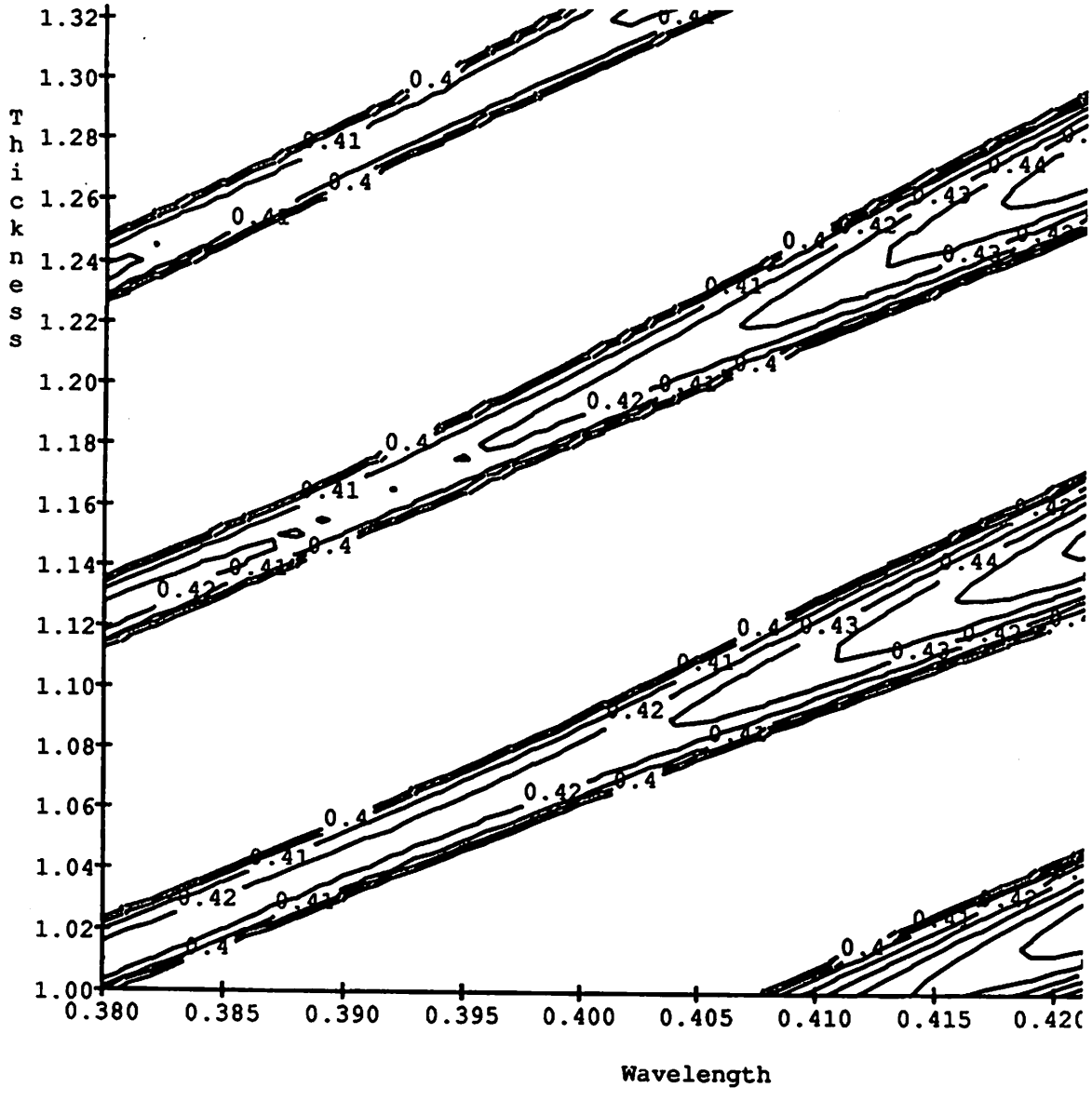


Figure 5 Contour plot of photoresist reflectance vs. photoresist thickness and probing light wavelength, with a fixed oxide thickness of 1190 Å.

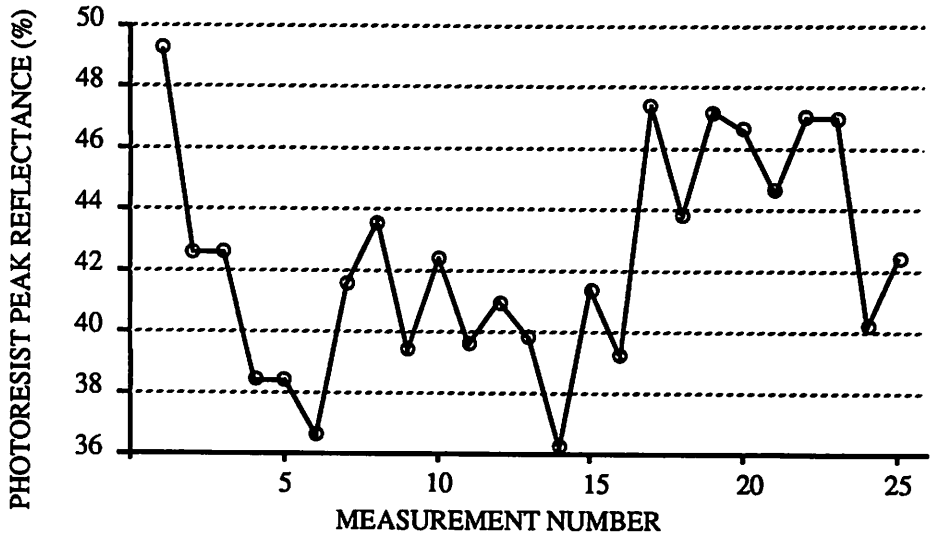


Figure 6 Peak Reflectance using Repetitive Probing. (Pulse Duration = 0.1 second)

A potential problem of Ling's metrology is that the measurements expose the photoresist, since peak reflectance measurements are made at wavelengths where the resist is sensitive. This raises the question of the optimum probing duration. On one hand, it should be made as short as possible, so that the photoresist will be minimally exposed. Experimental results show, however, that very short probing times result in poor replication. This is clearly demonstrated by Figure 6, where the probing time was 0.1 second.

If, on the other hand, probing times are too long, the measurements would be consistent, but the photoresist would get exposed. Figure 7 demonstrates this with repeated probing with pulses of 1.0 second.

Therefore, the objective should be to get measurements as reliably as possible, without exposing the resist too much. We have determined experimentally that by setting the probing time to 1.0 second, the photoresist does not get exposed significantly if less than four measurements are made. Yet the measurements will be quite consistent and reliable to

within ±1%. For the case of photoresist on silicon, we found that 0.7 second was sufficient, since the measurements are less noisy (±0.7%.)

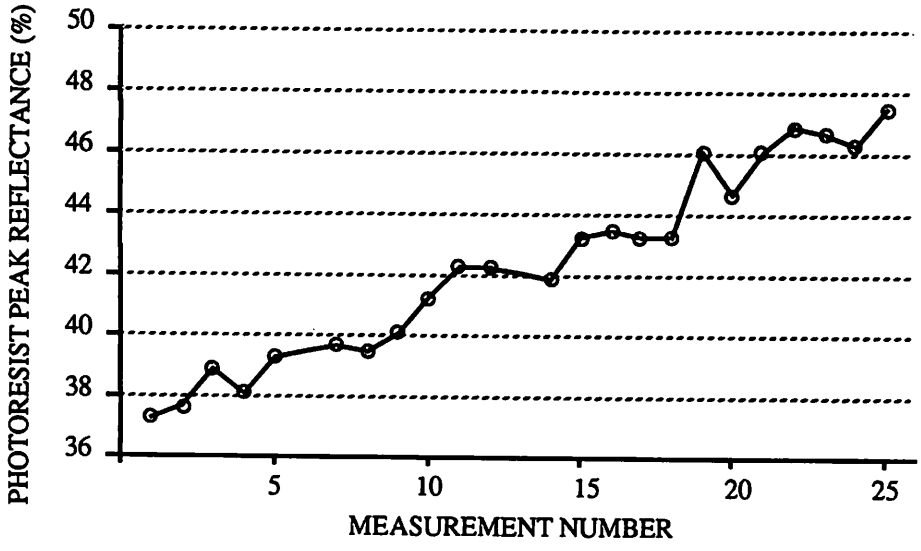


Figure 7 Peak Reflectance using Repetitive Probing (Pulse Duration = 1.0 sec).

2.4 Choosing the Range of the Probing Wavelength

The photospectrometer in the Berkeley Microfabrication Laboratory does not return any reflectance reading for wavelengths below 365 nm. Above that threshold wavelength, there is a region between 365 nm and 380 nm, where the reflectance reading is reliable only within ±5 %. Above 380 nm, the measurements become more stable. Note that this limitation is purely related to this equipment.

Since we are developing a response surface model for the change in peak reflectance after exposure, the resist must have different peak reflectances before and after exposure. For the KTI 820 photoresist, the absorbance, and therefore the peak reflectance, has such property only between λ = 306 nm and λ = 450 nm (Figure 8). Above 450 nm, the absorbance is not a function of exposure. Therefore, we cannot use M, and thus peak reflectance, beyond 450 nm to characterize the chemical status of the photoresist.Taking the

intersection of both regions, the range of useful probing wavelengths is therefore between 380 nm and 450 nm.

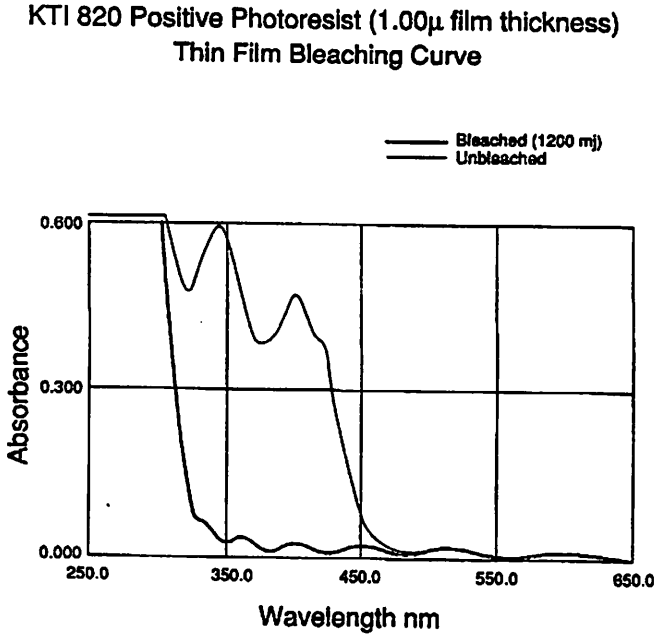


Figure 8 KTI 820 Bleaching Curve [10]

2.5 Summary

We have shown in this chapter that resist thickness and development time measurements complemented with peak reflectance measurements model CDs more accurately than resist thickness and development time measurements alone. Although the peak reflectance metrology is based on weak assumptions, it provides additional information on the status of the photoresist, not available through simple resist thickness measurements. A detailed analysis of the metrology has been done and a procedure to collect data reliably and consistently has been developed. Therefore, while an improved reflectance metrology is being investigated, the present metrology will be used by the model generation system, which will be described next.

Chapter 3 Model Generation System

3.1 Introduction

Reliable equipment models are a critical component of the supervisory control system. They help to determine when a process is in or out of statistical control, and help generate feedback and feed-forward recipe corrections. Ideally physical models should be used since they are valid over a wide process range; but since their outputs are not easily measurable parameters and since their accuracy around the operating point is inadequate for manufacturing purposes, they are replaced by empirical equipment models. Numerous approaches can be used to create reliable empirical equipment models, most notably Response Surface Modeling (RSM) [11][12]. These approaches are rigorous and well understood.

3.2 Experimental Design

To develop a statistically designed experiment, we must decide which process inputs and responses should be monitored (Figure 9). The input factors are easy to choose: they are the control settings of the machines. So for the spin-coat and bake machine, the input factors are the spin speed of the chuck, the spin time, the soft-bake plate temperature, and the soft-bake time, while the responses are photoresist thickness and peak reflectance. The choice of these responses and the metrology have been discussed in chapter 2.

For the stepper, the input factors are the two outputs of the spin-coat and bake machine, the exposure dose, which is regulated by the exposure time, (the light intensity is fixed by the lamp) and the focus. The response of the stepper is the change in peak reflectance, i.e, the difference between the resist reflectance after exposure and before exposure.The difference in peak reflectance is also a measure of the intensity of the latent image.

For the developer, the input factors are the resist peak reflectance after exposure, the photoresist thickness and the development time. The response of the developer is the CD of the photoresist patterns.

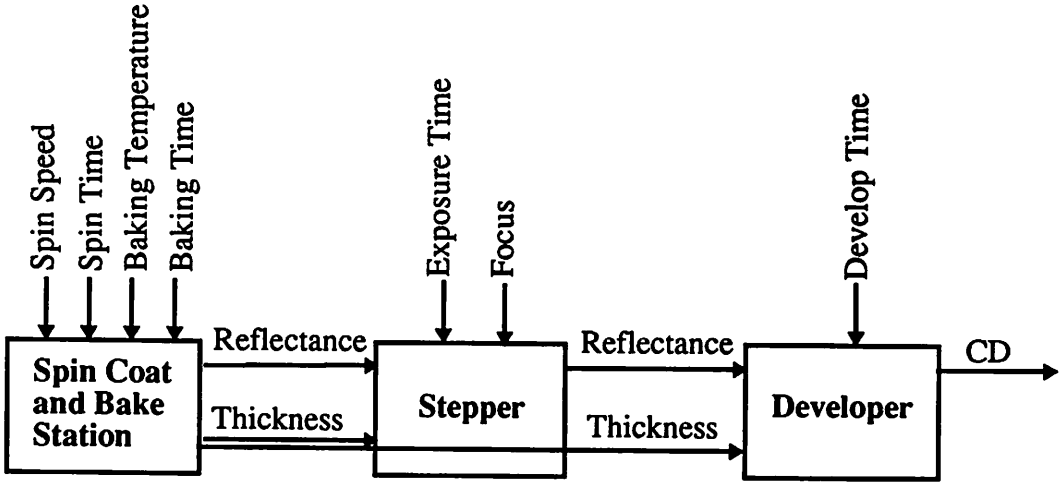


Figure 9 Inputs and Outputs of the Photolithographic Supervisory Controller

3.2.1 Empirical Model Generation for the Spin-Coat and Bake Equipment

The wafer processing system has four inputs and two outputs as described previously. Therefore, to develop a full factorial experiment, 16 experiments are needed. These must be augmented by additional replicated experiments at the operating point. The experimental settings are shown in TABLE 1.

Input Factors	Lower Setting (-)	Standard Setting (0)	Higher Setting (+)
Spin Speed	3600 RPM	4600 RPM	5600 RPM
Spin Time	15 sec	30 sec	90 sec
Soft-Bake Temperature	105 °C	120 °C	135 °C
Soft-Bake Time	20 sec	60 sec	100 sec

TABLE 1 Range of modeling experiment for the spin-coat and bake equipment.

The runs are summarized in TABLE 2. The actual order of execution has been randomized to avoid any blocking effects [11], which occur when similar experiments are executed within a short time span. This causes the results to be more homogeneous than

they would have been, had they been executed further apart in time, and therefore causes the actual equipment replication error to be underestimated.

Experiment #	Spin Speed	Spin Time	Baking Temp	Baking Time
1	−	−	−	−
2	+	−	−	−
3	−	+	−	−
4	+	+	−	−
5	−	−	+	−
6	+	−	+	−
7	−	+	+	−
8	+	+	+	−
9	−	−	−	+
10	+	−	−	+
11	−	+	−	+
12	+	+	−	+
13	−	−	+	+
14	+	−	+	+
15	−	+	+	+
16	+	+	+	+
17	0	0	0	0
18	0	0	0	0
19	0	0	0	0
20	0	0	0	0

TABLE 2 Full-factorial experimental design for the spin-coat and bake equipment

These experimental designs provide a framework from which linear models, with interaction terms, can be developed for the photoresist thickness and peak reflectance. The following peak reflectance model has been developed using step-wise regression [13]:

$$R = 134.4 - 0.046SPS + 0.32SPT - 0.17BTE + 0.023BTI \quad (1)$$
$$-4.34 \cdot 10^{-5}(SPS \cdot SPT) + (5.19 \cdot 10^{-5})(SPS \cdot BTE) - (1.07 \cdot 10^{-3})(SPT \cdot BTE)$$
$$- (4.11 \cdot 10^{-4})(SPT \cdot BTI) + (5.15 \cdot 10^{-6})(SPS)^2$$

where R is the resist reflectance in %, SPS is the spin speed in RPM, SPT is the spin time in seconds, BTE is the soft-baking temperature in ^{o}C, and BTI is the baking time in sec-

onds. The peak reflectance model predicts the mean response of the spin-coat and bake equipment with a one-sigma prediction error of ±0.6 %. The actual measured response of the equipment will vary around the mean value with a one-sigma replication error of ±1.96 %.

The Analysis of Variance (ANOVA) table [11] for the reflectance model is shown in TABLE 3. A measure of the overall variation of the data can be obtained by the sum of squares of the data. A property of the sum of squares is that it is additive, i.e, the sum of squares about the model and the sum of squares about the residuals add up to the total sum of squares of the data. Therefore, since the sum of squares about the model is almost as big as that about the total data, the model accounts for most of the data's variations. The residual sum of squares corresponds to the part that is unaccountable by the model and is therefore attributed to experimental error. Among the residuals, we can further distinguish two types of error, if replicate runs have been done. The sum of squares attributed to the "lack of fit" corresponds to the discrepancies among results obtained from replicated runs. That error corresponds to the lower limit of the total error, and corresponds to the inherent variation of the process. The rest of the residuals corresponds to all other causes of error, such as, for example, operator error.

Another important criterion found in an ANOVA table is the F-ratio of the model, which is the ratio of the mean square of the variation explained by model over the mean square of the residuals. The mean square of a term is equal to the sum of squares about that term divided by the degrees of freedom used by that term. Assuming the residuals are random and normally distributed around zero, if the model was not significant, i.e, if it does not represent the data well, the F-ratio of the model would follow an F distribution with 9 and 6 degrees of freedom. These corresponds to the degrees of freedom used by the model and by the residuals, respectively (Figure 10). The area under this curve, to the right of the observed ratio 10.08, is the significance level for the hypothesis that the model is not sig-

nificant. Since the significance level of the hypothesis that the model is insignificant is a mere 0.5%, we conclude that the model is indeed significant.

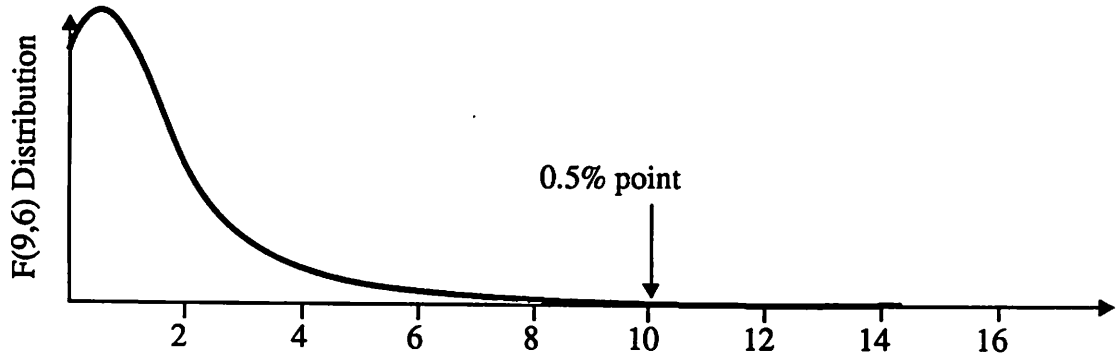


Figure 10 Observed value of an F-ratio of 10.08 in relation to an F distribution with 9 and 6 degrees of freedom. (for Peak Reflectance Model)

Finally, the adjusted R^2 is another parameter which measures the fraction of the total variation in the data accounted for by the model. The closer the adjusted R^2 is to 1, the better the model fits the data. Since the adjusted R^2 obtained for the peak reflectance model is 0.897, it only confirms furthermore that our peak reflectance model is significant. The adjusted R^2 is a variant of the typical R^2, in the sense that it takes into account how many degrees of freedom have been used by the model. The fewer degrees of freedom used by the model, while keeping the model significant, the higher the adjusted R^2. A scatterplot of the resist reflectance values predicted by the model versus the corresponding experimental values is given in Figure 11. Here, the closer the experimental data are to the y=x line, the more accurate the model prediction is.

Source	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Significance
Total	15	3723010	248201		
Model	9	3491988	387999	10.08	0.005
Residual	6	231022	38504		
Lack of Fit	4	196564	49141		
Error	2	34458	17229		

Adjusted $R^2 = 0.897$

TABLE 3 ANOVA table for the reflectance model of the spin-coat and bake equipment.

The thickness model has also been developed using linear step-wise regression at first, but the fit was very poor. In order to improve the fit, we used linear regression on non-linear transformations of the input parameters. These non-linear transformations were derived from residual analysis of each input parameter. And the model we arrived at is:

$$T = -13814 + \frac{2.54 \cdot 10^6}{\sqrt{SPS}} + \frac{1.95 \cdot 10^7}{BTE\sqrt{SPS}} - 3.78BTI - 0.28SPT - \frac{6.16 \cdot 10^7}{SPS} \quad (2)$$

This model predicts the mean response of the spin-coat and bake machine with a one-sigma prediction error of ±30Å. The actual measured response of the equipment will vary around the mean value with a one sigma replication error of ±68Å. The ANOVA table for the thickness model is shown in TABLE 4. The F-test shows that this model is highly significant, since the probability that F(5,13) > 1466 is extremely low (2e-17) [13]. A scatter-plot of the resist thickness values predicted by the model versus the corresponding experimental values is given in Figure 11.

Source	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Significance
Total	18	34016040	1889780		
Regression	5	33955810	6791162	1466.00	0.000
Residual	13	60231	4633		
Lack of Fit	11	59642	5422		
Error	2	589	294		

Adjusted $R^2 = 0.998$

TABLE 4 ANOVA table for the thickness model of the spin-coat and bake equipment.

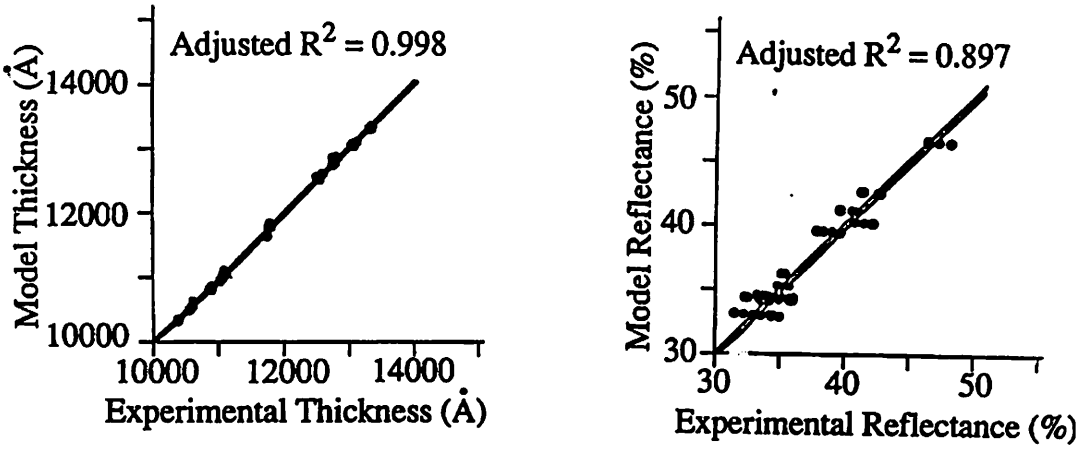


Figure 11 Thickness and Reflectance Model Verifications for the Spin Coat and Bake Station

3.2.2 Stepper

The stepper has four inputs and one output. As noted earlier, the stepper affects only the peak reflectance of the photoresist, not the thickness. The four input factors are: the resist thickness and peak reflectance at the output of the spin-coat and bake machine, the focus and the exposure time. The last two inputs can be directly controlled on the stepper. Therefore, to develop a full factorial experiment, 16 experiments are needed, in addition to a few replicated runs at the normal operating point. The experimental settings are shown in TABLE 5.

Input Factors	Low Setting (-)	Standard Setting (0)	High Setting (+)
Input Reflectance	36 %	40 %	44 %
Input Thickness	11500 Ang	12300 Ang	13200 Ang
Exposure Time (Dose)	0.67 sec	0.77 sec	0.87 sec
Defocus	-8 μm	0	+8 μm

TABLE 5 Experimental Range for the Stepper.

The experimental runs have also been randomly executed to avoid time dependent blocking effects. The stepper model has been developed using step-wise regression [13]. Notice that the defocus did not have any significant effect on the incremental reflectance within our experimental range.

$$\Delta\text{Rout} = 4808 - 850\text{Rin} - 945\text{Tin} + 767\text{Din} + 1134\,(\text{Din} \cdot \text{Tin}) \qquad (13)$$

This model predicts the change in peak reflectance after exposure with a one sigma prediction error of ±0.3%. The actual measured response of the equipment will vary around the mean value with a one-sigma replication error of ±1.08 %. The ANOVA table for the stepper model is shown in TABLE 6. The F-test shows that this model is highly significant. A scatterplot of the values predicted by the model versus the corresponding experimental values is given in Figure 12.

Source	DF	Sum of Squares	Mean Square	F-Ratio	Significance
Total	18	3371107			
Regression	4	3207582	801895	68.65	0.000
Residual	14	163525	11680		

Adjusted $R^2 = 0.938$

TABLE 6 ANOVA table for the reflectance model of the stepper.

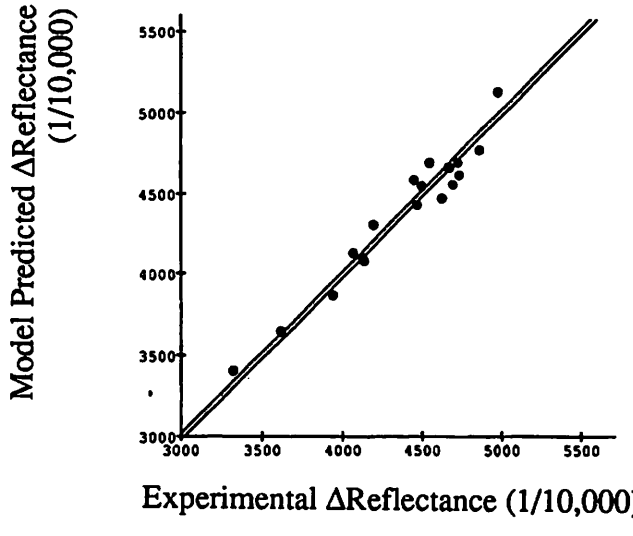


Figure 12 Verification of the stepper model.

3.2.3 Developer

The developer has three inputs and one output. The three input factors are: the resist thickness after spin-coat and bake, the resist peak reflectance after exposure and the development time. Only the last one is directly controllable on the developer. The output is the CD of the photoresist pattern. The slope of the developed patterns is potentially another output parameter, but it is unfortunately not easily measurable. Therefore, to develop a full factorial set, we must run 8 experiments, and a few more experiments at the operating point. The range of this experiment is shown in TABLE 7.

Input Factors	Low Setting (-)	Standard Setting (0)	High Setting (+)
Input Reflectance	75 %	84 %	89 %
Input Thickness	11500 Å	12300 Å	13200 Å
Development Time	40 sec	60 sec	90 sec

TABLE 7 Modeling experiment for the developer.

The 8 runs have been randomized to avoid blocking effects. The CD model has been developed using step-wise regression [13], and the model is:

$$CD = 8.86 - 0.076Rin - 0.00075Tin + 0.00000375 \cdot Tin \cdot Devtime$$
$$+ 8.95 \cdot 10^{-6} (Rin \cdot Tin) - (4.27 \cdot 10^{-8}) (Tin \cdot Rin \cdot Devtime)$$

This CD model predicts the mean response of the stepper with a one-sigma prediction error of ±0.025 μm. The actual measured response of the equipment will vary around the mean value with a one sigma replication error of ±0.025 μm. The ANOVA table for the critical dimension model is shown in TABLE 8. A scatterplot of the resist critical dimension values predicted by the model versus the corresponding experimental values is given in Figure 13.

Source	DF	Sum of Squares	Mean Square	F-Ratio	Significance
Total	10	0.063753			
Regression	5	0.060517	0.012103	18.70	0.003
Residual	5	0.003237	0.000647		

Adjusted $R^2 = 0.949$

TABLE 8 ANOVA table for the developer model.

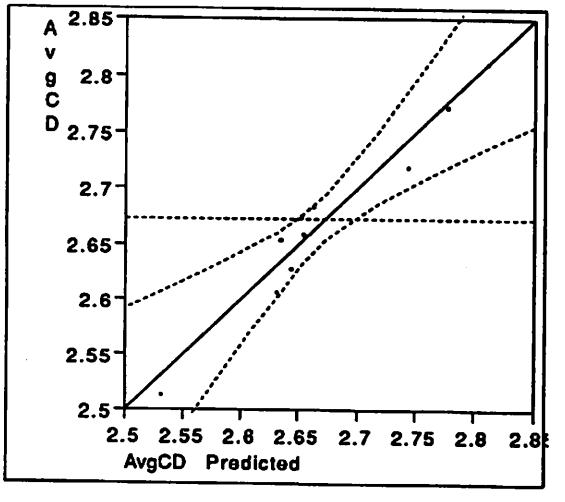


Figure 13 Verification of the developer model.

3.3 Summary

Reliable empirical response surface models have been developed for all three photolithography machines. They have been developed using step-wise regression and are fully characterized through statistical analysis. They have been proven to represent the whole lithography sequence quite accurately, and are robust enough to serve our supervisory control system. This system is described next.

Chapter 4 Feedback Control System

4.1 Introduction and Background

A control system can improve the reliability and the accuracy of equipment without significantly increasing its cost; furthermore, by allowing the same equipment to be used for tasks that need higher accuracy, the control system could also postpone costly hardware upgrades. Such a system would include feedback and feed-forward control. This chapter describes a feedback control algorithm and the next one describes a feed-forward control algorithm. Although both control algorithms are generic and can therefore be applied to any semiconductor manufacturing sequence, both have been applied in this case to a lithography sequence.

The primary task of the feedback control mechanism is to adjust the process to bring the output response back on target. In addition, the feedback control mechanism must be able to foretell upcoming needs for preventive maintenance, and to correctly identify uncontrollable situations that require human intervention. Although heuristic algorithms for control have been reported in the semiconductor industry [15], we have chosen to base our approach on formal statistical methods. Statistically based algorithms offer several advantages over heuristic approaches, since they can be adapted to a large number of processes and, once in place, are robust enough to be useful in an actual manufacturing environment.

Several attempts have already been made to formalize this procedure, including the MIT run-by-run Controller [16] and Ultramax™ [17], a commercial package for sequential process optimization and process control.

The current version of the MIT Run-by-Run Controller allows the control of one process parameter by adaptively changing the constant term of a linear process model [16],

although recent reports talk about more general model adaptation and multivariate applications. Ultramax is a commercially available software [17] that can handle multiple inputs and outputs. Although the details of its operation are proprietary, Ultramax uses a variant of the evolutionary operation algorithm (EVOP) [12] to find the optimum operating point. Ultramax offers the significant advantage that no prior model of the process is required; however, it does require continuous changes on the process in order to derive such a model. The control algorithm presented here, on the other hand, has the additional advantages that it offers multivariate control and complex adaptation of non-linear models, without introducing extraneous variance in the process.

The goal is to improve the overall capability of photolithographic pattern transfer by applying feedback control on the photolithographic sequence. Towards that objective, our general control methodology consists of three components: the monitoring of several critical process parameters, as discussed in chapter 2; the creation of adaptive response surface models, as discussed in chapter 3; and finally the detection and compensation of any systematic process disturbances, which will be discussed below.

4.2 Detection of Process Disturbances

Disturbances can be classified into two main types. The first one manifests itself through sudden significant changes in the process output. This clearly indicates the presence of a problem that needs to be corrected by an operator. This type of disturbance triggers what we call *malfunction alarms*. The second type of disturbance manifests itself as a systematic process drift, which can be corrected by an appropriate recipe change. This type of disturbance triggers what we call *control alarms*. The two types of alarm are described next.

4.2.1 Malfunction Alarms

Malfunction alarms are needed to identify conditions which require operator attention. These are cases where the variation of a monitored parameter increases, or when we encounter sudden changes that are not consistent enough to be compensated by recipe adjustments. Finally, a malfunction alarm is also generated if the change cannot be compensated unless one (or more) of the controlling parameters moves beyond its acceptable range.

These conditions can be identified with the application of a special SPC scheme that can accommodate multiple parameters (as several process parameters are being monitored). This scheme must be able to ignore intentional changes in equipment settings such as those that might occur due to control algorithms. Such a SPC scheme has been developed using a combination of the Regression Chart [22] and Hotelling's T^2 statistic [23].

Using this scheme, malfunction alarms are generated in two stages: first, we use the response surface models to predict the new measurements. Then, we plot the difference between the reading and the model prediction. When the process is under statistical control, this difference is a random number with a known mean and variance. This variance is calculated using the prediction error of the model, as well as the observed variation of the equipment. The method is described for univariate regression models in [22] and it has been generalized for multivariable response surface models in [24].

After the prediction errors and their variances and covariances are calculated, the multiple responses are merged together using the T^2 statistic. This statistic is plotted on a single-sided control chart whose upper control limit (UCL) can be formally set at the desired probability of erroneously stopping a good process, by using the x^2 distribution. If this limit is exceeded, the automated control system stops and a human operator investigates

the malfunction, the same way the operator would have investigated a traditional SPC out of control condition (Figure 14).

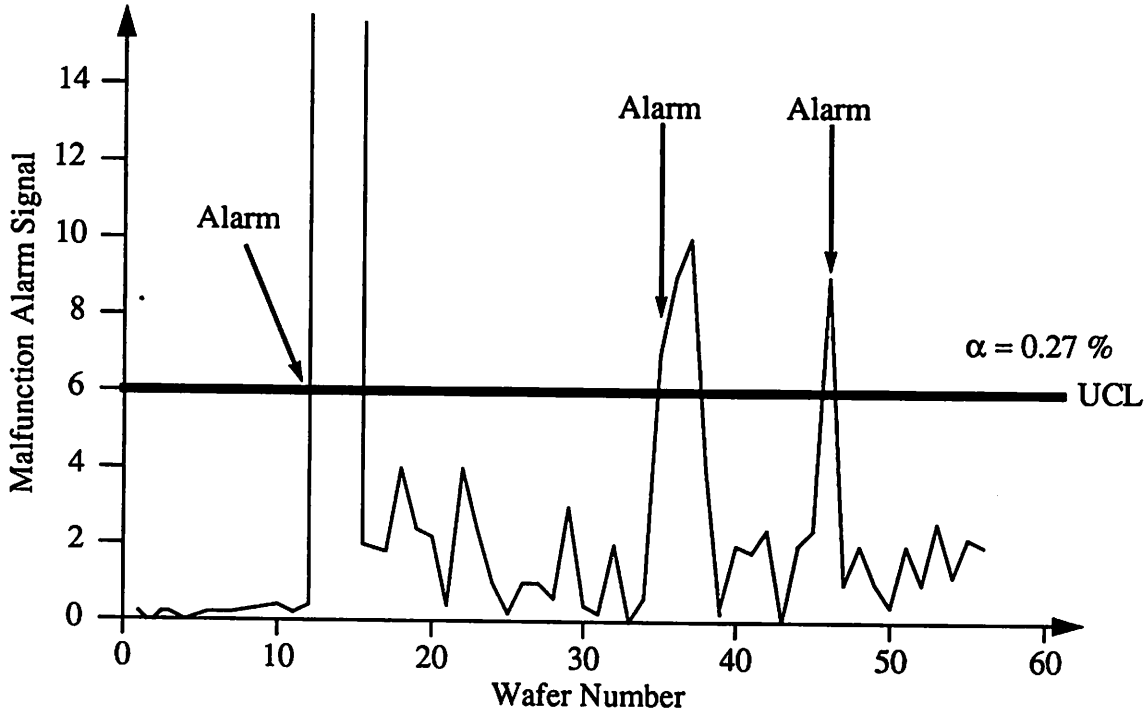


Figure 14 Malfunction alarm generation.

4.2.2 Control Alarms

Control alarms are needed to identify process drifts that can be compensated by appropriate recipe changes. These drifts and disturbances are detected by means of a multivariate cumulative sum (CUSUM) scheme that is very efficient at identifying small, consistent changes, while ignoring outliers that are not useful for feedback corrections.

This alarm generation is based on the multivariate CUSUM scheme described by Crosier [25]. Crosier's scheme forms a CUSUM vector directly from the experimental data. This vector is then used in conjunction with the baseline covariance matrix which is calculated from historical data obtained when the equipment is known to be in control, to yield an alarm indicator. The sensitivity of the alarm depends purely on the upper control limit of that alarm indicator, and can be adjusted for the desired probability of stopping erroneously a good process.

As in any CUSUM scheme, the multivariate scheme discussed here is based on comparing the parameter readings against a target. However, since a process might not be able to return to its ideal target, a comparison between the experimental results against the ideal targets would generate control alarms too often. The comparison of the experimental data to the model predictions, on the other hand, will properly generate an alarm only if the updated models do not represent the experimental data well. Since this is exactly what is desired, the control alarm is then set off only when the model inadequately represents the data. Figure 15 depicts the generation of control alarms during the feedback experiment that will be described in chapter 7.

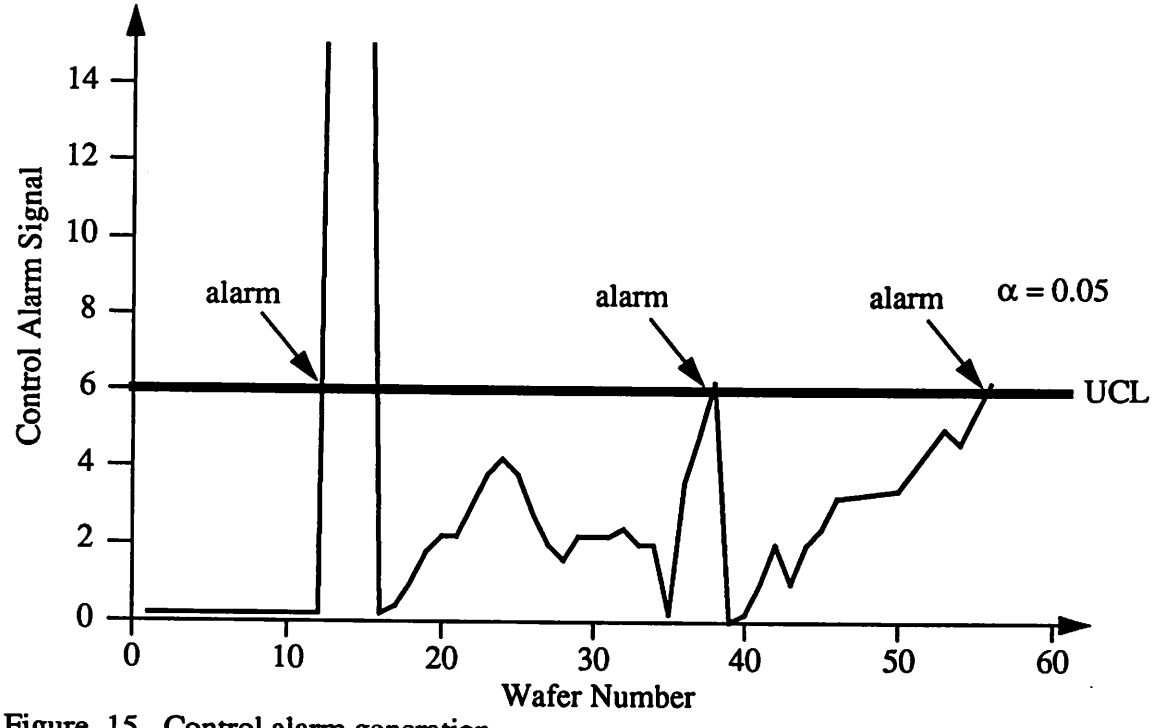


Figure 15 Control alarm generation.

4.3 Algorithm for Adaptively Updating Equipment Models

Although the original RSM models obtained from off-line factorial experiments offer a comprehensive representation of the process, they are unable to describe the process over a period of time, since equipment age, components are replaced, and incoming materials change. Thus, it is important to be able to adapt these models to the current status of

the process. In our control methodology, adaptation occurs every time a control alarm is issued, since that is when the original RSM model does not represent the equipment accurately anymore. Since the additional data that we have at our disposal do not come from an off-line controlled experiment, the method used for this update is important.

Since the machine is continuously changing over time, the model is updated according to an exponential weighting scheme that attributes more importance to recent *groups of observations* and less importance to older ones. For this weighting scheme, a *group of observations* is defined as a set of data that has been obtained under the same recipe; in other words, *groups of observations* are delimited by control alarms.

So, once a control alarm is generated, the model updating scheme calculates the deviations of the experimental results from the original RSM model, and generates a correction model using linear regression to model these deviations. Then, this correction model is added onto the original model to generate the updated model. Mathematically, this algorithm is summarized as follows:

$$f_{updated}(\underline{x}) = f_{original}(\underline{x}) + \Delta f_{corr}(\underline{x}) \quad (1)$$

where $\underline{x}$ is the vector of the process settings, $f_{updated}$ is the updated model and thus currently the most accurate model, $f_{original}$ is the model obtained from the original off-line factorial design of experiments, and $\Delta f_{corr}(\underline{x})$ is the correction model. Subsequently, the current model will be adaptively updated to reflect the new status of the process.

The correction model is supported by the additional experimental data obtained during the normal operation of the process. To make the new model adapt to the more recent state of the process, we fit the correction model while incorporating a forgetting factor that gives more weight to the recent *groups of observations*. The correction model is obtained by a weighted least square minimization process, summarized by the following equation:

$$\min \sum_{i=0}^{n} w(i) \left[\Delta f_{corr}(\underline{x}_i) - (\underline{y}_i - f_{original}(\underline{x}_i))\right]^2 \quad (2)$$

where w(i) corresponds to the weight of the *i*-th group of data, *n* is the total number of *groups of observations*, $\underline{x}_i$ corresponds to the input vector of the *i*-th group, $\underline{y}_i$ to the output vector of the i-th group of data, and $f_{original}$ and Δf_{corr} are defined previously. The weighting factor w(i) is set to be w^i, with the value for w chosen between 0 and 1.

Whether w(i) should be kept constant or varying has not been investigated yet, and is part of future work. If the process is susceptible to slow, continuous drifts, old batches of data are still quite relevant, and the appropriate value of w is chosen very close to 1. If, on the other hand, the process is susceptible to abrupt step changes, old data are not very relevant, and a small value of w should be chosen.

This model updating scheme allows updating any or all model coefficients, if new coefficients can be supported by new *groups of observations*. More specifically, immediately after the first *group of observations* is collected, the correction model consists of just a constant term. After the recipe is changed and the second *group of observations* is collected, the two data sets can support a maximum of two terms. So now, the constant term plus one of the first order terms can be evaluated in the correction model. With more recipes and more *groups of observations*, more complex correction models can be supported, the highest complexity being limited by the maximum number of "relevant" *groups of observations* given the forgetting scheme described above. In order to automate the generation of the optimal correction model, we use step-wise regression that is set to select the model with the highest adjusted R^2 whose terms are significant at the 5% level [12].

4.3.1 Automated Recipe Generation

Once the original response surface model has been updated with the addition of a correction model, a new recipe is generated to bring all process responses back on target. However sometimes, this is not possible because some of these output responses are corre-

lated. Therefore, a compromise recipe that brings each of the responses as close as possible back on target must be reached. Such a compromise recipe is obtained by minimizing the deviations of the model predicted responses from their targets. The optimum equipment setting is found by minimizing that objective function while observing the constraints that limit each of the process settings to a certain range. For small problems, this minimization problem can be solved using an exhaustive search. Larger problems can be solved with the application of a constraint minimization method [28].

4.4 Summary

In this chapter, we have presented an algorithm for the implementation of feedback control on a semiconductor manufacturing step (Figure 16). The feedback control algorithm is based on the formal generation of malfunction and control alarms, and on an adaptive model updating strategy. This algorithm has been implemented on various machines in the Berkeley Microfabrication Laboratory and experimental results, presented in chapter 7, show that process responses can be significantly improved by applying this control algorithm.

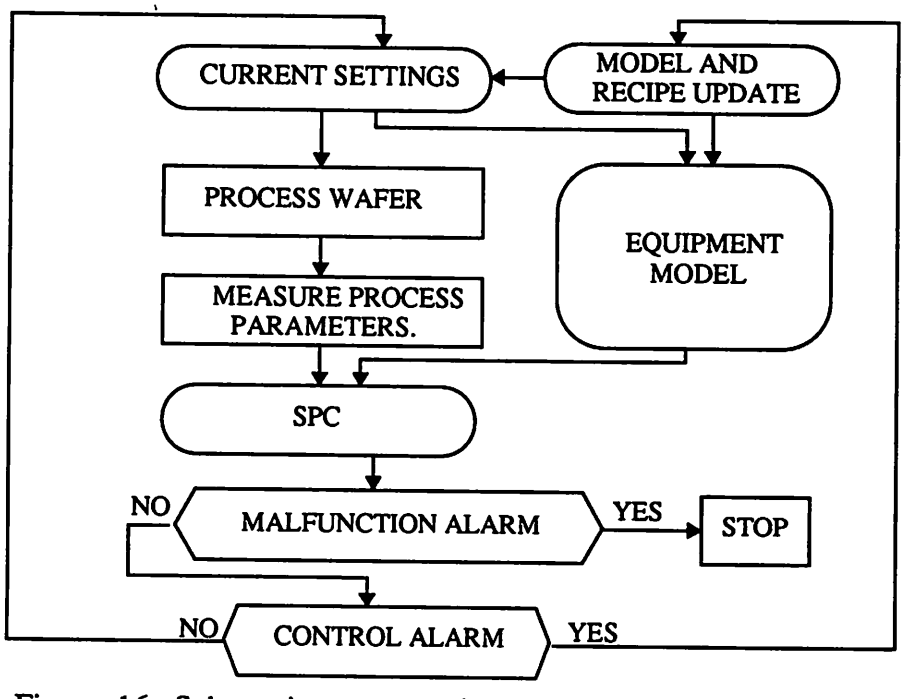


Figure 16 Schematic representation of the feedback procedure.

Chapter 5 Feed-Forward Control System

5.1 Introduction

The feed-forward controller complements the feedback controller which centers the process mean back on target, by reducing the process variability. It achieves this task by monitoring and compensating for the variability that wafers inherit from past process steps. Relying on accurate equipment models, the feed-forward controller implements the corrective action by generating a new recipe for a downstream process step, so that the final properties of the wafer are within specifications.

5.2 Feed-Forward Control Alarm Generation

Before applying feed-forward control on an equipment, a criterion that decides whether feed-forward correction is possible must be developed [1]. If the corrected output is predicted to be within process specifications, then feed-forward control is activated; if not, the photoresist is stripped and the wafer is sent back to the wafer track to be coated again.

A process target and its tolerance are defined by a lower specification limit (LSL) and an upper specification limit (USL). The following statistical test is then used to decide whether or not to process the wafer with the standard recipe:

$$LSL + Z_\delta \sigma < \hat{Y} \pm 3\sigma_{pred} < USL - Z_\delta \sigma \qquad (3)$$

where $\hat{Y}$ is the predicted output according to the equipment model, σ_{pred} is the prediction error of the empirical equipment model, δ is the acceptable yield loss, σ is the replication error of the controlled equipment, and Z_δ is the probability of stopping a good process, given an accepted maximum yield loss. If, with standard recipe settings, the predicted output falls between the limits set in Eq.(3), no feed-forward action is taken. If it falls outside

these control limits, then feed-forward action is activated. The feed-forward compensation, however, is only implemented if the new recipe falls within the range of acceptable recipes. If not, then the controller sends the wafer back to be stripped and re-coated.

5.3 Feed-Forward Control Methodology

Before processing the wafer on an equipment, the feed-forward controller predicts first whether or not the wafer will meet specifications if run at normal settings. If it doesn't, the feed-forward controller looks for the next step that is dependent on the present step, and adjusts the input settings of that machine downstream in the process sequence to correct for the shortcomings of the present equipment that has gone out of specifications. The feed-forward controller therefore treats wafers on a run-by-run basis.

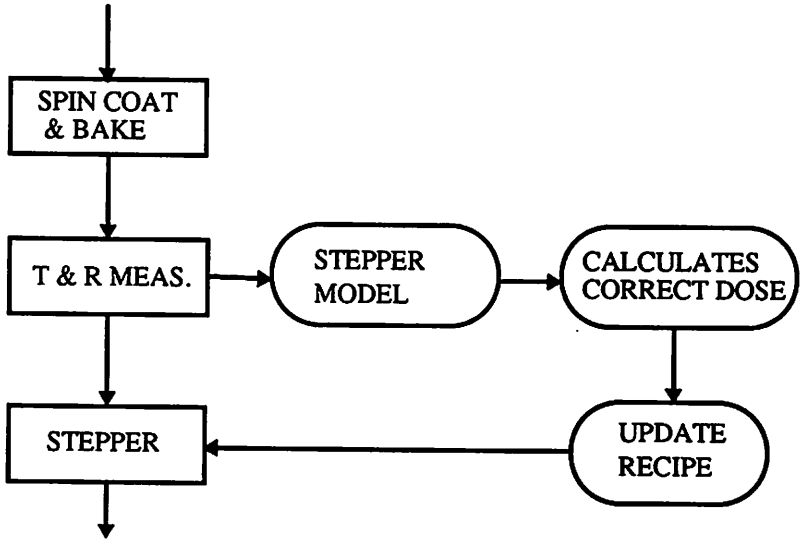


Figure 17 Example of the feed-forward control procedure applied to a stepper.

More specifically, the feed-forward procedure starts by measuring the characteristics of the incoming wafer. The controller then evaluates these measurements and decides whether to activate feed-forward control or not. If it does, a new recipe designed specifically to bring that wafer under control is generated for the process. This is accomplished using equipment models that have been obtained previously either from a designed experiment or from an in-line modeling experiment. Then the wafer is processed under that new

recipe. Once the process is finished, the wafer parameters are inspected again to verify that the equipment model for the present equipment is still accurate. If it is not, feedback control will be activated to update the present equipment model. Note that feedback control plays a crucial role in feed-forward control, since feed-forward control is based on equipment models whose accuracy is solely dependent on feedback control. To function efficiently, both controllers must therefore be tightly coupled to each other.

Applied to the stepper, for example, the feed-forward algorithm works as follows: the thickness and the peak reflectance of the photoresist layer are measured before exposure. At this point the feed-forward controller evaluates the measurements and recommends an exposure recipe. If the measurements show that the wafer parameters were exactly on target, the standard recipe will be used. If, however, the parameters are not on target, then a new recipe will be generated, using the current stepper model. After exposure, the status of the resist is inspected again, and the measurements are compared with the model prediction. If a statistically significant discrepancy is found, we conclude that the model does not represent the process well anymore. This results in a control alarm and the feedback controller updates the stepper model to adapt to the new state of the process.

5.4 Automated Recipe Generation

The feed-forward controller tries to optimize the setting for each individual wafer by generating, if needed, a new recipe for each wafer. This reduces significantly the variability due to all previous process steps. The recipe generation has been totally automated for photolithography machines in the Berkeley Microfabrication Laboratory. For example, after the photoresist has been applied onto the wafer, its parameters are measured and evaluated by the feed-forward controller, using the stepper model. If they lie beyond the control limits, the feed-forward controller solves for the exposure time that would bring the peak reflectance back on target, and the wafer is exposed using this new exposure time.

5.5 Summary

A robust feed-forward control system based on accurate equipment models and SPC has been implemented for the lithography sequence. The control system uses SPC to reduce process variability and to bring the process back on target on a run-by-run basis. The limitations of the feed-forward controller depend on the accuracy of the equipment models which are adaptively updated by the feedback controller. When the feedback controller performs well, the feed-forward controller is highly successful in reducing process variability and in centering the process back on target, as will be shown in chapter 7.

Chapter 6 Automation of the Supervisory Control System

The supervisory control system described in this work is fully automated using the Work-In-Progress management system developed by Dr. Hegarty [29].

After the measurements are extracted using a photospectrometer, they are uploaded into a database via an interface program called Nanotalk [31]. Once inside the database, the measurements are manipulated using the Berkeley Process Flow Language (BPFL) [29]. First, a program extracts output responses from the measurements, screens them for outliers and calculates relevant statistics about them.

Next, the output responses are sent to a statistical package which automatically generates relevant alarms and adaptively updates equipment models using the Berkeley Interactive Statistical System (BLSS) [30]. The model update algorithms use the step-wise regression algorithm in BLSS and return the model that reaches the highest adjusted R^2. Finally, the Han-Powell non-linear optimizer [28] generates new process recipes to bring the process back in control, using these updated equipment models.

In summary, the lithography controller is fully automated using robust public domain software developed at UC Berkeley. Next, we will describe experimental implementations of the control algorithms discussed in this report. These experimental implementations have been performed in the Berkeley Microfabrication Laboratory.

Chapter 7 Experimental Results

7.1 Description of the Experimental Set-Up

7.1.1 Test Pattern Design and Fabrication

To obtain reliable readings, the locations where measurements are taken must be carefully chosen. The photoactive compound concentration (PAC) is generally not dependent on position, but the resist thickness is radially dependent on position. As deposited on our wafer track, the photoresist is thicker in the middle of the wafer and thinner at the edges [14]. Therefore, to obtain repeatable thickness measurements, measurements are taken at positions 1, 2, 3, and 4, since they lie on the same radius; and to make measurements even more repeatable, these exact same locations will be used for all measurements. A test pattern was designed for this purpose and is shown below in .

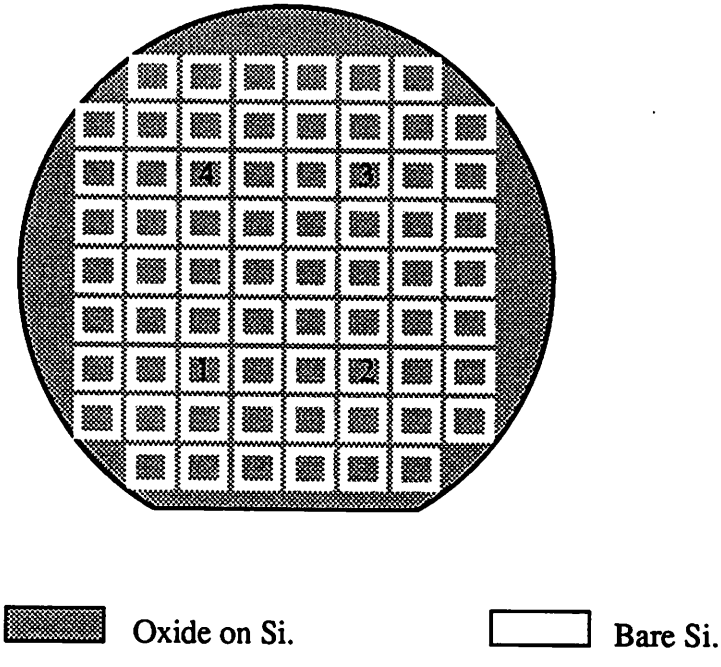


Figure 18 Pattern used for test wafers.

To obtain such a pattern, we performed the following processes. First, we grew 1000 Å of thermal oxide on Si wafers. Our samples included a random mixture of both <111> and <110> wafers. These wafers were spin-coated with a positive photoresist (KTI 820) at 4600 RPM for 30 sec, soft-baked at 120°C for 60 sec, and exposed with a I-line stepper using the mask shown in Figure 19. Then we post-baked the wafers at 120°C for 60 sec, and developed the pattern. This procedure results in a wafer patterned as in .

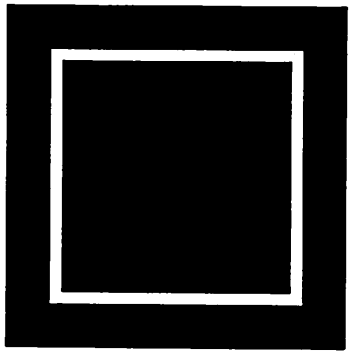


Figure 19 Mask for the wafer test pattern.

Before producing resist patterns in order to measure CDs, all wafers have to be RCA cleaned and dehydrated in a convection oven for 30 mins. The RCA cleaning procedure involves fully immersing the wafers for 10 minutes in a piranha sink, and then rinsing them in DI water for around 15 minutes, until the cleanliness of the DI water exceeds 10 MΩ·cm. Finally, the wafers are dried in a spin-dryer.

Next, we go through the whole lithography sequence to pattern the squares covered with oxide with lines in order to measure CDs. We develop the pattern shown in Figure 20 inside each square. These line patterns contain different sets of linewidths (d = 1 μm, 2 μm, 3 μm, 4 μm, and 5 μm), and each set has three different pitches (Figure 20). One set of line pattern has a pitch 4 times the linewidth, another has a pitch 3 times the linewidth, and another has a pitch twice the linewidth. The choice of using which set of line pattern in our standardized metrology depends on the quality of the photospectrometer and the

capability of the stepper. After investigating all sets of line patterns for maximum measurement reliability, we chose to measure the 3 μm linewidth pattern that has a pitch twice the linewidth.

This concludes the description of the lithography process we intend to control. Next, we will describe the equipment used in each process.

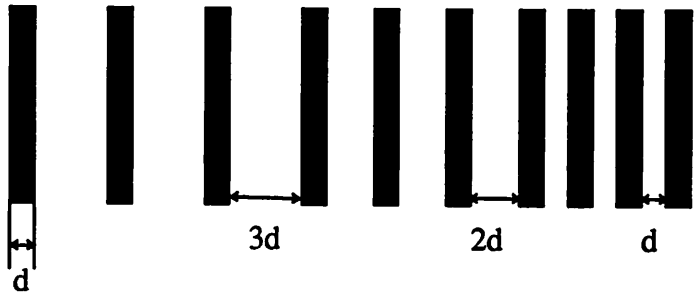


Figure 20 Mask for measuring CDs.

7.1.2 The Spin-Coat and Bake Equipment

The first equipment in the lithography sequence is the EATON LSI 45/60 Wafer Processing System and it is designed to spin-coat and bake 4" wafers. It consists of a chuck that can spin wafers at different speeds, a photoresist dispenser, a hot baking plate, and a cold plate. The duration of each step can be controlled in increments of one second by an internal computer. The wafers are loaded into the equipment in a cassette containing up to 24 wafers and are processed individually one after another.

During the coating operation, the wafer is held to the chuck by a vacuum. The spin speed of the chuck can be set from 0 to a maximum of 7000 RPM, in increments of 100 RPM. The chuck actual spin speed is within 20 RPM of the set value. The photoresist dispenser can be moved so that it can dispense the resist at any radial position on the wafer. Throughout our experiment, we dispense the resist at the center of the wafer. The hot bake plate can be programmed in increments of $1^{o}C$, but the actual temperature is kept within $\pm 2\ ^{o}C$ of the set value. This limited temperature control turned out to be a major

cause of experimental error, since the PAC concentration, and therefore the peak reflectance, is very sensitive to temperature. The thickness is also sensitive to temperature variation, because the solvent evaporation rate depends on temperature, but its sensitivity is less severe than that of reflectance. The cold plate is set at room temperature which varies between 20°C and 27°C. It has been verified that the cold plate temperature variation does not affect the measurements.

7.1.3 The Stepper

The next equipment in line is the 10X reduction GCA stepper, model 6200. It is also configured to handle 4" wafers, which are loaded in batch mode. The wavelength of the light source is 365 nm (I-line). The 6200 is a fully automated stepper that handles wafers one at a time. The stepper has an embedded controller that keeps the dose constant throughout the life of the lamp. The inputs to the stepper are focus and exposure time, which controls the input dose. The focus is measured in μm and is adjusted in increments of 0.5 μm. The exposure time is measured in seconds, and can be controlled to within ±0.01 second.

7.1.4 The Developer

After the wafer is exposed, it is post-baked on the wafer track and then developed by the MTI Omnichuck Photoresist Development Station. Like the other two machines, it is also made to handle 4" wafers and processes one wafer at a time, with the wafers loaded in batch mode. Up to three liquids can be used in the wet process. The MTI Omnichuck can therefore serve as a developing station as well as a photoresist stripping station. The Omnichuck is linked to a computer where its recipe programs are stored. Although many parameters can be changed such as the spin speed of the chuck, only the development time has been used in our control scheme. The development time can be controlled in increments of one second.

Now that the lithography machines have been described, the machines used in our monitoring scheme will be discussed.

7.1.5 The Photospectrometer for Reflectance and Thickness Measurement

The photospectrometer used for collecting reflectance and thickness measurements is the 210 Nanospec/AFT Film Thickness System, made by Nanometrics. The photospectrometer is quite accurate when it calculates the resist thickness from reflected interference patterns. Used for reflectance readings however, its precision is heavily dependent on the scanning speed and on the number of transparent layers on the wafer. When measuring a single film of photoresist, the reliability of the Nanospec is acceptable. However, as described in chapter 2, monitoring photoresist on oxidized wafer requires a long scanning time, in order to obtain reliable measurements. Despite all its limitations, its monitoring capabilities are sufficient for our control system implemented in the Berkeley Microfabrication Laboratory. It must be noted, however, that the experimental error would be significantly reduced, had a better photospectrometer been available.

7.1.6 The Critical Dimension Measurement Computer

Next, the instrument used for measuring CDs is described. The Nanoline IV Critical Dimension Computer, made by Nanometrics is mainly designed to handle masks, but it can also handle wafers. This instrument measures the CD by scanning horizontally across a portion of the wafer. The reflectance plot is then displayed on the computer screen. The photoresist covered regions have higher reflectance than the rest of the wafer. The CD is then measured at the 35% level of the maximum reflectance, since for our system parameters, measurements are least sensitive to defocus at that particular threshold. This level can be adjusted if needed. The Nanoline system also incorporates several other specific programs, such as one that measures the pitch. The reliability of the Nanoline depends heavily on the type of lens used. We use the strongest lens available (100X) to increase the reliability of our data collection. The Nanoline measures the CD in μm, with a stated pre-

cision of ±0.01 μm. Experimentally, we find that the Nanoline precision is limited to ±0.03 μm.

7.2 Data Collection and Screening

The measurement procedure has been described in detail in chapter 2. However, not all the readings are used in the model generation system nor in the control system, because some of the measurements are not representative of the process. For example, occasionally, the spin-coat and bake equipment leaves streaks on the wafer, because the hot plate was not baking the wafer evenly. This was due for example to photoresist scums left on the hot plate. The photoresist thickness on the streak is then different from the mean thickness on the wafer, and is thus not representative of the general status of the wafer.

Normally, these situations reveal an out of control condition, and the machine would undergo a cleaning and maintenance process where all vacuum sensors would be checked and all scums removed from the baking plates. Unfortunately, since these situations happen frequently in the Berkeley Microfabrication Laboratory, it is unfeasible to take the machine down for maintenance so often.

Therefore, we chose instead to isolate these outliers when we characterize the process. To that end, we developed the following screening process. During the analysis of, for example, the thickness measurements, we plot the range of the thickness measurements in a range chart [12] (Figure 21). For a sample group of 4, the upper control limit of the range chart is $2.282\overline{R}$. Any wafer whose range violates that upper control limit is investigated, and the outlying point is deleted. Notice that we do not discard a data point just because the range of the results is too big; in order to discount the value of a data, it must clearly be far away from the other data points. This simple algorithm provides a rigorous and robust way to screen outliers. Once these outlying measurements are casted out, the mean response of each wafer, whether it is thickness, peak reflectance, or CD, is determined,

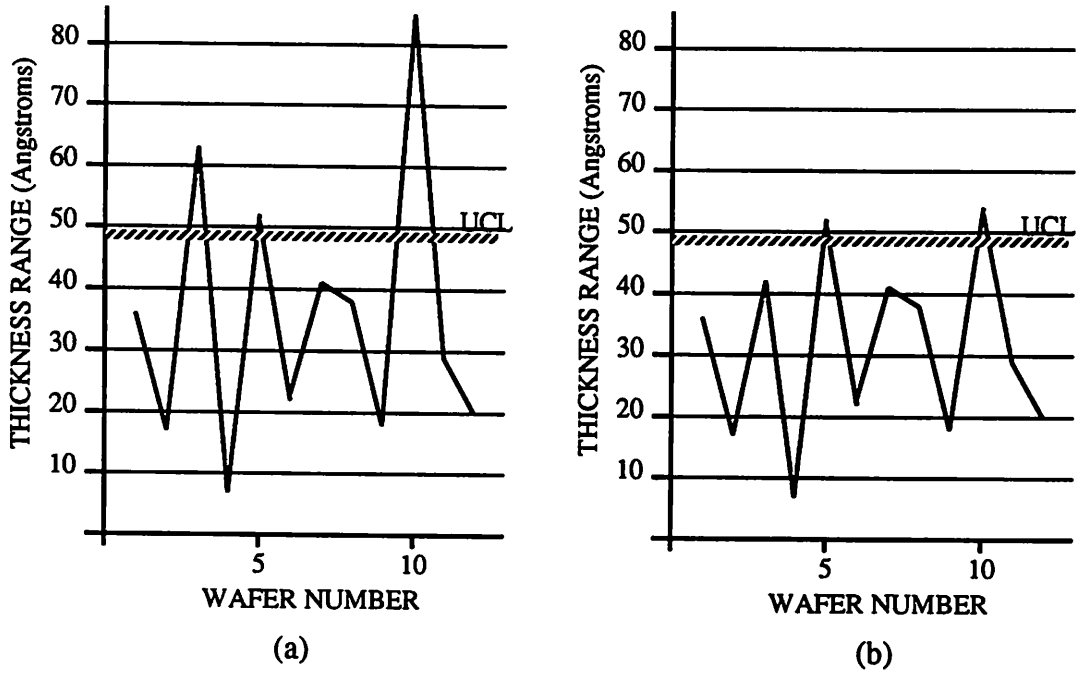


Figure 21 Thickness R-chart (a) before rejecting outlying measurements, and (b) after rejecting outlying measurements.

and entered into the model generation scheme or control scheme, which have been described in chapter 4 and 5.

7.3 Experimental Results of the Feedback Control Mechanism

Now that the experimental set-up, the monitoring system, the model generation system and the control system have been described, actual experiments can be discussed. First, we will evaluate the efficacity of the feedback controller, by applying it to the wafer track.

We operated the wafer track for several days and populated our database with historical data. The first malfunction alarm was generated at wafer #13 (Figure 23): both resist thickness and peak reflectance went out of specs (Figure 26 and Figure 27). The malfunction alarm was investigated and a probable cause of the problem was found to be the newly installed photoresist dispenser. To verify that the process has indeed been changed by the new dispensing method, two more wafers were processed. Once the change was

confirmed, the feedback controller was given a chance by the operator to bring the process back on target. Note that feedback control has not been activated by a control alarm as it typically would have been; feedback control, this time, has been activated by the operator, after analyzing the malfunction alarm.

The models were then updated by changing their constant terms to adapt to the new state of the process (point (a) in Figure 24 and Figure 25). Note that only one term can be updated at this point, since only one recipe change has occurred up to now, or statistically speaking, only one additional degree of freedom has been added to the set of experimental data. Although our objective is to bring both the thickness and the peak reflectance back to their initial levels, the optimizer could not find such a recipe using the current equipment models. Therefore a compromise recipe obtained by minimizing the deviations of the model predicted thickness and peak reflectance from their previous levels was used. This new recipe called for a spin speed of 4100 RPM, a spin time of 30 seconds, a baking temperature of 110°C and a baking time of 50 seconds. Normally, the baking temperature should not be changed, since that would change the characteristics of the photoresist. However, to bring the peak reflectance back to its normal level, i.e, to prove the efficacity of feedback control, this was needed. In an actual lithography production line, however, it would have been preferable to bring just the thickness back on target using feedback control while letting the feed-forward controller of the stepper correct the peak reflectance deviation.

Shortly afterwards, another *malfunction* alarm was generated (Figure 23). The problem was investigated and the cause has been traced to the reference wafer being unusually dirty. This problem was solved by cleaning the reference wafer. Since this affected our peak reflectance measurement, we "reset" the peak reflectance updating model algorithm by eliminating the correction model and starting from scratch. The thickness model, on the other hand, was allowed to be updated without interruption.

Figure 26 and Figure 27 show that the updated thickness model reflect the experimental data well, but the updated peak reflectance model does not. This causes a control alarm to be generated (point (b) in Figure 24 and Figure 25) and the models to be adaptively updated. The new models are:

$$T = \mathbf{-14794} + \frac{2.54 \cdot 10^6}{\sqrt{SPS}} + \frac{\mathbf{1.86 \cdot 10^7}}{BTE\sqrt{SPS}} - 3.78BTI - 0.28SPT - \frac{6.16 \cdot 10^7}{SPS}$$

$$\begin{aligned} R = {} & \mathbf{131.5} - 0.046SPS + 0.32SPT - 0.17BTE + 0.023BTI + 5.19 \cdot 10^{-5} \cdot SPS \cdot BTE \\ & -4.34 \cdot 10^{-5} \cdot SPS \cdot SPT - (1.07 \cdot 10^{-3} \cdot SPT \cdot BTE) + 5.15 \cdot 10^{-6} \cdot (SPS)^2 \\ & -4.11 \cdot 10^{-4} \cdot SPT \cdot BTI \end{aligned}$$

In these expressions the coefficients that have been updated are shown bold-faced. From these models a new recipe is generated, calling for a spin speed of 4000 RPM, a spin time of 60 sec, a baking temperature of 113°C, and a baking time of 50 sec. The same remarks concerning the change of baking temperature are applicable again. This time, the updated models allow for a recipe that brings both thickness and peak reflectance to their original levels. Experimental results have become much closer to the model predicted results. In fact, considering the models' prediction errors, these experimental results can be considered on target compared to the model predicted results. Several wafers later a third malfunction alarm was generated, which was investigated and dismissed as an outlier. Finally, a third control alarm is generated. This alarm is the result of a documented, continuous drift in the photoresist dispensing mechanism.

In summary, the feedback controller has proven to be very efficient and successful in bringing a process back on target. This is largely due to the robustness and accuracy with which the model update algorithm adapts the wafer track model to the new state of the process.

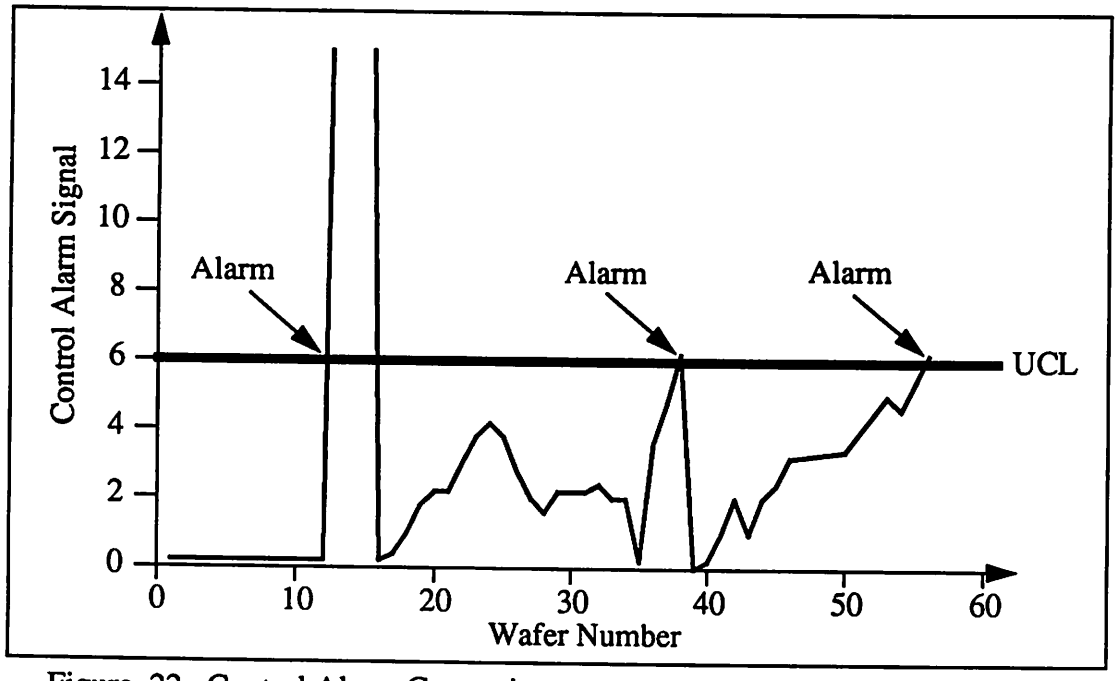


Figure 22 Control Alarm Generation

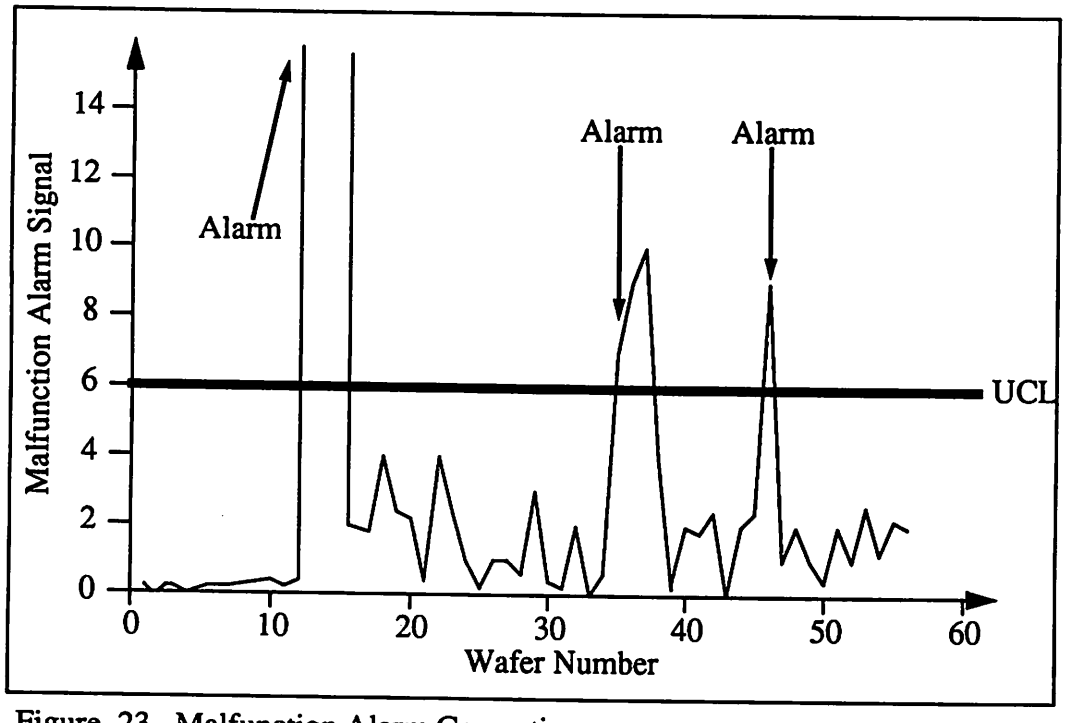


Figure 23 Malfunction Alarm Generation

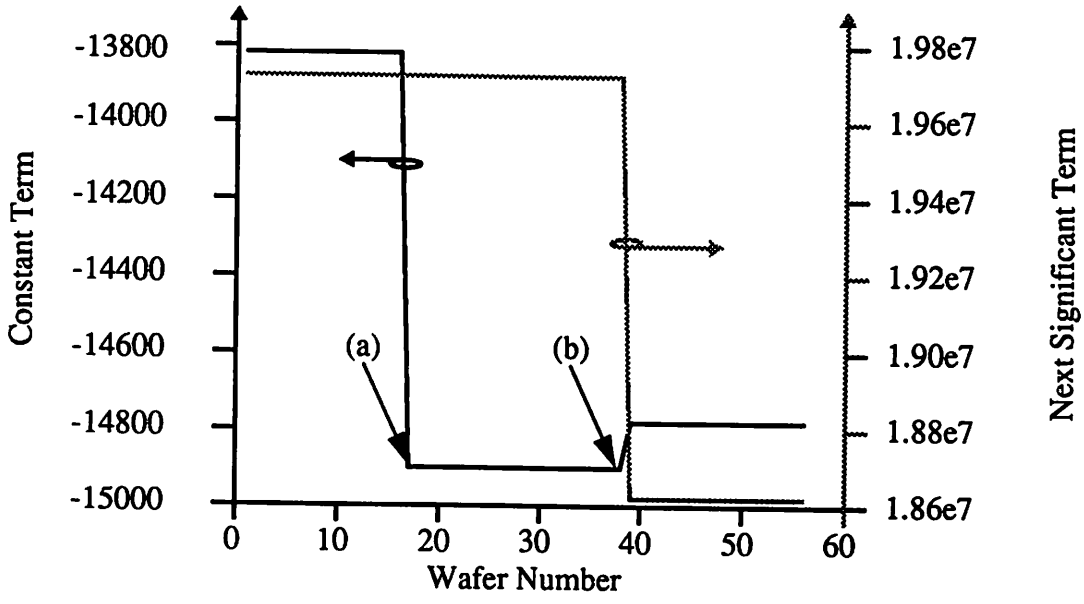


Figure 24 Thickness model corrections.

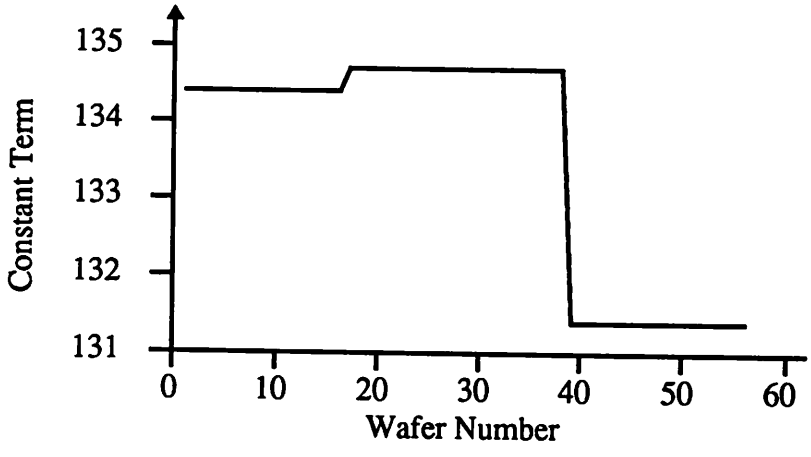


Figure 25 Reflectance model corrections.

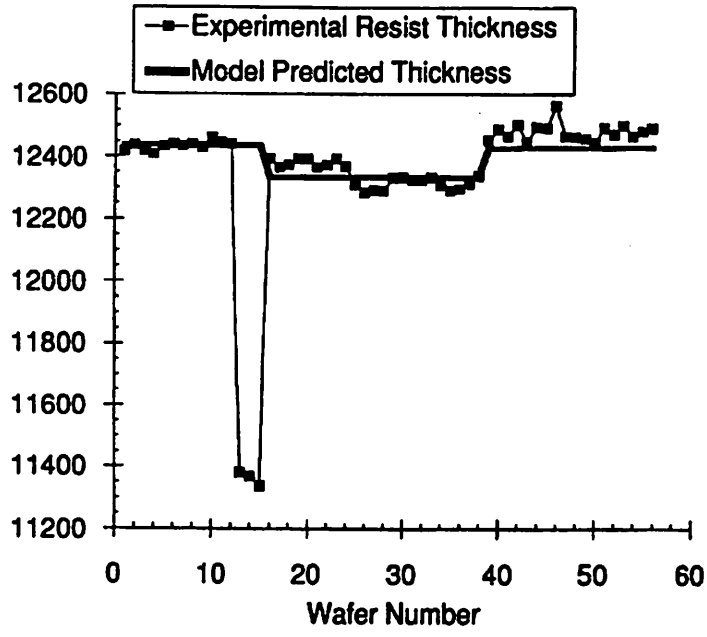


Figure 26 Effect of feedback control on thickness measurements.

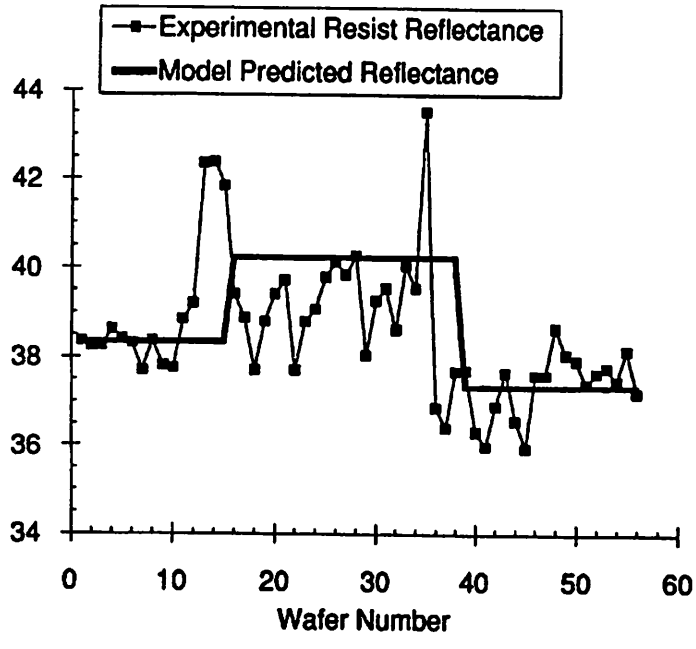


Figure 27 Effect of feedback control on reflectance measurements.

7.4 Experimental Results of the Feed-Forward Control Mechanism

Next, we evaluate the efficacity of the feed-forward controller by applying it to the stepper. The reason why the stepper has been chosen over the developer is because: 1) its input parameters can be changed simply and accurately; 2) the output is easily monitored; 3) we have an excellent empirical model for the stepper. This model has an adjusted R^2 of 94%; and 4) the wafer track that comes before the stepper is noisy and susceptible to large drifts within a few days. This allows us to simulate the behavior of an industrial machine in a shorter amount of time.

As done previously in the feedback experiment, we first operated the wafer track and the stepper using the standard recipes for several days, while passively monitoring the process. Once enough historical data has been collected, the target peak reflectance was chosen to be 87.75%, which was the average output peak reflectance of the historical data. Then, we let the wafer track run under open loop conditions, while we operated the stepper alternatively under open loop and closed loop conditions. Open loop conditions simply refer to operating the equipment without any control action, while closed loop conditions refer to operating it under feedback and feed-forward control. As noted in chapter 5, the feed-forward controller needs to be constantly coupled to the feedback controller, because it relies heavily on the accuracy of the equipment models which are updated by the feedback controller.

Using 14 wafers, we divided them into 2 lots of 7 wafers after they have been processed by the wafer track. The first lot was processed using the standard exposure time, while each wafer of the second lot had its own custom recipe, generated by the feed-forward controller. Then the wafers were post-baked and the peak reflectance was measured on each wafer. This experiment was repeated four times over a period of four days and the results are shown in Figure 29.

The shaded regions correspond to wafers run under open loop conditions, while the clear regions correspond to wafers run under closed loop conditions. The status of the resist after exposure has noticeably drifted within those four days. Since the GCA 6200 stepper has an internal feedback control loop that compensates for drifts in light source intensity, and therefore ensures that the exposure dose stays constant over the life of the lamp, this observed drift can be attributed to the "aging" of the wafer track. Figure 28 shows more clearly the advantages of the feed-forward controller. Under closed loop conditions, the standard deviation of the peak reflectance after exposure has been reduced from 3.4% to 1.7%, while at the same time, the process mean has been kept closer to target: under open loop operation, the average output peak reflectance has drifted away from 87.75% to 84.8%, while under feed-forward control, it stayed around 87.1%. Since the replication error of the stepper is ±1.08%, 87.1% relative to a target of 87.75% is considered on target. In summary, the application of feed-forward control results in a much tighter process response distribution that is well centered on target.

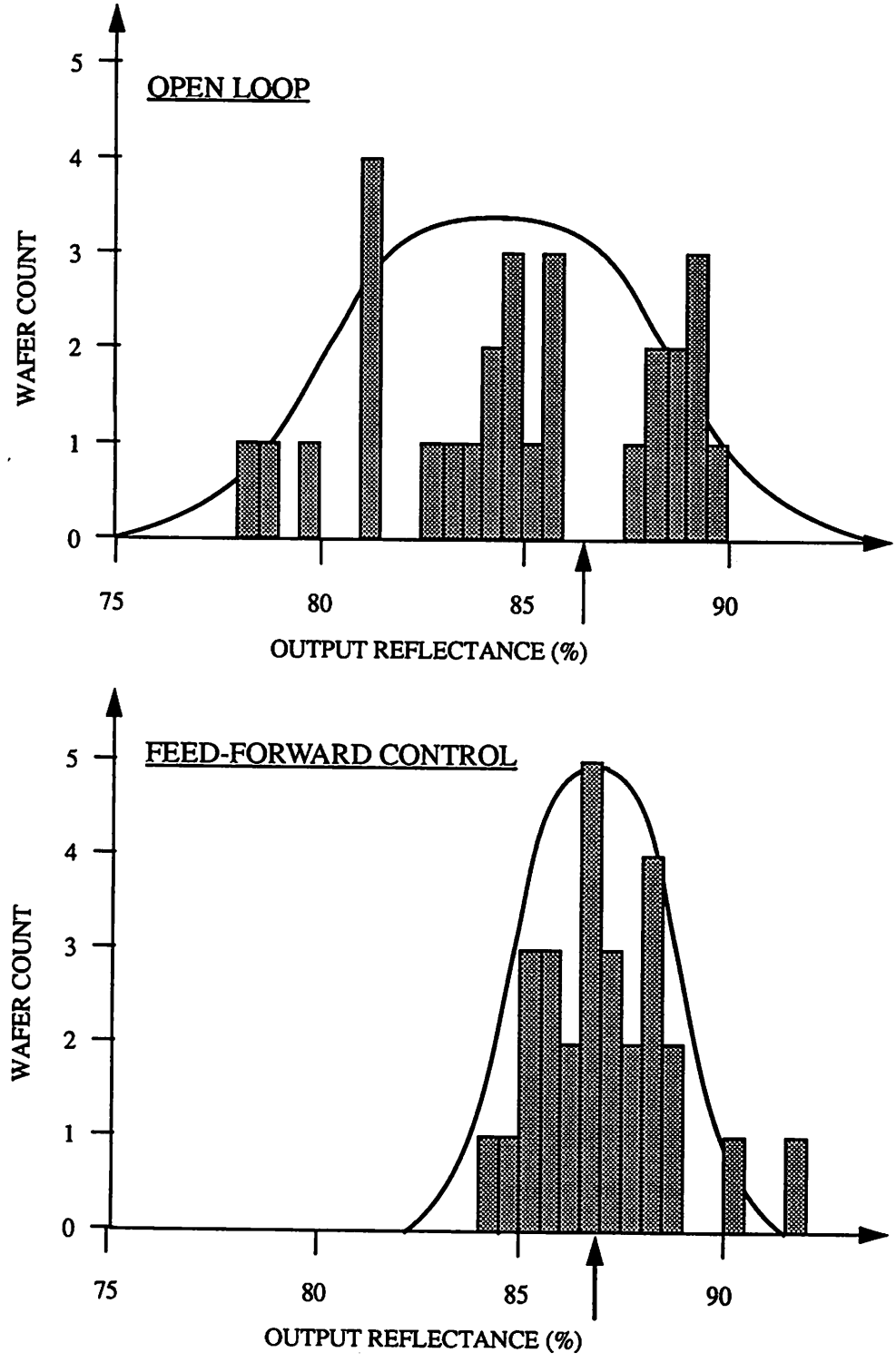


Figure 28 Process Variation Under (a) Open Loop and (b) Feed-Forward Control

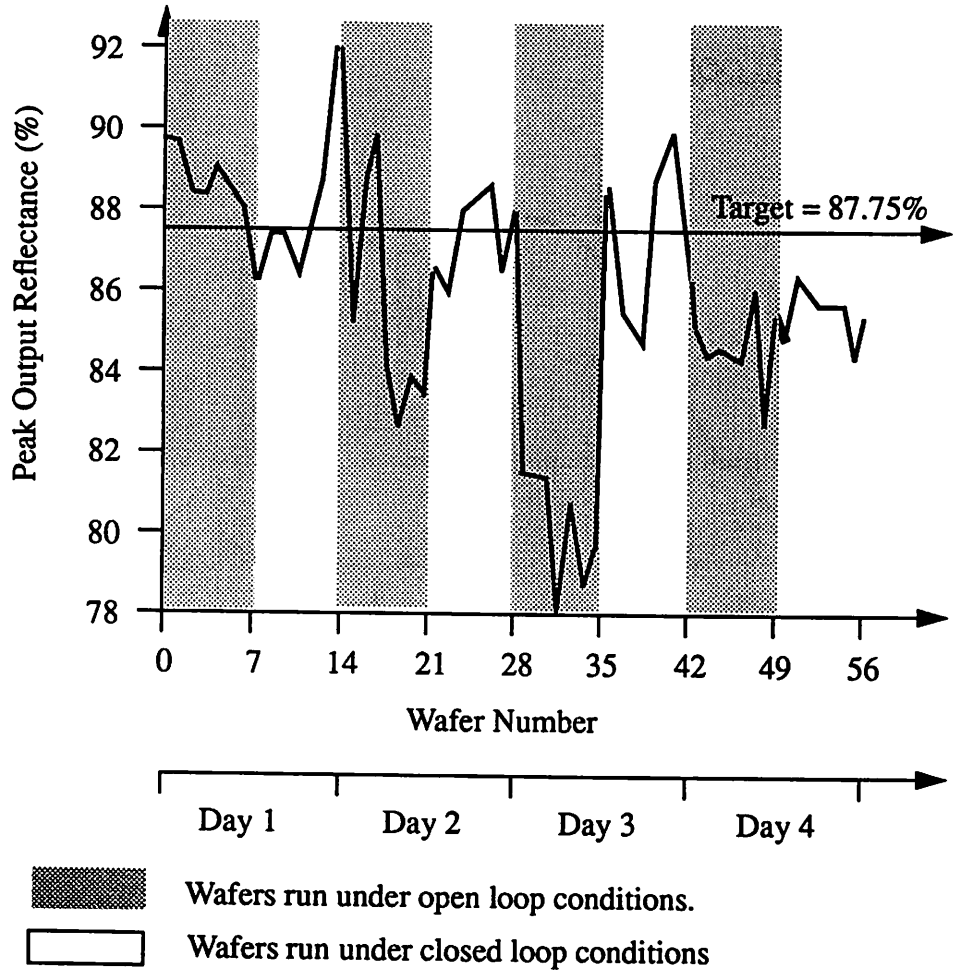


Figure 29 Experimental Results of Feed-Forward Control on the Stepper

7.5 Summary

Encouraging experimental results of both feedback and feed-forward controllers have been presented in this chapter. The feedback controller is very efficient in detecting drifts and in bringing the process back on target, while the complementing feed-forward controller is very efficient in reducing process variations and in bringing the present workcell back on target. The combination of both controllers has proven to be a powerful tool in increasing the process capability of the lithography sequence.

Chapter 8 Conclusions

In conclusion, a complete and fully automated supervisory control system for the photolithography process sequence has been developed in order to control CDs of photoresist patterns. To understand the big picture, it is best to describe a wafer run under the supervisory control system.

After a wafer has been spin-coated and baked, its status is evaluated using robust metrologies which measure resist thickness and peak reflectance. These measurements are then sent via ethernet to the fully automated feedback and feed-forward controllers. The feedback controller checks that the wafer track is in control. If it is not, it updates the wafer track equipment model and generates a new recipe so that the next wafer is processed properly. Meanwhile, the feed-forward controller predicts if the peak reflectance of the present wafer would be on target, using the current stepper model. If not, the feed-forward controller calculates a new exposure time for the present wafer in order to bring it closer to target, and the wafer gets exposed under that new exposure time. To confirm the reliability of the stepper model, the status of the present wafer is checked again after the wafer has been exposed and post-baked. If the stepper model has been unreliable, the feedback controller updates the stepper model and generates a new recipe to bring the next wafer closer to target. Meanwhile, that last reflectance measurement is sent to the developer model which determines if a change in development time is needed. If a change is needed, the wafer is processed using the new development time. Finally, the CD of the photoresist pattern is inspected. By this time, most of the process variations that would have caused CD variations have been eliminated by the supervisory control system. If CD variations still exist, they are used by the feedback controller to update the developer model.

Finally, some words about limitations and caveats of the supervisory control system. Although the system works very well in the Berkeley Microfabrication Laboratory, the reader is reminded that the metrology used to measure peak reflectance contains some weak assumptions and therefore needs to be improved. Another needed improvement is to enlarge the range of equipment models, and one way of achieving that task is to develop easily measurable process parameters that are related to process parameters used in physical simulators such as SAMPLE [33].

REFERENCES

[1] Zhi-min Ling, Sovarong Leang and Costas Spanos, "In-Line Supervisory Control in a Photolithographic Workcell", SPIE, Symposium on Microelectronic Processing Integration, Santa Clara, Sept. 1990.

[2] R. Azari et al, "Dynamic Statistical Process Control", SPIE vol.921, p.258, 1988

[3] F. H. Dill, W. P. Hornberger, P. S. Hauge, and J. M. Shaw, "Characterization of Positive Photoresist", IEEE Transactions on Electron Devices, Vol. ED-22, No.7, July 1975.

[4] W.G. Oldham et al, "A General Simulator for VLSI Lithography and Etching Process", IEEE trans. Electron Devices, vol. ED-27, p.717, 1979.

[5] C.A. Mack, "PROLITH: A Comprehensive Optical Lithography Model", SPIE vol.538, p.207, 1985.

[6] M. Watts, et al, "Photoresist as its Own Process Monitor", Solid State Technology, p.59, July 1988.

[7] Zhi-min Ling and Costas Spanos, "A Novel Approach for Film Reflectance Measurement and its Application for the Control of a Photolithography Workcell", ECS Conference Proceedings, Washington, Sept. 1990.

[8] F.H. Dill et al, "Modeling Projection Printing of Positive Photoresist", IEEE trans. on Electron Devices, vol. ED-22, No. 7, p.456, July 1975.

[9] P.H. Berning, "Theory and Calculations of Optical Thin Films", in Physics of Thin Film, vol. I, New York: Academic Press, p. 69, 1963.

[10] KTI Chemicals Company.

[11] G.Box, W. Hunter & S. Hunter, "Statistics for Experimenters: an Introduction to Design, Data Analysis, and Model Building" 1st ed., New York, John Wiley & Sons, 1978.

[12] Douglas C. Montgomery, "Introduction to Statistical Quality Control", 2nd ed., New York: John Wiley & Sons, 1990.

[13] BBN Software Products Corporation, RS/1 Release 3.0 Software.

[14] Warren Flack, David Soong, Alexis Bell, and Dennis Hess, "A Mathematical Model for Spin-Coating of Polymer Resists", Journal of Applied Physics, Vol. 56, pp. 1199-1206, Aug. 1984.

[15] Richard Guldi, C.D. Jenkins, G.M. Damminga, T.A. Baum, and T.A. Foster, "Process Optimization Tweaking Tool (POTT) and its Application in Controlling Oxidation Thickness", IEEE trans. on Semiconductor Manufacturing, Vol.2, No. 2, pp. 54-59, May 1989.

[16] Emanuel Sachs, R.S. Guo, S. Ha, and A. Hu, "Process Control System for VLSI Fabrication", IEEE Trans. on Semiconductor Manufacturing, Vol. 4, No. 2, pp. 134-144, May, 1990.

[17] ULTRAMAX, Version 4.1, by ULTRAMAX Corp., 1990.

[18] Sovarong Leang and Costas Spanos, "Statistically Based Feedback Control of Photoresist Application", ASMC Proceedings, Boston, Oct. 21, 1991.

[19] IEEE International Solid State Circuits Conference Proceedings, High Density DRAMS session, San Francisco, February, 1991.

[20] Sovarong Leang, "Requirements for a 1 Gbit DRAM", EE243 Spring 1991 Term Report, UC Berkeley, Berkeley, May, 1991.

[21] Stephen Parke, "Requirements for a Gigabit DRAM", EE243 Spring 1991 Term Report, UC Berkeley, Berkeley, May, 1991.

[22] B.J. Mandel, "The Regression Control Chart", Journal of Quality Technology, Vol.1, No.1, pp. 1-9, Jan. 1969.

[23] Richard Harris, "A Primer of Multivariate Statistics", Academic Press, 1975.

[24] Sherry Lee, "A Strategy for Adaptive Regression Modeling of LPCVD Reactors",

Special Issues in Semiconductor Manufacturing, pp.69-80, University of California, Berkeley / ERL M90/8, January 1990.

[25] Ronald Crosier, "Multivariate Generalizations of Cumulative Sum Quality-Control Schemes", Technometrics, Vol. 30, No. 3, Aug. 1988.

[26] William Woodall and Matoteng Ncube, "Multivariate CUSUM Quality-Control Procedures", Technometrics, Vol. 27, No. 3, Aug. 1985.

[27] Zhi-min Ling, Sovarong Leang, and Costas Spanos, "A Lithography Workcell Monitoring and Modeling Scheme", Proceedings of Micro-Electronique 1990, Sept. 1990, Leuven, Belgium.

[28] M.J.D. Powell, "A Fast Algorithm for Nonlinearly Constrained Optimization Calculations.", Proceedings of Dundee Conference on Numerical Analysis, 1977.

[29] Chris Hegarty, "Process-Flow Specification and Dynamic Run Modification For Semiconductor Manufacturing", Ph.D thesis 1991, Memorandum No. UCB/ERL M91/40, April 1991.

[30] M. Abrahams and F. Rizzardi, "BLSS: The Berkeley Interactive Statistical System", UC Berkeley, Berkeley, 1988.

[31] Nanometrics Inc., 690 East Arques Ave., Sunnyvale, CA 94086, 1990

[32] Hegarty, Chris. Private Communications. July, 1991

[33] Andrew Neureuther et al, "SAMPLE", U.C. Berkeley, Berkeley, 1982.