Copyright © 1994, 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.

THE APPLICATION OF DYNAMIC SPECIFICATIONS IN A MULTISTEP LITHOGRAPHIC SEQUENCE

by

Shang-Yi Ma

Memorandum No. UCB/ERL M94/84

17 October 1994

COVER PAGE

THE APPLICATION OF DYNAMIC SPECIFICATIONS IN A MULTISTEP LITHOGRAPHIC SEQUENCE

by

Shang-Yi Ma

Memorandum No. UCB/ERL M94/84

17 October 1994

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

THE APPLICATION OF DYNAMIC SPECIFICATIONS IN A MULTISTEP LITHOGRAPHIC SEQUENCE

by

Shang-Yi Ma

Memorandum No. UCB/ERL M94/84

17 October 1994

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Acknowledgments

I would like to express my greatest appreciation and gratitude to my research advisor, Professor Costas J. Spanos, for his invaluable guidance and constant support throughout the duration of this project. I also thank Professor Kameshwar Poolla for serving on my dissertation committee.

I also want to thank Sovarong Leang for developing the monitoring system and providing his valuable experience on the photo process. Bart J. Bombay, Herb Huang, David Mudie, and Dong-Wu Zhao developed and improved the BCAM environment which is the heart of this supervisory control system. John Thomson proposed and implemented the code for dynamic specifications.

I am also grateful to Eric Boskin, Sherry Lee, Tony Miranda, Crid Yu, Sean Cunningham, Roawen Chen, Xinhui Niu, Mark Hatzilambrou, Manolis Terrovitis, David Rodriguez, Pamela Tsai, and Dr. Raymond Chen, for their friendship and assistance during my stay at the BCAM research group.

I am thankful to the staff of the Berkeley Microfabrication Laboratory for their excellent technical support on equipment and the process. Special thanks should be given to Robert Hamilton, David Hebert, Debra Hebert, Evan Stateler, Mike Linan, Mark Kraitchman, Marilyn Kushner, Maria Perez, and Katalin Voros.

I like to express my deepest appreciation to my parents and my wife. Without their encouragement, support and sacrifice, I could not have completed this research project.

This research has been sponsored by the Semiconductor Research Corporation and Sematech, under the SRC contract 94-YP-700.

Table of Contents

Chapter 1. Introduction 1

1.1 Motivation 1

1.2 Thesis Organization 3

Chapter 2. Multistep Run-to-Run Control Using Dynamic Specifications 4

2.1 Principles and Procedures of Dynamic Specifications 4

2.1.1 Specification Propagation Using Monte Carlo Simulation 5

2.1.2 Selection of Acceptable Input Points 5

2.1.3 Principal Component Analysis 6

2.1.4 Cost Function Derivation 9

2.2 Simulated Examples of Dynamic Specifications 11

Chapter 3. The Infrastructure of the Supervisory Control System On The Photolithographic Workcell 16

3.1 An Overview of the Photolithographic Workcell 16

3.2 The Measuring System 19

3.3 The BCAM Control System 21

3.4 Summary of the Supervisory Control System 25

Chapter 4. Experimental Results and Discussion: Part I 28

4.1 Introduction 28

4.2 Experiment #1: Calibration of the PR Inspector 28

4.3 Experiment #2: Building Empirical Equipment Models 30

4.3.1 Design of Experiment for Building Equipment Models 31

4.3.2 Model Development 35

4.4 Experiment #3: The Baseline Historical Runs 43

4.4.1 Experimental Setup 43

4.4.2 Trendcharts of Four Output Parameters 45

4.4.3 Correlation Between the Output Parameters 48

4.4.4 Distribution of Output Parameters 51

4.5 Summary 51

Chapter 5 Experimental Results and Discussion: Part II 54

5.1 Experiments to Compare Different Control Methods 54

5.1.1 Experimental Setup 54

5.1.2 Thickness Trend and Distribution of Control Comparison Experiments 55

5.1.3 Trend and Distribution of PAC and PACxp of Control Comparison Experiments 61

5.1.4 CD Trend and Distribution of Control Comparison Experiments 70

References 78

Chapter 1 Introduction

1.1 Motivation

The goal of this project is to implement the experimental infrastructure for a flexible supervisory controlled sequence within a photolithographic workcell. We intend to create a hardware demonstration of distributed run-to-run control across the photolithographic steps required to define a polysilicon gate. One important objective of this novel implementation is to realize the idea of a "global" supervisory controller which oversees the operation of individual controllers by means of dynamic modification of the targets the individual controllers must follow.

In a VLSI manufacturing environment, there is a high demand for reliable process control. To obtain better control for individual process steps, various control systems (feedback control and feed-forward control) have been developed. However, most control systems are only "local" controllers designed to regulate only one piece of equipment. In order to develop a global, supervisory control system, which can coordinate and integrate the control for a workcell consisting of multiple steps and many pieces of equipment, the use of dynamic-specifications offers a promising approach.

Traditionally, the specifications of any individual process step are considered rigid references and in most cases remain fixed. The concept of *dynamic specifications* allows intermediate specifications to change in response to a final specification change or process drift. The spec change or process drift of an individual step can be propagated "upstream" to all the preceding steps whose specifications are changed accordingly. Since the specifications are the major references by which local controllers function, a global control system is thus built. In this way, we add flexibility and capability to the control system.

In order to build an effective and reliable hardware infrastructure for this demonstration, we developed new models for the three pieces of equipment used in the photolithographic

sequence. These models bear more physical meaning than earlier empirical models. First, the "SAMPLE" process simulator [1] was used to derive the equipment model for the photoresist developer. Second, a new parameter, photoactive compound concentration (PAC), is used in the place of photoresist reflectance, since PAC is more physically meaningful than reflectance.

A new measuring system capable of computing the PAC has been developed at the BCAM. This system includes an in-situ photoresist measuring instrument (PR Inspector) and supporting software developed at BCAM. The PR Inspector, which contains an in-situ spectrometer, can obtain an *in-situ* spectrograph by reflecting light on the coated wafer surface for a few seconds.

The photolithographic workcell consists of three pieces of equipment: The wafer track (or the photoresist coater), the stepper, and the photoresist developer. The *final output* of the entire workcell is the critical dimension (CD) of the developed resist pattern. This CD is also the output of the photoresist developer. The *intermediate output* parameters are the outputs of the wafer track and of the stepper. These are the photoresist thickness, the PAC-before-exposure and the PAC-after-exposure. For each measured output parameter a *specification* (including an absolute target and an acceptable range) is assigned.

For the experiments that will be described in this document, the BCAM control and monitoring environment is used for local control within this infrastructure. The programs used to propagate the specifications, as well as the concept and procedures of dynamic specifications, were developed earlier at the BCAM research group [2]. For this project, the programs were partially modified and updated to match the new equipment models and parameters.

A series of experiments was performed to demonstrate the superiority of a supervisory control system using dynamic specifications. These experiments demonstrate that the dynamic-specification approach is a promising method for improving the capability and flexibility of a controlled sequence.

1.2 Thesis Organization

The subject of this thesis is supervisory control for the photolithographic sequence. The emphasis is on the experimental demonstration of process control by different methods. A description on how to handle dynamic specifications and some simulation results is provided in Chapter 2. Chapter 3 discusses the hardware and software infrastructure of this project, including brief descriptions of the photolithographic workcell, the measuring system, and the alarm and control system in the BCAM environment. Detailed experimental results and discussion are presented in Chapters 4 and 5. In Chapter 6 we present our conclusions and we discuss potential future work.

Chapter2 Multistep Run-to-Run Control Using Dynamic Specifications

2.1 Principles and Procedures of Dynamic Specifications

The general concept and computer implementation of a dynamic specification scheme were first described in [2]. The major advantages of this novel approach are briefly described as follows: First, the independent local controllers are "linked" and integrated as a whole by specification propagation, thus a "global" supervisory control system is realized. Second, the overall control system can respond more effectively to process fluctuation as well as to specification change. Finally, it can provide the basis for synthesizing a new process, by determining the specifications and the recipe of all the steps within a new process sequence more efficiently.

It is noteworthy that the *inputs* to the equipment model of a certain processing step can be categorized into two types. The first type of inputs are *controllable inputs* which are easily modified by adjusting a machine parameter (e.g., the develop time on the developer). The second type of inputs are not controllable on that machine because they are the measured outputs from the preceding machine, thus called *upstream inputs*. For example, the thickness and after-exposure-PAC (PACxp) are the outputs from the stepper as well as the upstream inputs to the developer. Figure 2.1 gives a high level view of a photolithographic workcell. Those parameters in rectangular boxes (spin speed, baking time, exposure dose....etc.) are the controllable inputs while the parameters inside ellipses (T, PAC, and PACxp) are upstream inputs.

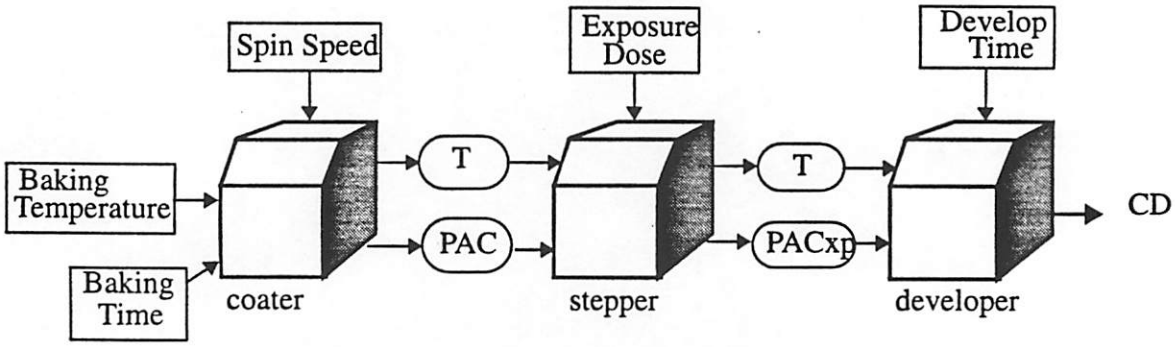


Figure 2.1 Two Types of Inputs Within A Photolithographic Workcell

The procedure that handles the dynamic specifications is described next. For simplicity, we discuss only two consecutive processes, where step B comes after step A. The procedure is used to propagate the specifications of step B upstream to the specifications of step A. This approach can be repeated and generalized for any number of processing steps. The sequence of spec generation is described next.

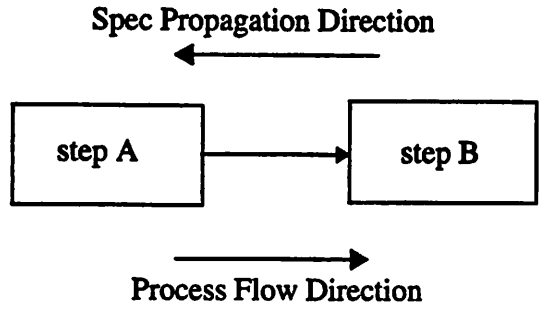


Figure 2.2 The Relation Between Step A and Step B

2.1.1 Specification Propagation Using Monte Carlo Simulation

In order to generate the specifications at the input of process step B, we first select the maximal range of all upstream inputs of step B. The range is set by the user to reflect machine and process limitations. Then, Monte Carlo simulation is done to randomly generate uniformly distributed points across the range of upstream inputs to step B. Each point represents a unique set of input combinations of step B. Consequently, these points are fed into the equipment models of step B to compute the corresponding points in the output space of step B. If step B only has one output, then one such point represents one output value. In the case of multiple outputs, one point stands for a set of output values.

2.1.2 Selection of Acceptable Input Points

Not all points in the input set for B will lead to outputs of B that meet their specifications. This is established by the use of a cost function, defined as follows:

(1) For multiple outputs whose specifications are independently defined, the total cost is the sum of the costs of all outputs:

$$Cost = \sum_{i=1}^{n} k_i (y_i - y_i^{target})^2$$

where n is the number of outputs.

(2) For multiple outputs with dependent specifications (i.e., the specification of one output depend on the values of other outputs), the cost function is defined as:

$$Cost = \sum_{i=1}^{n} \sum_{j=1}^{n} k_{ij} ((y_i - y_i^{target}) \cdot (y_j - y_j^{target}))$$

Once the appropriate cost function is defined, only those points with cost values less than a given number (the default value is 1.0) are considered to meet the specification. Based on this, we choose those points in the input space of B that lead to output points that meet their specifications. These points at the input space are the *acceptable inputs* of step B as well as the *desired outputs* of step A.

2.1.3 Principal Component Analysis

Once the acceptability region has been mapped, principal component analysis is done on those desired output points of step A in order to determine the specifications of step A. Principal component analysis (PCA) is a multivariate technique used to transform a correlated multi-dimensional data set into fewer uncorrelated variables. The purpose of using PCA here is to simplify and analyze the data to produce a set of specifications for multiple outputs of a machine. The PCA method [3] is briefly described next for a two-dimensional case. This can be easily generalized to n-dimensional problems. First, the covariance matrix **S** of the data set is calculated. For the two-output case, **S** is a 2×2 matrix. Then, the covariance matrix is diagonalized using singular value decomposition[1]:

$$\mathbf{U'SU} = \mathbf{L}$$

where $\mathbf{U} = [\, \mathbf{u_1} | \mathbf{u_2} \,]$, $\mathbf{u_1}$ and $\mathbf{u_2}$ are the eigenvectors of **S**.

1. Matrices are symbolized with upper case bold letters. Column vectors are symbolized with lower case bold letters.

The direction of the principal component can be obtained by taking the elements of the eigen-vectors as the rotation angles of a principal component axis related to the original axis. For example, Figure 2.3 shows 20 data points in a two-dimensional space, as well as the first derived principal component direction (PC_1) and the two angles ($\theta_{11}$and θ_{21}) between PC_1 and the two original axes. It also shows the second principal component (PC_2), which is orthogonal to PC_1, as well as the two angles (θ_{12} and θ_{22}) between PC_2 and the two original axes. The following relations define the rotation:

$$cos\theta_{11} = u_{11}, \quad cos\theta_{21} = u_{21}, \quad cos\theta_{12} = u_{12}, \quad cos\theta_{22} = u_{22}$$

where u_{11}, u_{21}, u_{12}, u_{22} are the elements of the matrix U.

As shown in Figure 2.3, PC_1 explains the most variation in the data set. PC_2 is the direction orthogonal to PC_1. This can be generalized to more than two dimensions. In the original space the data points are correlated, but in the PC_1-PC_2 space they are independent.

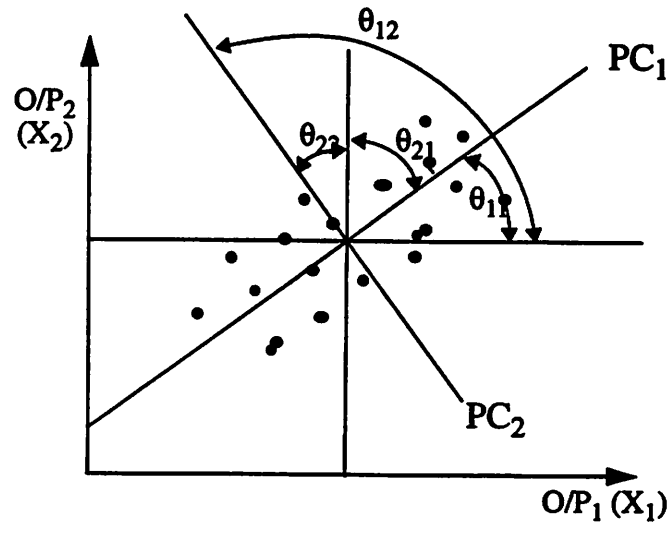


Figure 2.3 Axis Rotation Based On Principal Component Analysis

Looking now at the PCA results in the context of defining specifications for step A, the point where PC_1 and PC_2 intersect is also the target point with coordinates (x_1 , x_2); where x_1 and x_2 are the average values of output 1 (X_1) and output 2 (X_2), respectively. If we draw a rectangular box along the PC_1 and PC_2 directions, centered at the target point (as the dashed-line box shown in

Figure 2.4), the box can be seen as the approximate boundary of a set of specifications. In the rotated space defined by PC_1 and PC_2, this is a set of *independent* specifications. From the view-point of X_1 and X_2 (the actual outputs of the equipment), it is a set of *dependent* specifications.

In addition to the target and orientation of principal components, the range of the specifications must be determined. If the width of specifications along the PC_2 direction is W (as shown in Figure 2.4) and the variance of the distance from acceptable points to the PC_1 axis is σ^2, the two values are related as:

$$\frac{W}{2} = \sqrt{3\sigma^2}$$

since the variance of a uniform distribution of width W is $W^2/12$.

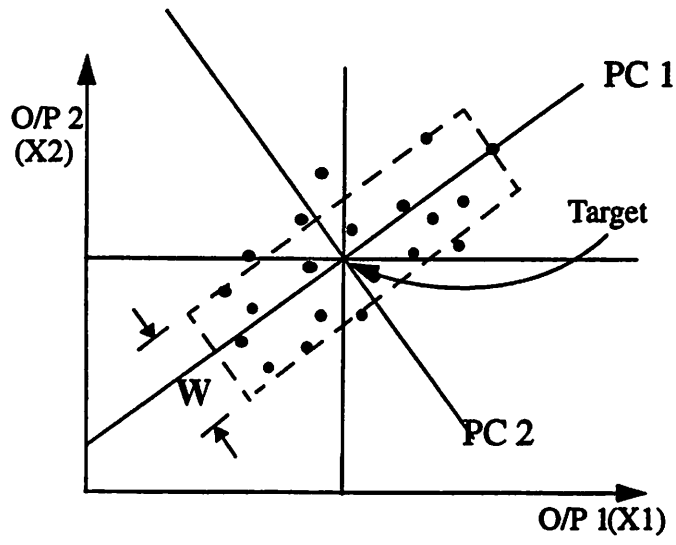


Figure 2.4 Specifications derived by Principal Component Analysis

In Figure 2.4, the distribution of acceptable points in PC_1 direction is much wider than that in PC_2 direction. This means that the process is less sensitive to variation along PC_1 compared to PC_2. In other words, a tighter specification is required along the PC_2 direction. It is noteworthy that in many other PCA applications PC_1 is more important since it explains more variation of the data

set. However, in the application of specification derivation, PC_2 deserves more attention because the process is more "sensitive" to change along the PC_2 direction.

Finally, note that before doing PCA, the X_1 and X_2 (outputs of step A, also inputs to step B) values need to be normalized such that they carry equal weight in the PCA. All X_1 and X_2 values need to be linearly scaled to be within the range -1 to 1.

2.1.4 Cost Function Derivation

After the specifications of step A are derived, the next step is to derive a cost function for step A which can be used by the controller of step A in order to center that process. The same cost function will also be used to further propagate the specifications to equipment upstream of step A. Usually the derived specifications of step A are *dependent* specifications and their cost function looks like:

$$Cost = \sum_{i=1}^{n} \sum_{j=1}^{n} k_{ij} ((y_i - y_i^{target}) \cdot (y_j - y_j^{target}))$$

Considering the two-dimensional case (n = 2), the cost function is also an equation of a tilted ellipse exactly contained in the specification box (as shown in Figure 2.5). In the PC_1-- PC_2 space the specifications are independent. Therefore, after coordinate transformation, the cost function has the form:

$$Cost = \sum_{i=1}^{n} k_i' (y_i' - y_i'^{\,target})^2$$

The coefficient k_i' is chosen such that if one output is right on its spec limit, the contribution to the cost value by this output *alone* will be 1. Therefore, each k_i' is calculated *independently* simply by: (1) setting all other k_i's to zero, (2) setting y_i'-$y_i'^{\,target}$ to the tolerance (distance from the spec limit to the target, or W/2 in Figure 2.4), and (3) setting the cost equal to 1. By coordinate transformation the k_{ij} coefficients of the cost function of dependent specifications can be calculated from the k_i coefficients of the cost function of independent specifications.

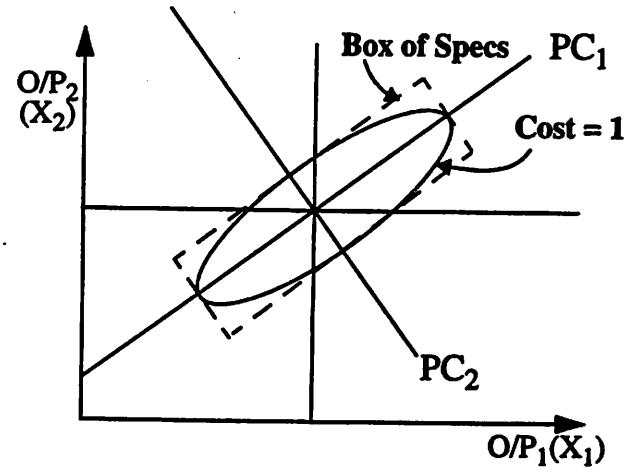


Figure 2.5 Cost Function and the Specification Box

The overall procedure to handle dynamic specifications is summarized in Figure 2.6. This procedure assumes nearly linear models in the vicinity of the derived specifications. As we will see next, even when this assumption is not always met, the above procedure offers a decent approximation to the actual acceptability region of a process.

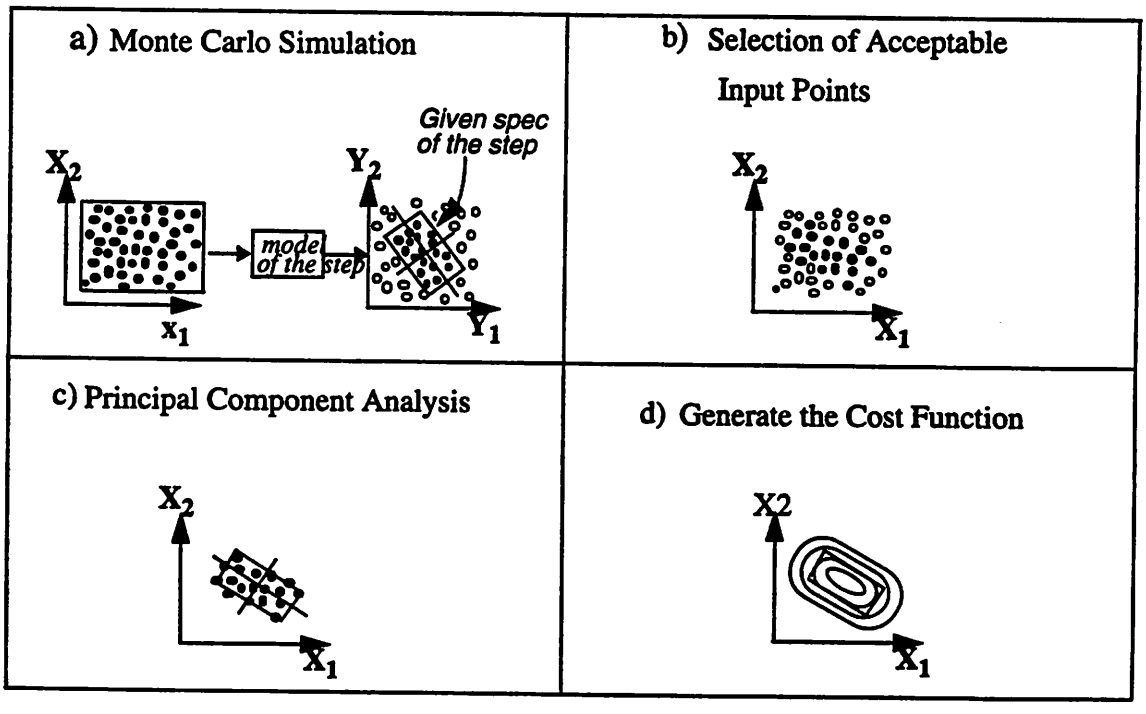


Figure 2.6 Procedure to Handle Dynamic Specifications

2.2 Simulated Examples of Dynamic Specifications

Three pieces of equipment, including a photoresist coater, a wafer stepper, and a developer are used in the photolithographic sequence. In this section, simulation results of spec propagation are presented. The simulation involved: (1) the spec propagation from the developer output (CD) to the wafer stepper outputs (PACxp and thickness[1]), and (2) the spec propagation from the wafer stepper output to the photoresist coater outputs (Thickness and PAC).

Figure 2.7 shows the result of spec propagation from the developer output (CD) to the wafer stepper outputs (thickness and PACxp). The acceptable points and the derived specifications (within the dashed-line box) in the thickness-PACxp space are shown. The equipment model of the developer (CD model) is described in detail in Chapter 4.2. The inputs of the CD model include thickness, PACxp, develop time, and exposure dose, but no PAC. The CD target and tolerance was 1.72 +/- 0.045 μm. The thickness input range was 11000 Å ~ 14400 Å (thus, on the normalized thickness axis in Figure 2.7, -1.00 represents 11000Å, +1.00 stands for 14400Å, 0.00 corresponds to 12700Å). The PACxp input range was 0.2 ~ 0.68. The develop time was fixed at 60 seconds. The exposure dose was maintained at 163 mj/cm^2.

Note that due to the highly nonlinear characteristics (a cosine term is included) of the CD model, the acceptable region in Figure 2.7 shows some "holes". Also, even when a point is within the specification limits (inside the box in the thickness-PACxp space) of the stepper, the corresponding output of the developer is not guaranteed to meet the CD specification. However, this usually happens to points near to the edge of the spec box. Usually the center of the box gives the targets and it can produce an output of the next step well within the specifications. So, although the PCA derived bounding box is a poor approximation of the acceptability region, the derived specs still effectively guide the controller towards the center of that region.

---

1. Thickness is not changed by the process done on the wafer stepper since it is an output of the photoresist coater and an input to the wafer stepper. It is also seen as an output of the wafer stepper (directly transmitted from the input to the output without any change) such that the output space of the wafer stepper contains two variables.

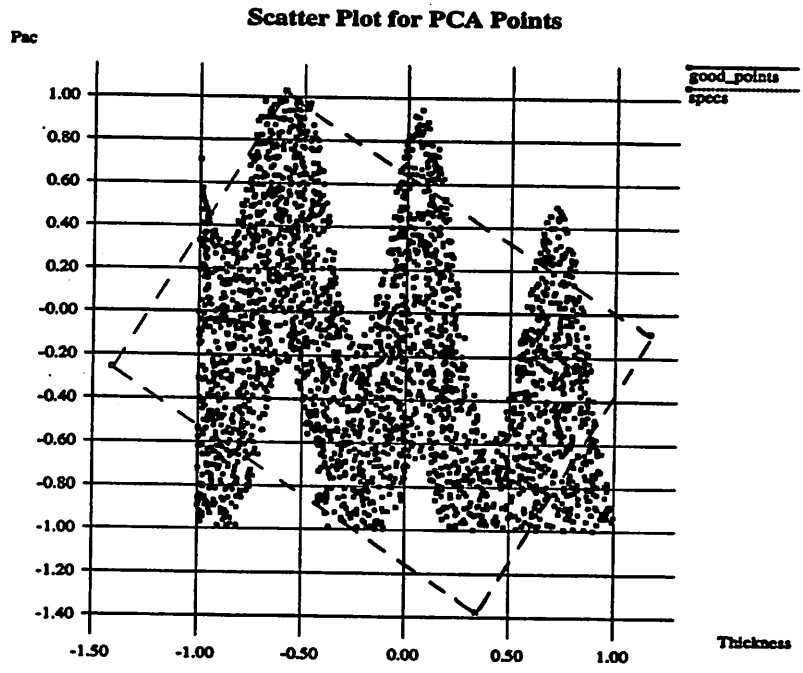


Figure 2.7 Acceptable Points and Derived Specifications in the Output Space of the Stepper Assuming a Fixed Develop Time

As it is discussed next, this situation is greatly improved when one assumes a *variable* develop time, as when the developer is continuously adjusted by a run-to-run controller. Figure 2.8 shows another simulation result which was run with the same conditions as that of Figure 2.7 except that the develop time was not fixed at 60 seconds, but allowed to vary between 40 to 80 seconds. It is seen in Figure 2.8 that the "holes" become less obvious than those in Figure 2.7. This means that with the addition of feed-forward control (the develop time can be adjusted to produce a better CD value), the "box" approximation of the acceptability region is more meaningful.

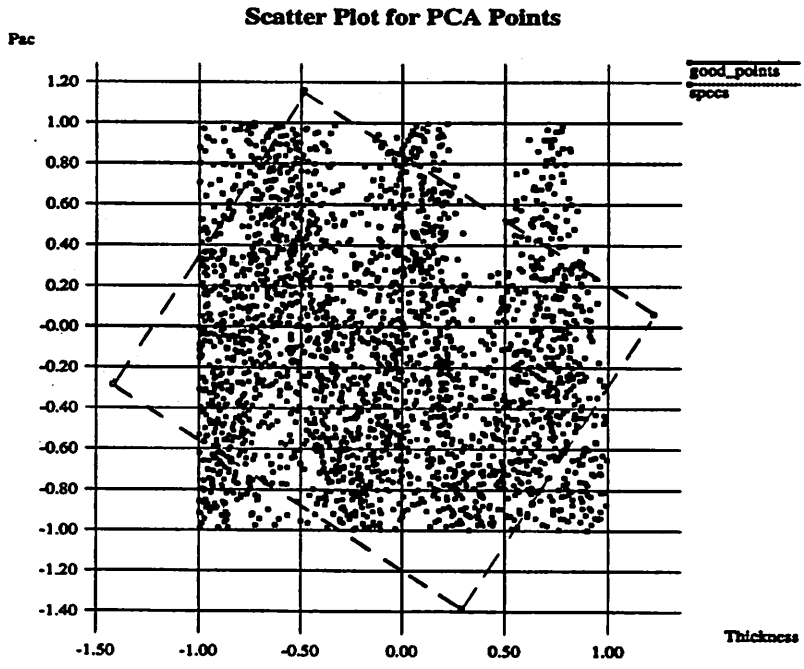


Figure 2.8 Acceptable Points and Derived Specifications in the Output Space of the Stepper with Variable Develop Time

Figure 2.9 shows the result of spec propagation from the wafer stepper outputs (thickness and PACxp) to the photoresist coater outputs (thickness and PAC). The acceptable points and the derived specifications (within the dashed-line box) in the thickness-PAC space are shown. The equipment model of the wafer stepper (PACxp model) is described in detail in Chapter 4.2. The PACxp target and tolerance was 0.44 +/- 0.02. The thickness input range was 12000 Å ~ 13400 Å. The PAC input range was 0.86 ~ 1.06. The exposure dose was fixed at 163 mj/cm^2. For each parameter which is the output of one step and input to the next, the target of one step is equal to the center of input search range of the next step.

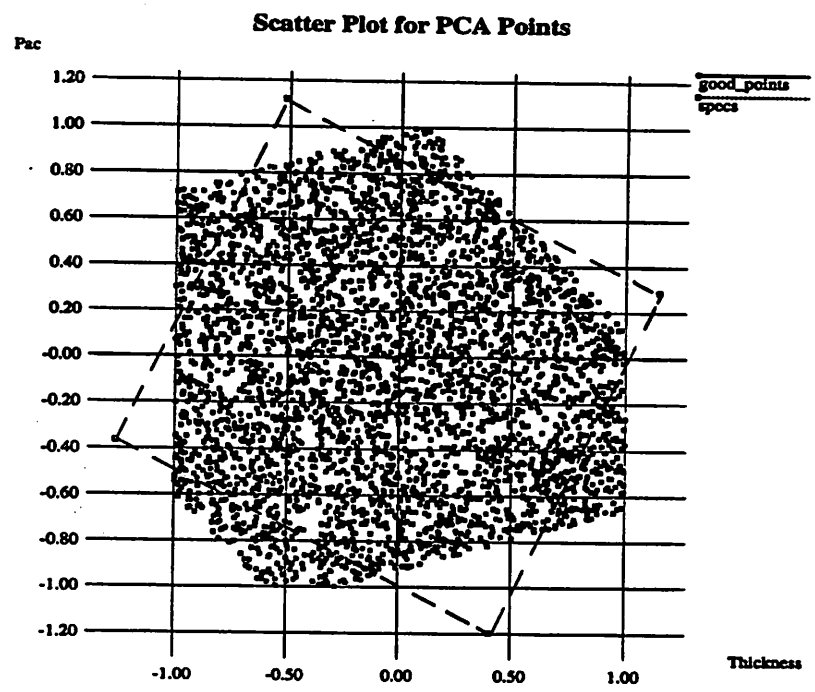


Figure 2.9 Acceptable Points and Derived Specifications in the Output Space of the Photoresist Coater with Fixed Exposure Dose

Figure 2.10 is the same as Figure 2.9 except that the exposure dose is not fixed. The exposure dose is allowed to vary from 143 to 183 mj/cm^2. As a result, the acceptable region is enlarged.

In the next chapter we discuss the hardware and software infrastructure of our supervisory control system for a photolithographic workcell.

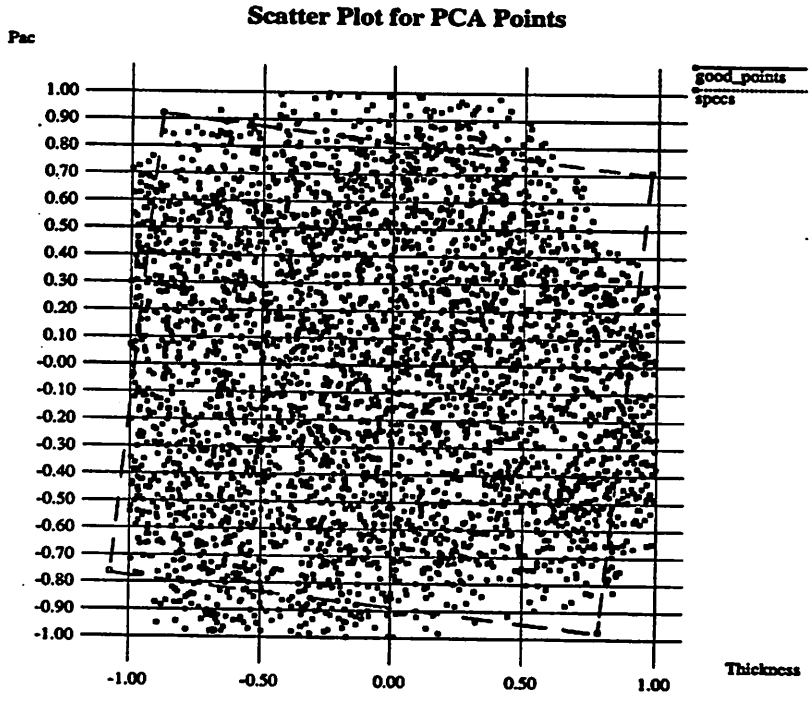


Figure 2.10 Acceptable Points and Derived Specifications in the Output Space of the Photoresist Coater with Variable Exposure Dose

Chapter 3 The Infrastructure of the Supervisory Control System On The Photolithographic Workcell

In this chapter, we present the hardware and software infrastructure of the supervisory control system applied on the photolithographic workcell. Descriptions on the photolithographic workcell, the monitoring system, and the process alarm/control system are included.

3.1 An Overview of The Photolithographic Workcell

The photolithographic workcell consists of three pieces of equipment. These three machines and their input/ output parameters of the equipment models are shown in Figure 3.1.

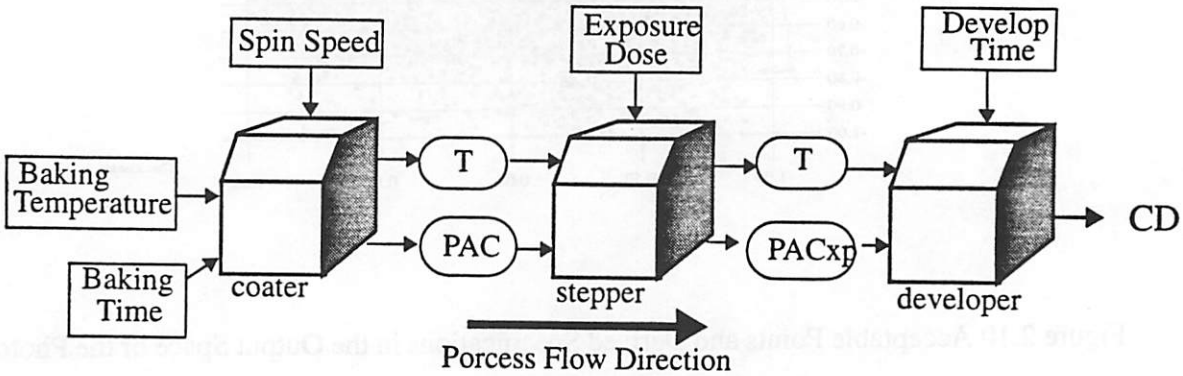


Figure 3.1 Photolithographic Machines And Their Input/Output Parameters

Those parameters in rectangle boxes (spin speed, baking time, exposure dose....etc.) are the controllable inputs. T (Thickness), PAC (PhotoActive Compound concentration), and PACxp (PAC after exposure) are not only the outputs from a certain step but also the upstream inputs to the next step. The CD (Critical Dimension) is the final output of the workcell. The CD is the measured width of the photoresist pattern after development (but before etching). The CD of the polysilicon layer is very important because it eventually determines the physical channel length of the devices, a feature that dominates the overall performance of the IC.

In most IC fabs the CD has been given lots of attention, but the measures taken for CD control are limited, and require considerable judgement and experience to employ. The supervisory control system introduced here is intended to be more flexible and automated than traditional methods, so that better capability of CD can be obtained with less cost. This is accomplished by the effective control of individual machines, as well as of the workcell as a whole. The equipment models and the intermediate specifications (T, PAC, and PACxp) for local control are introduced. Also, the dynamic specifications are used in order to create a flexible supervisory control system. The thickness and the PAC of photoresist are essential parameters in our photolithographic control scheme. Therefore, some of their important physical properties are briefly introduced as follows.

3.1.1 Some Physical Properties of the PAC and Thickness of Photoresist

The PAC (PhotoActive Compound Concentration, or inhibitor concentration) is a measure of the degree of exposure. For positive photoresist, the absorption of light destroys the inhibitor so that the PAC is decreased. The relation between PAC (M) and the exposure light intensity is given by the following equations [4]:

$$\frac{\partial}{\partial x} I(x,t) = -I(x,t)\,[AM(x,t) + B]$$

$$\frac{\partial}{\partial t} M(x,t) = -I(x,t)\,M(x,t)\,C$$

where $I(x,t)$ is the light intensity at depth x from the surface and exposure time t. $M(x,t)$ is the PAC at depth x and exposure time t. The constants A,B and C depend on photoresist material and exposure wavelength.

Development of positive resist can be described as a surface-limited etching reaction and the developing rate depends strongly on the PAC. F. H. Dill et. al. proposed an empirical relationship between the rate of development (R) and the PAC (M) based on experimental results [4]:

$$R = \exp\,(E_1 + E_2 \cdot M + E_3 \cdot M^2)$$

where E_1, E_2, and E_3 are fitting constants that change with different photoresist. E_3 is usually

negative. A typical relation between developing rate and PAC is shown in Figure 3.2. We can see that the PAC (M) is a suitable parameter to include in the equipment models since it is a good indicator of exposure energy as well as an important factor affecting the CD.

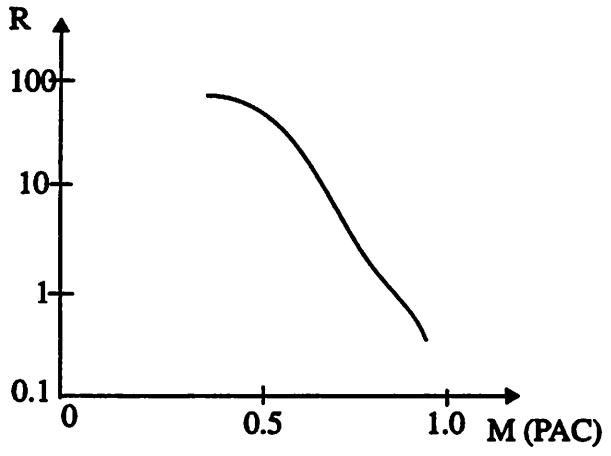


Figure 3.2 An Approximate Curve of The Relation Between Developing Rate And M (PAC) (Graph Based on a Figure From [4])

Photoresist is a semi-transparent thin film. In practical applications it is usually exposed with a stack of underlying layers (e.g., polysilicon and oxide layers) and this makes things more complex. Inside the photoresist, the interference between incident light and reflected light results in a standing wave of the exposure light intensity [5], as shown in Figure 3.3.

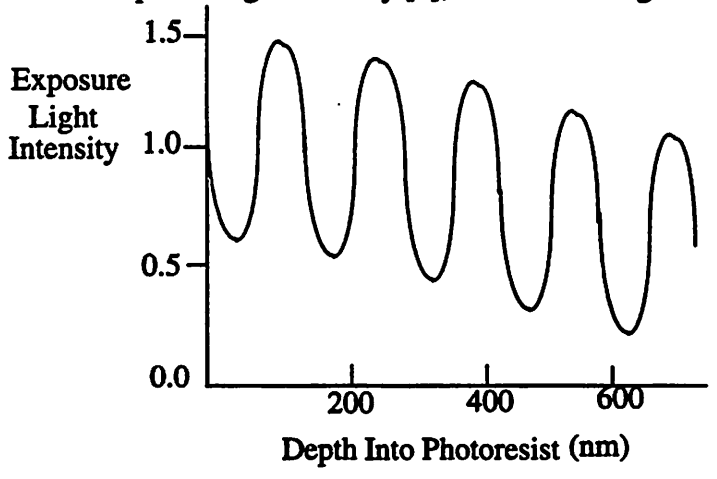


Figure 3.3 Exposure Light Intensity Within a Photoresist Film on Bare Silicon at the Beginning of Exposure (Graph Based on A Figure From [5])

The thin film properties of photoresist help to explain the significant effect of thickness on the CD and the way thickness must be included in the CD model. Therefore, photoresist thickness is a very important parameter in the control of CD on a photolithographic workcell.

3.2 The Measuring System

The goal of the new measuring system is to provide automated and reliable thickness and PAC measurements *in-situ*. To accomplish this goal, a new measuring tool, the PR (PhotoResist) Inspector, has been used. Also, a set of programs to compute PAC from measured data has been developed at the BCAM research group [6].

The PR (PhotoResist) Inspector [7] is an instrument used to obtain the thickness and other important properties of a thin-film substance, especially photoresist. It can make in-situ measurements during the normal production of IC wafers. It consists of three major parts: (1)the fiber optics/lens probe module, (2)the inspector module, and (3) the data station.

The fiber optics/lens probe module includes a fiber optic cable and a lens probe assembly. The function of the fiber optic cable is to transmit incident light from the light source to the wafer, and reflected light from the wafer to the detector. It is protected by a stainless steel shield and bifurcated to allow light transmission in both directions. The lens probe is used to collimate light onto the sample and collect reflected light. The inspector module contains a filtered 38,000 lux quartz halogen illuminator and a multi-channel spectrometer. The spectrometer, which includes a 512-element photodiode array, is used to receive the reflected light signals collected by the fiber optics/lens probe module, and transform these signals into reflected light intensity spectrographs. The data station consists of a 486 computer system. In addition to storing and managing programs and data of samples and setup conditions, the data station acquires data from the inspector module to provide fast and accurate thickness computation.

The working principle is briefly described as follows. A sample consisting of a thin film and underlying material is illuminated with white light. The light reflected from the top surface of the film combines with the reflected light from the underside of the film to create interference resulting in maxima and minima in the reflected light intensity spectrograph, from which the

thickness is computed by the PR Inspector. PAC can also be derived from this spectrograph.

The PR Inspector does not provide the PAC measurement, but it creates spectrographs from which PAC can be computed. In the BCAM group, Sovarong Leang developed a set of programs "HITIME" [6] that can compute PAC and thickness from the spectrograph. HITIME optimizes the parameters (thickness, PAC, and index of refraction) to fit a theoretical curve to the measured curve. The optimization of thickness and Na is done in a long wavelength range (about 500 ~ 620 nm). The optimization of PAC and Na is done in a short wavelength range (about 350 ~ 380 nm). Na is the constant and dominant term in the empirical expression for index of refraction:

$$n \text{ (index of refraction)} = Na + Nb/\lambda^2$$

Shown in Figure 3.4 is an example of reflected light spectrograph. Both the measured curve and the optimized theoretical curve are shown.

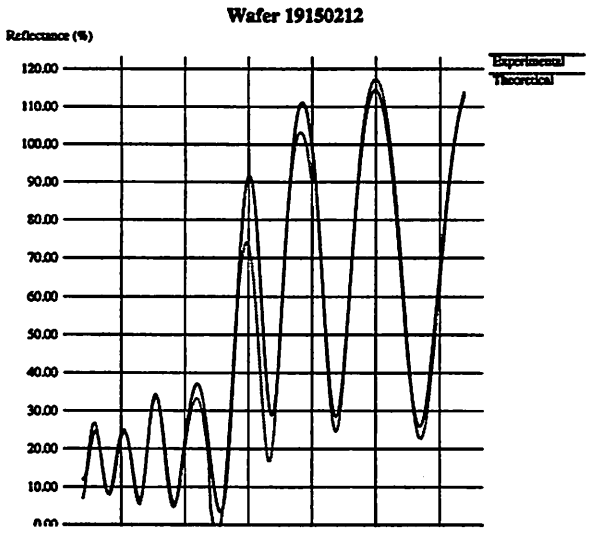


Figure 3.4 A Spectrograph Measured By The PR Inspector and The Fitting Curve by HITIME

Thin film thickness is computed by the PR Inspector as:

$$Th = \frac{\lambda_{max} \times \lambda_{min}}{4\,(\lambda_{max} \times n_{max} - \lambda_{min} \times n_{min})}$$

where λ_{max} and λ_{min} are the wavelengths at which reflected light intensity is the local maximum and minimum, respectively. Similarly, n_{max} and n_{min} represent the indices of refraction at which

reflected light is the maximum and minimum, respectively. It is assumed that the refractive index is a function of wavelength, and that it can be approximately computed by the following empirical relation:

$$n = Na + (Nb)/\lambda^2 + (Nc)/\lambda^4$$

where Na, Nb and Nc are fitting constants.

This empirical fitting equation is a little different from that of HITIME and the Nanospec, a widely accepted thickness-measuring tool in the IC industry. On the Nanospec, n is approximated by:

$$n = Na + Nb/\lambda^2$$

In order to match the thickness measured by the PR Inspector and the Nanospec, we calibrated the PR Inspector to find suitable Na, Nb and Nc. The result is in Chapter 4.2. It is shown that the PR Inspector can produce thickness values very close to the Nanospec across a wide range of thickness.

3.3 The BCAM Control System

The BCAM (Berkeley Computer Aided Manufacturing) system [8] is a powerful and still growing computer-integrated manufacturing environment developed at the BCAM research group at the EECS department of UC Berkeley. It includes and integrates many functions such as feed-back and feed-forward control, model-based SPC and simulation within a workcell, regression equipment model editing and update, recipe generation, malfunction alarm and control alarm generation, process/equipment diagnosis, various plotting capabilities, and interfaces to other BCAM applications. The BCAM is written in C++ and features user-friendly X window display. Figure 3.5 (courtesy of Bart Bombay) depicts the BCAM environment.

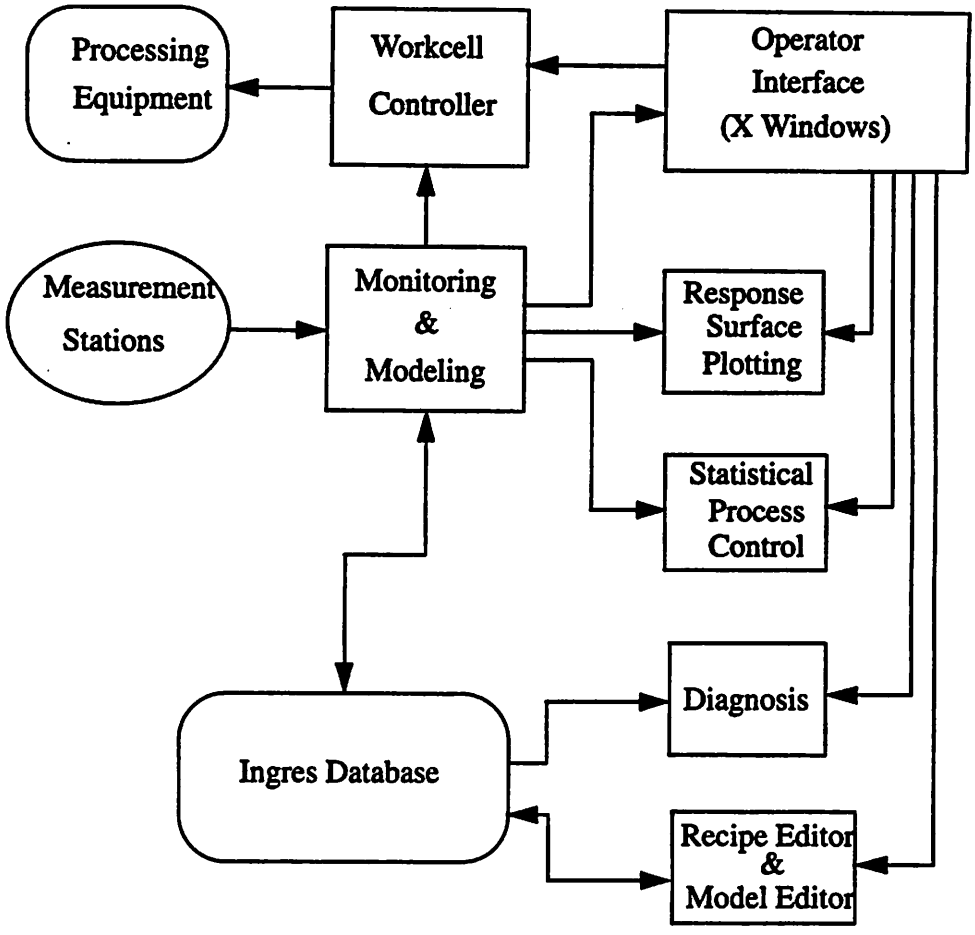


Figure 3.5 The BCAM Environment (courtesy of Bart J. Bombay)

The BCAM environment makes use of the Ingres database program which allows multiple users to share recipes and models. The Ingres library functions are used to store and retrieve equipment models and recipes in the database.

In this project the BCAM environment is mainly used to provide control/malfunction alarms and local feedback control(including statistical tests, equipment model update and editing, and recipe generation and update). The dynamic specification feature is not yet part of the BCAM environment. For the examples described in this document, all the dynamic spec generation was

done outside the BCAM environment.

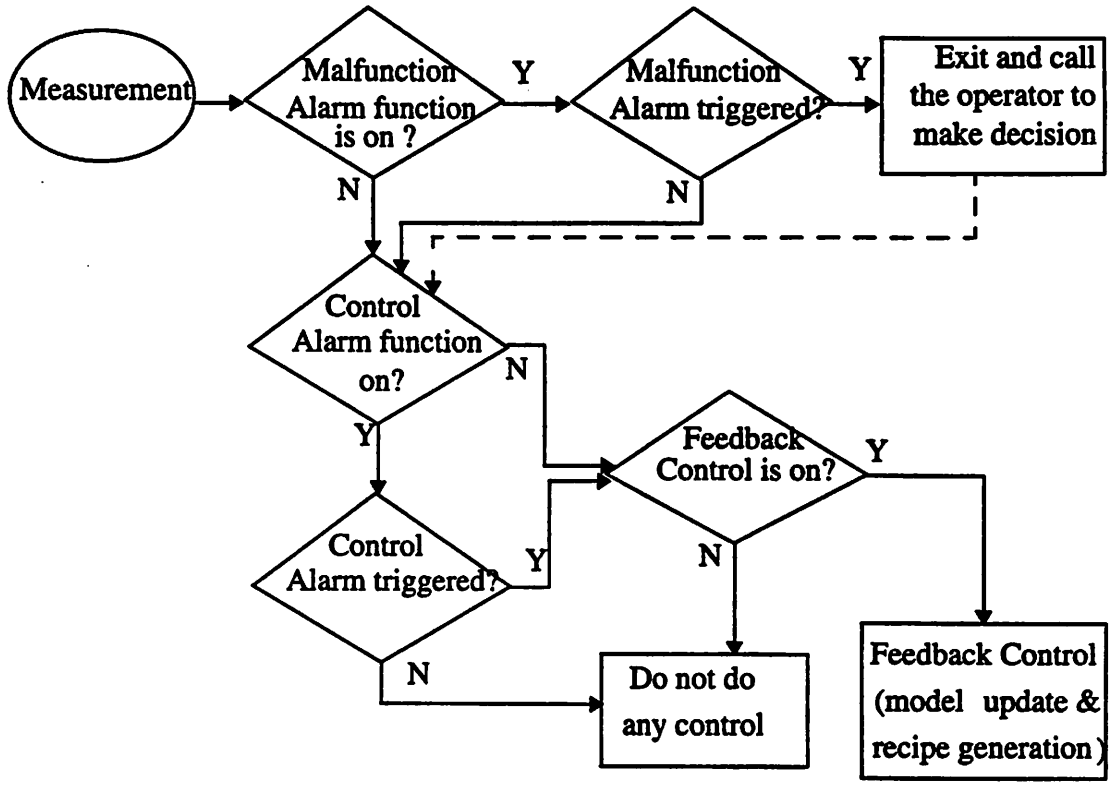


Figure 3.6 The Decision-making Process of Doing Feedback Control on the BCAM Environment

If the process has drifted significantly (as determined by the control alarm system), the equipment model needs to be updated. The model update algorithm performs statistical regressions to determine the optimum correction to the model based on historical records. For multiple outputs of a machine, the model update is done separately for each output since one model only describes one output parameter. In order to give more weight to the latest measurement, a forgetting factor is assigned to earlier data. In addition, a "window size" can be assigned, determining the total number of observations to be included in the model update calculations.

The recipe generation function can be used to generate the input settings which can produce a desired output as predicted by the equipment model. Usually the recipe update happens under

three conditions: the target change, the model update (in feedback control) and compensations in feed-forward control. Unlike the model update algorithm, the recipe generation algorithm does not need historical records. In the application of feedback control, if the process has drifted and the control alarm is triggered, the equipment model is updated to reflect the current condition of the process. During feedback control, the recipe is updated to bring the output near the target when the model is updated. The recipe can be updated due to a target change at any time.

During feed-forward control, the output of one step is used to predict the output of the next step. If significant difference between the model prediction and the output target of the next step is detected, feed-forward control is triggered and the recipe of the next step is calculated by the recipe generation algorithm. In this case, neither the target nor the model of the next step is changed. However, the measured output of this step (also the upstream input to the next step) is known such that the recipe (i.e., the controllable input settings) of the next step can be calculated accordingly. Detailed algorithms of model update and recipe update are described in Bart J. Bombay's work [8].

The malfunction and control alarm system has been developed and integrated into the BCAM environment. A malfunction alarm is a model-based SPC scheme used to detect large, abrupt shifts which may result from sudden abnormality of the process or the equipment. No historical records are required to generate a malfunction alarm. For the equipment with only one output (e.g., the CD of the developer), a model-based Shewhart chart is used to handle the malfunction alarm. If the difference between the measured value and the model prediction is larger than a certain number (usually 3) times sigma (the standard deviation of the process), a malfunction alarm is triggered. For a machine with multiple correlated outputs (like thickness and PAC of the photoresist coater), the T-square statistic is defined as follows:

$$T^2 = n\,(x_{measure} - x_{model\text{-}pred})^{transpose}\, S^{-1}\,(x_{measure} - x_{model\text{-}pred})$$

where S is the variance-covariance matrix and x is the vector including all outputs x_i's. A malfunction alarm is triggered when the T^2 exceeds a certain value, defined at the appropriate type I error of the T^2 distribution.

A control alarm is designed to detect smaller, systematic shifts which indicate that the process may be drifting out of control. A multivariate, model-based CUSUM (Cumulative-Sum) chart is suitable for the control alarm since it accumulates and manifests small shifts. Unlike the malfunction alarm, historical records are required in the control alarm system. In the single-output case, a control alarm is triggered whenever either Sh[n] or Sl[n] exceeds a certain value h. Sh[n] and Sl[n] can be calculated as follows:

$$Sh[\,initial\,] = Sl[\,initial\,] = 0$$

$$Sh[\,n\,] = max(\,0.0,\ x_{measure} - (\,x_{model\text{-}pred} + b\,) + Sh[\,n\text{-}1\,]\,)$$

$$Sl[\,n\,] = max(\,0.0,\ (\,x_{model\text{-}pred} - b\,) - x_{measure} + Sl[\,n\text{-}1\,]\,)$$

where b is a number proportional to sigma (the standard deviation of the process). From the definitions of Sh and Sl we notice that they both are cumulative, so they must be reset-to-zero after an alarm. Sh [n] increases when most measured data are significantly larger than the model prediction until a control alarm is triggered. On the other hand, Sl [n] detects consistently low measured values.

For multiple-output equipment, the control alarm requires not only the historical records but also the variance-covariance matrix and the standard deviation vector of the outputs. Only one parameter Y[n] is computed from the measured data, historical records, variance-covariance matrix and sigma (the standard deviation vector) of the multiple outputs. The algorithm is complex and not shown here. A control alarm is triggered if Y[n] > h , and h is selected in order to achieve the proper ARL of this scheme [9].

During feedback control, the control actions (model update and recipe update) are enabled by a control alarm. A malfunction alarm does not trigger feedback control. Rather, the process is stopped and the operator is notified.

3.4 Summary of The Supervisory Control System

Given the photolithographic workcell, the measuring system and the BCAM control/alarm environment, the infrastructure of the supervisory control system can be constructed. The logical,

or conceptual connections are shown in Figure 3.7. The physical connections of machines are presented in Figure 3.8. In Figure 3.8, the neural network stepper calibration belongs to another project in the BCAM research group, thus it is not applied to this project. The SVG computer linking the SVG coater and the SVG developer is not employed in this project, either. Within this infrastructure, several experiments have been done to demonstrate dynamic specifications in supervisory control. The experimental results are presented next, in chapters 4 and 5.

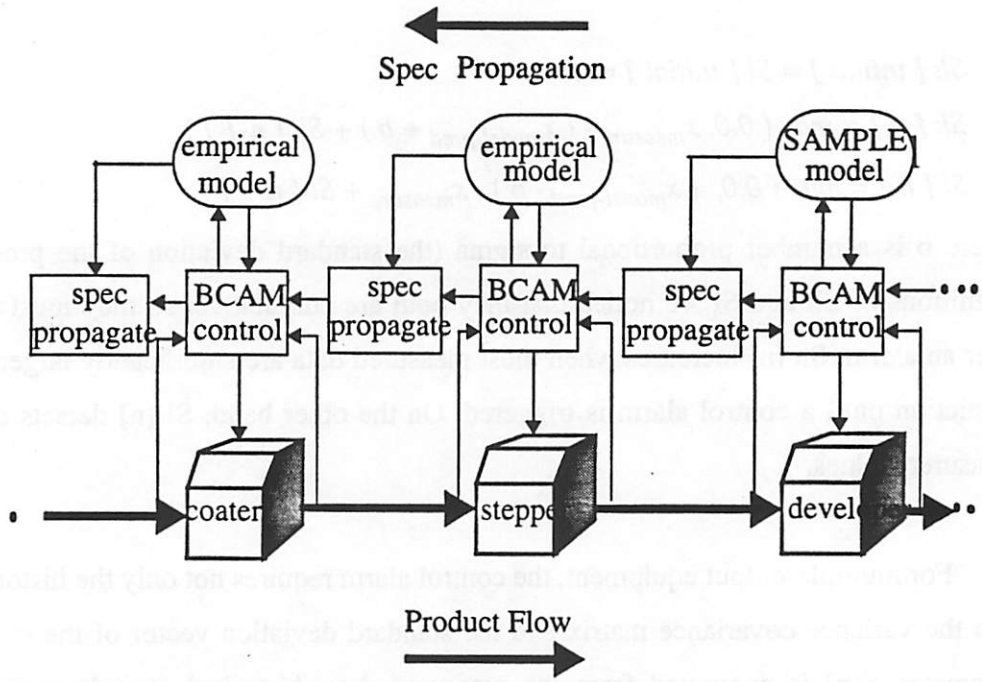


Figure 3.7 Logical Connections of the Supervisory Control System

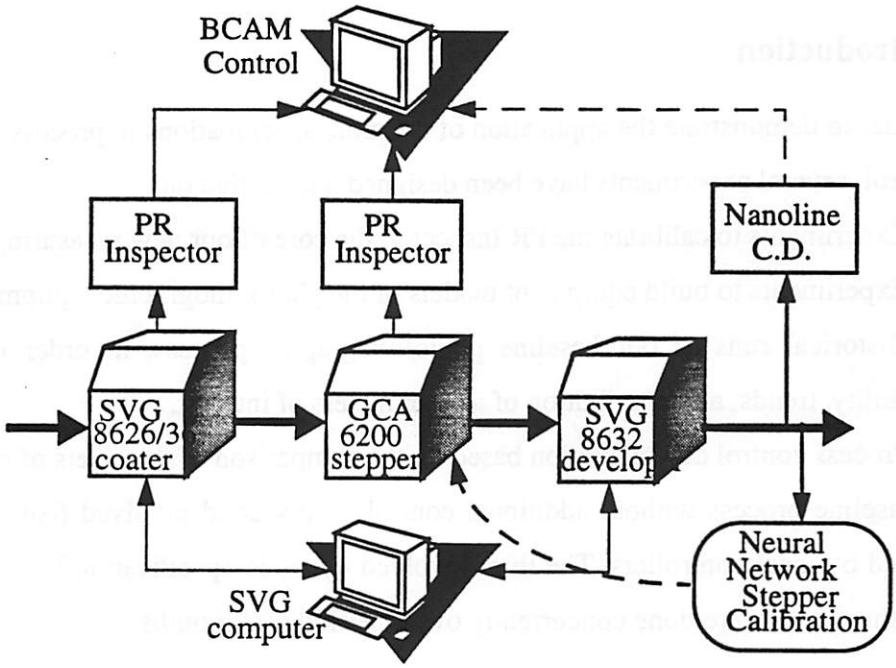


Figure 3.8 Physical Connections of Equipment

Chapter 4 Experimental Results and Discussion: Part I

4.1 Introduction

In order to demonstrate the application of dynamic specifications in process synthesis and process control, several experiments have been designed and carried out:

(1) Experiments to calibrate the PR Inspector, the core of our new measuring system.

(2) Experiments to build equipment models of the photolithographic equipment.

(3) Historical runs of our baseline photolithographic process, in order to obtain the process capability, trends, and distribution of all parameters of interest.

(4) Process control demonstration based on the comparison of three sets of runs. The first set used a baseline process without additional control. The second involved fixed-specification runs controlled by local controllers. The third involved dynamic-specification based supervisory control. The three runs were done concurrently over a period of 3 months.

The design of these experiments, experimental setup, as well as the results and discussion are described in the following two chapters. Experiments (1), (2), and (3) are included in Chapter Four. Experiments (4) and (5) are described in Chapter Five.

4.2 Experiment #1: Calibration of the PR Inspector

The PR Inspector is a novel instrument capable of measuring photoresist thickness in-situ. However, it is not as widely accepted for thin-film monitoring as Nanospec from Nanometrics. Thus, Nanospec was used as a reference to calibrate the PR Inspector.

Thin film thickness is computed by the PR Inspector as:

$$Th = \frac{\lambda_{max} \times \lambda_{min}}{4\,(\lambda_{max} \times n_{max} - \lambda_{min} \times n_{min})}$$

where λ_{max} and λ_{min} are the wavelengths at which reflected light intensity is at the local maximum and minimum, respectively. Similarly, n_{max} and n_{min} represent the indices of refraction at which reflected light is at the maximum and minimum, respectively. It is assumed that the

refractive index is a function of wavelength, and that it can be approximately computed by the following empirical relation:

$$n = Na + (Nb)/\lambda^2 + (Nc)/\lambda^4$$

where *Na, Nb* and *Nc* are fitting constants.

The purpose of calibration is to find suitable *Na, Nb* and *Nc* around the thickness of interest (*Na, Nb* and *Nc* may be different at two different film thicknesses) in order to match the PR Inspector readings to that of the Nanospec. The best way to accomplish this is by measuring the refractive index (n) at many different wavelengths and by curve-fitting the above empirical equation to get *Na, Nb*, and *Nc*. Unfortunately, the required instrument to do this was unavailable. Instead, a large amount of measurements were taken by both the PR Inspector and the Nanospec. By trial and error, a set of *Na, Nb*, and *Nc* was obtained that gives very good match between the PR Inspector and the Nanospec across the required thickness range (from about 11700 Å to 13700 Å). The results are shown below:

$$Na = 1.88, Nb = 1.1 \times 10^6, Nc = 1.918 \times 10^{13}$$

After these values were set, ten wafers with ten different thicknesses were measured by both the Nanospec and the PR Inspector. On each wafer, measurements were taken at six different sites. The location of the six sites is shown on a wafer map (Figure 4.1). All measured data were included to compare the thickness values measured by the Nanospec and the PR Inspector (Figure 4.2).

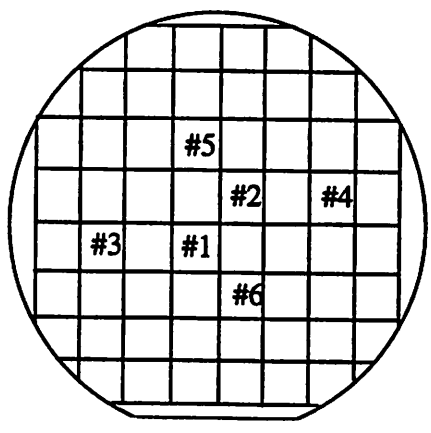


Figure 4.1 Wafer Map of Sites where Measurements were Taken (4" wafer)

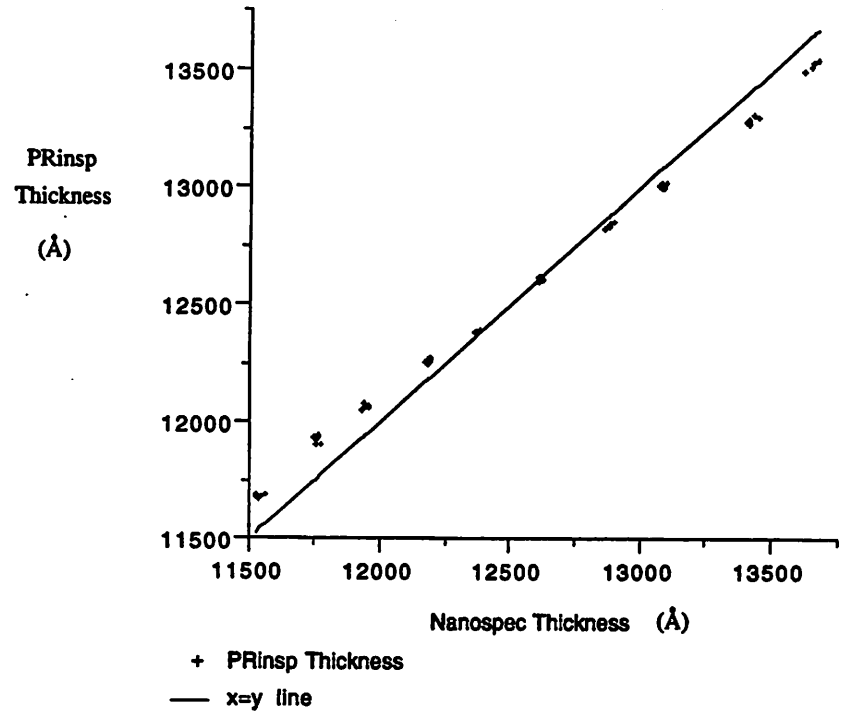


Figure 4.2 Comparison of Thickness Measured by PR Inspector and Nanospec

4.3 Experiment #2: Building Empirical Equipment Models

In order to build effective equipment models for process control, it is desirable to determine a minimal set of input/output parameters including all factors which significantly affect the process. For the photolithographic sequence, we have selected the following input/output parameters for the model of each equipment:

(1) For the photoresist coater (wafer track), the input parameters include (a) spin speed, (b)bake temperature, and (c) bake time. These are all user-controllable inputs. The output parameters include photoresist thickness and photoactive compound concentration (PAC).

(2) For the wafer stepper, the input parameters include photoresist thickness, PAC, and exposure dose. Only the exposure dose is a controllable input; the other two parameters are "upstream inputs" which are also the outputs from the upstream machine (wafer track) and they are not directly user-adjustable. The output parameters are photoresist thickness and after-

exposure-PAC (PACxp).

(3) For the developer track, the input parameters are photoresist thickness, PAC, PACxp, and develop time. Again, only develop time is a controllable input, and the remaining three are upstream inputs. The output of the developer track and the final output of the whole photolithographic sequence, is the critical dimension (CD).

Shown in Figure 4.3 is the photolithographic workcell containing these parameters.

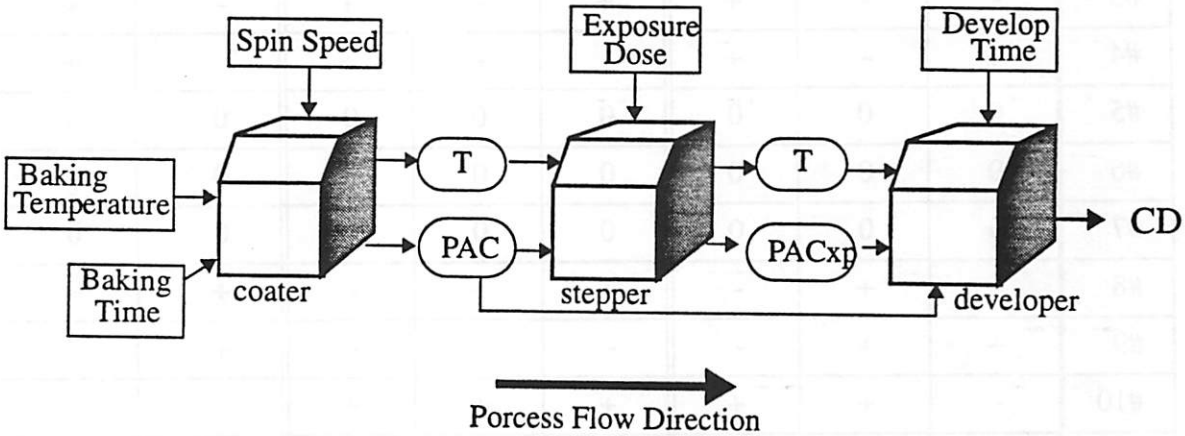


Figure 4.3 The Inputs and Outputs of the Machines within a Photolithographic Workcell

4.3.1. Design of Experiment for Building Equipment Models

The experimental design for the entire photolithographic sequence (including three pieces of equipment) is shown in Tables 4.1, 4.2 and 4.3. Table 4.1 describes an economical combination of three factorial experiments (one for each machine) each including three input variables. This experiment also includes three repetitions at the center point (standard setting). Eleven wafers were included in this basic part of experimental design. The symbol "+" represents high values, "0" for standard setting, and "-" for low values.

The actual values for all controllable inputs are listed in Table 4.2. The values were chosen so that the models can be valid within the wide range of the inputs. Table 4.1 was a minimal yet complete experimental design for this type of input combination. However, in order to accommodate second order terms across the whole range of the inputs, additional wafers with emphasis on

finding effects of individual inputs were added. This part of the experiment is shown in Table 4.3. The two sets of experiments together form an approximate central composite experiment.

Table 4.1 Design of Experiment for Model Building (Part 1)

wafer	SPS†	BTE†	BTI†	THK*†	PAC*	DOSE	PACxp*	DevTime
#1	-	-	-	+	-	-	+	+
#2	+	-	-	-	-	-	+	+
#3	-	-	+	+	-	+	-	+
#4	+	-	+	-	-	+	-	+
#5	0	0	0	0	0	0	0	0
#6	0	0	0	0	0	0	0	0
#7	0	0	0	0	0	0	0	0
#8	-	+	-	+	+	-	+	-
#9	+	+	-	-	+	-	+	-
#10	-	+	+	+	+	+	-	-
#11	+	+	+	-	+	+	-	-

* Since these parameters (THK, PAC, PACxp) were not set directly by the user, the "+" "-", and "0" indication are only approximate.

† "SPS" = Spin Speed, "BTE" = Bake Temperature, "BTI" = Bake Time, "THK"=Thickness

Table 4.2 Values of the Inputs (from Table 4.1)

	SPS (rpm)	BTE (°C)	BTI (sec)	THK* (Å)	PAC* (no unit)	Dose (mj/cm^2)	PACxp* (no unit)	DevT (sec)
-	3600	75	20	11900	0.935	105.5	0.26	50
0	4600	90	60	13100	0.96	167	0.28	60
+	5600	105	100	15000	0.98	246	0.38	90

*Approximate values of the uncontrollable inputs

Shown in Table 4.3 is the experimental design of additional wafers processed for equipment model building. Thirty three wafers were included in this part (data courtesy of Sovarong Leang). For the entire modeling experiment, a total of forty-four (11 + 33) wafers were used. On each wafer, six measurements were made for each output parameter. The average of the six measurements on one wafer was taken as one data point. Therefore, forty-four data points were used to fit each equipment model. The location of the six measurements is shown on the wafer map (Figure 4.4).

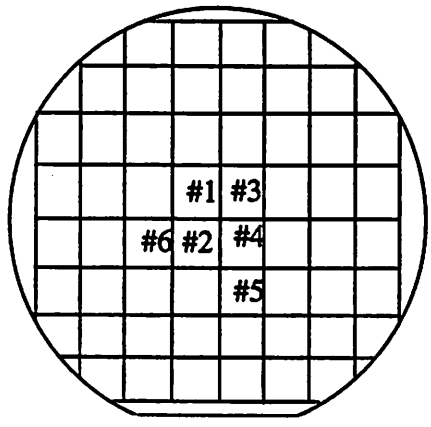


Figure 4.4 Location of Measurements for Experiments to Build Equipment Models

Table 4.3 Design of Experiment for Model Building (Part II)

	SPS(rpm)	BTE(°C)	BTI(sec)	Dose(mj/cm^2)	DevT(sec)
#12	4600	74.6	60	76	75
#13	4600	78.1	60	94	75
#14	4600	81.5	60	112	75
#15	4600	84.4	60	131	75
#16	4600	87.4	60	149	75
#17	4600	90.4	60	168	60
#18	4600	93.5	60	186	60
#19	4600	96.5	60	204	60

	SPS(rpm)	BTE(°C)	BTI(sec)	Dose(mj/cm^2)	DevT(sec)
#20	4600	99.5	60	223	60
#21	4600	102.5	60	241	60
#22	4600	105.5	60	260	60
#23	4600	90	20	89	75
#24	4600	90	30	107	75
#25	4600	90	40	127	75
#26	4600	90	50	147	75
#27	4600	90	60	167	75
#28	4600	90	70	187	75
#29	4600	90	80	207	75
#30	3600	90	60	89	60
#31	4000	90	60	107	60
#32	4300	90	60	127	60
#33	4600	90	60	147	60
#34	4900	90	60	167	60
#35	5200	90	60	187	60
#36	5600	90	60	207	60
#37	4600	90	60	168	60
#38	4600	90	60	168	60
#39	4600	90	60	168	60
#40	4600	90	60	168	60
#41	4600	90	60	168	60
#42	4600	90	60	168	60
#43	4600	90	60	168	60
#44	4600	90	60	168	60

4.3.2 Model Development

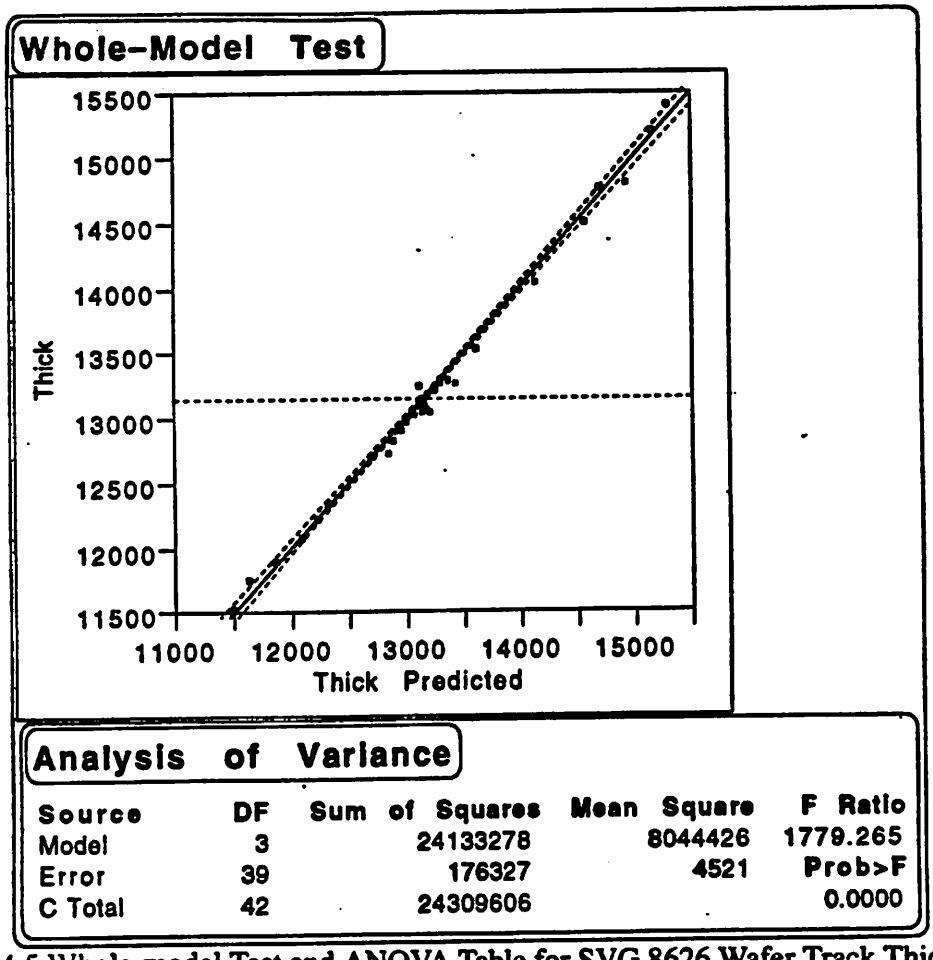
After the experiment was completed, the data were analyzed by using a statistical tool "JMP" installed on a power PC Apple computer. In determining the terms of the equipment models, some physical considerations were made. For example, we learned that thickness is about proportional to the inverse square root of the spin speed. Also, the critical dimension can be better modeled if a cosine-of-thickness term is included. After these transformations were made, linear regression analysis was used to fit the models.

Models of SVG 8626 Wafer Track

Two models, one for thickness and one for PAC, were derived for the SVG coater (wafer track). During model fitting, the following parameter transformation were used:

SPS_normal = SPS/1000, (SPS in revolution per minute)

BTE_normal = BTE/100 (BTE in °C)



Source	DF	Sum of Squares	Mean Square	F Ratio
Model	3	24133278	8044426	1779.265
Error	39	176327	4521	Prob>F
C Total	42	24309606		0.0000

Figure 4.5 Whole-model Test and ANOVA Table for SVG 8626 Wafer Track Thickness Model

As shown in Figure 4.5, the thickness model can predict the measured data very well. A very high F-ratio of 1779 was obtained, which meant most variation could be explained by the model. In Figure 4.6, four tables, including summary of fit, lack of fit, parameter estimates and effect test of thickness are presented. From the "parameter estimates" table, the constant and coefficients as well as their statistical significance can be obtained.

Response: Thick

Summary of Fit

Rsquare	0.992747
Root Mean Square Error	67.23993
Mean of Response	13150.09
Observations (or Sum Wgts)	43

Lack of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack of Fit	27	170587.45	6318.05	13.2092
Pure Error	12	5739.69	478.31	Prob>F
Total Error	39	176327.14		0.0000

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	1291.9774	225.832	5.72	0.0000
1/SPS_normal^.5	29353.311	411.734	71.29	0.0000
BTI	-1.624123	0.52852	-3.07	0.0039
BTE_normal	-1948.854	125.98	-15.47	0.0000

Effect Test

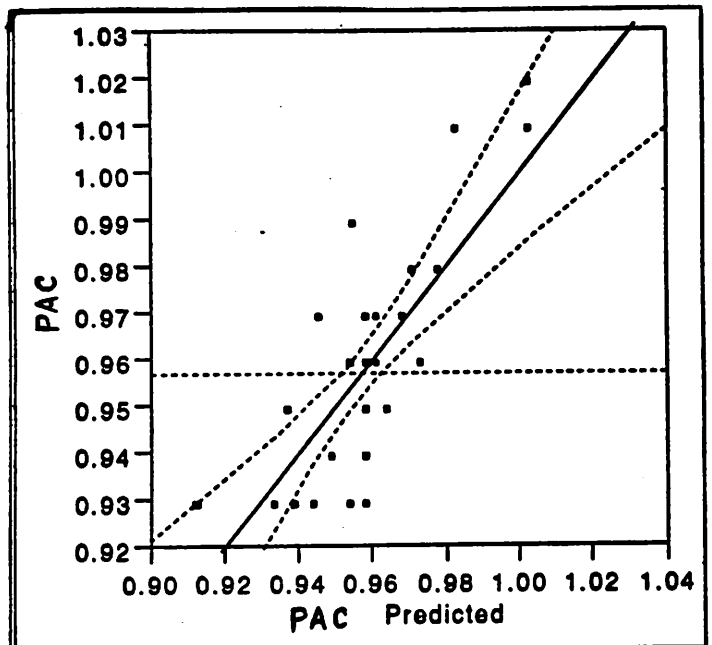
Source	Nparm	DF	Sum of Squares	F Ratio	Prob>F
1/SPS_normal^.5	1	1	22979205	5082.536	0.0000
BTI	1	1	42695	9.4432	0.0039
BTE_normal	1	1	1081958	239.3072	0.0000

Figure 4.6 Various Statistical Significance Tests for the SVG 8626 Wafer Track Thickness Model

From the above table, converting SPS_normal and BTE_normal back to SPS and BTE, the thickness model for the SVG 8626 wafer track was obtained as follows:

$$Thickness = 1291.98 + 928233 \times \frac{1}{\sqrt{SPS}} - 19.4885 \times BTE - 1.62412 \times BTI$$

In Figure 4.7, the whole model-test and ANOVA table for the PAC model of SVG 8626 wafer track are demonstrated. Although the fit is not as good as for the thickness model, the model is statistically significant.



Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	0.01160735	0.005804	24.3116
Error	37	0.00883265	0.000239	Prob>F
C Total	39	0.02044000		0.0000

Figure 4.7 Whole-Model Test and ANOVA Table for SVG 8626 Wafer Track PAC Model

In Figure 4.8, four tables, including summary of fit, lack of fit, parameter estimates and effect test of PAC are shown. Again, the constant and coefficients can be obtained from the "parameter estimates" table.

Response: PAC

Summary of Fit

Rsquare	0.567874
Root Mean Square Error	0.015451
Mean of Response	0.957
Observations (or Sum Wgts)	40

Lack of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack of Fit	18	0.00595487	0.000331	2.1842
Pure Error	19	0.00287778	0.000151	Prob>F
Total Error	37	0.00883265		0.0498

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	0.9092807	0.0348	26.13	0.0000
SPS_normal	-0.020977	0.00496	-4.23	0.0001
BTE_normal	0.1607003	0.02895	5.55	0.0000

Effect Test

Source	Nparm	DF	Sum of Squares	F Ratio	Prob>F
SPS_normal	1	1	0.00426619	17.8711	0.0001
BTE_normal	1	1	0.00735681	30.8177	0.0000

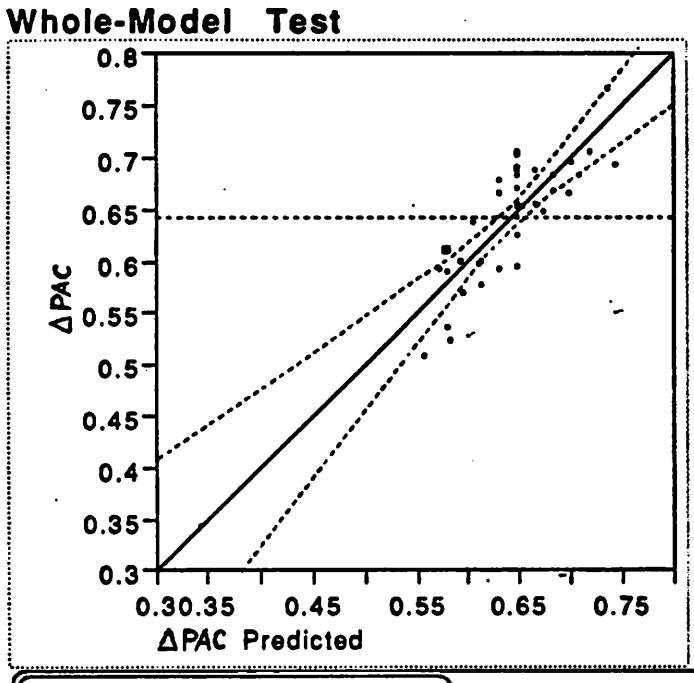
Figure-4.8 Various Statistical Significance Tests for SVG 8626 Wafer Track PAC Model

From the above table, converting SPS_normal and BTE_normal back to SPS and BTE, the model of PAC for the SVG 8626 wafer track was restated as follows:

$$PAC = 0.90928 - 2.0977 \times 10^{-5} \times SPS + 0.001607 \times BTE$$

Model for the GCA 6200 stepper

In Figure 4.9 the whole-model test and ANOVA table for ΔPAC are shown. ΔPAC is defined as ΔPAC = PAC - PACxp. The purpose of introducing ΔPAC is that it lead to more significant models. Once the model of ΔPAC is obtained, the model of PACxp can be obtained by simply moving the PAC term to the other side of the equal sign.



Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	0.07872493	.0.039362	34.0693
Error	37	0.04274844	0.001155	Prob > F
C Total	39	0.12147337		0.0000

Figure 4.9 Whole-Model Test and ANOVA Table for ΔPAC for the GCA 6200 Stepper

In Figure 4.10, four tables, including summary of fit, lack of fit, parameter estimates and effect test of ΔPAC are shown. Again, the constant and coefficients can be obtained from the "parameter estimates" table.

Response: ΔPAC

Summary of Fit

Rsquare	.6480838
Root Mean Square Error	.0339906
Mean of Response	0.642375
Observations (or Sum Wgts)	40

Lack Of Fit

Source	DF	Sum of Squares	Mean Square	F Ratio
Lack Of Fit	35	0.03656344	0.001045	0.3378
Pure Error	2	0.00618500	0.003093	Prob > F
Total Error	37	0.04274844		0.9351

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	.64294819	.108817	5.91	0.0000
dose	.00090945	.000123	7.34	0.0000
Thick	-.0000112	.000007	-1.48	0.1484

Effect Test

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
dose	1	1	0.06216812	53.8083	0.0000
Thick	1	1	0.00251769	2.1791	0.1484

Figure 4.10 Various Statistical Significance Tests for ΔPAC for the GCA 6200 Stepper

From the above table, by substituting ΔPAC with PAC-PACxp, the model of PACxp for the GCA 6200 stepper was obtained as follows:

$$PACxp = -0.64295 - 0.00090945 \times Dose + \left(PAC + 1.12 \times 10^{-5} \times Thickness\right)$$

Model for the SVG 8632 Develop Track

Fitting the CD model was more difficult, since linear model fitting alone did not produce a significant model. As stated before, it is necessary to include a term of the form cos(A*Thickness + B), where A and B are unknown constants. In order to obtain a significant model for CD, a process simulation program (SAMPLE) was used to compute a table of CD's based on 12,800 input combinations. After this, the computed CD's (by SAMPLE) were non-linearly fitted by using JMP. The result was a CD model including a constant term and terms of cos(A*thickness+B), thickness, PAC, exposure dose, and develop time. In this way we obtained values for the constant, coefficients, as well as A and B inside the cosine term. Since exposure dose is not a controllable input of the developer, it was replaced by PACxp, PAC, and thickness (from the model of PACxp). Finally, this SAMPLE-based model was compared to real measured data from our experiments. This comparison led to a new value for the constant term (SAMPLE data and nonlinear fitting data courtesy of Sovarong Leang).

Shown in Figure 4.11 is the result of nonlinear fitting based on 12,800 SAMPLE-measured CD's. This intermediate CD model was expressed as:

CD = a11 + a2*thickness + a3*cos(a4*thickness + a5) + a12 * dose + a9*PAC
+ a10*DevTime

Shown in Figure 4.12 is the scatter plot comparing the real measured CD from experiments and SAMPLE-calculated CD based on inputs from the experiments.

The CD model for the SVG 8632 develop track is as follows:

$$CD = 3.41735 + 0.358965 \times \cos\left(5.7125 \times 10^{-3} \times Thickness - 57.6204\right)$$
$$- 3.70536 \times 10^{-6} \times Thickness - 2.12452 \times PAC$$
$$+ 2.36433 \times PACxp - 2.518 \times 10^{-3} \times DevT$$

Solution

SSE	DFE	MSE	RMSE
16.337701583	12792	0.0012772	0.0357377

Parameter	Estimate	ApproxStdErr	Lower CL	Upper CL
a1	2.518273526	0	.	.
a2	0.0000227751	0.00000029	.	.
a3	0.0358964793	0.00044255	.	.
a4	0.0057125176	0.00001197	.	.
a5	-57.62038294	0.14543169	.	.
a6	-0.01188879	0	.	.
a7	0.0000453006	0	.	.
a8	-0.000000064	0	.	.
a9	0.2298082143	0.00403616	.	.
a10	-0.002518152	0.00001884	.	.
a11	1.8972053645	0.00543718	.	.
a12	-0.002150237	0.0000055	.	.

Figure 4.11 Nonlinear Fitting Results Based on SAMPLE-measured CD

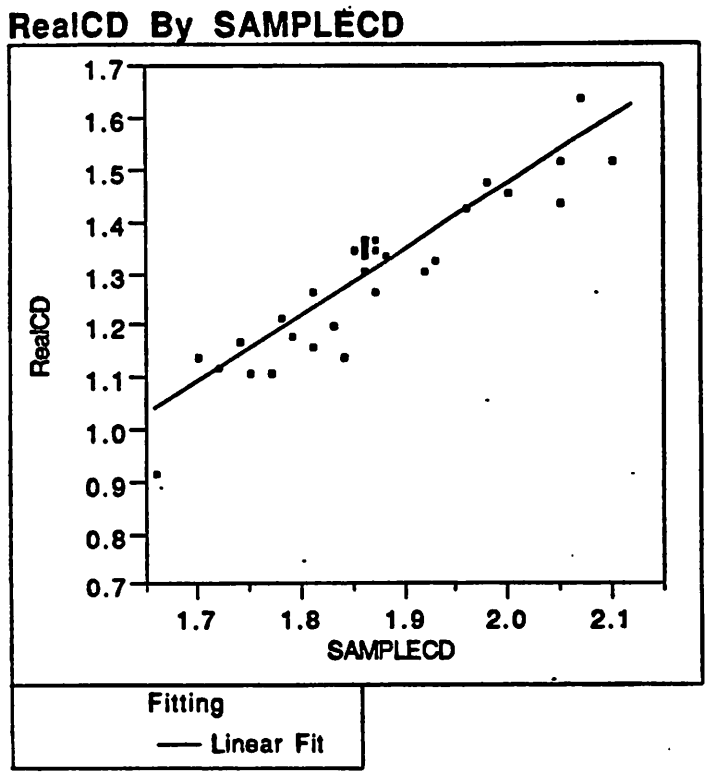


Figure 4.12 Comparison of Modelled CD and Real Measured CD from the SVG 8632

These equipment models are mainly empirical, response-surface models with the addition of more physically meaningful parameters (PAC), as well as certain knowledge of the photolithographic process. As a result, they are more efficient for control purposes than purely "physical" models like those used in Technology CAD process simulators, like SAMPLE.

There are some other factors that may affect the process but are not included in the models. For example, environmental parameters such as relative humidity and room temperature may have certain effects on thickness and PAC. They are not included in the models because they are not "controllable inputs" of the equipment. However, the sensors at the yellow-light room do provide the data of these factors. Over a long period, these environmental factors may drift; thus the trends of historical runs may be used to explore these effects. This will be discussed in the next section.

4.4 Experiment #3: The Baseline Historical Runs

In order to observe the process capability, including trends and distribution of the output parameters of our baseline process, several wafers were run over a period of time using standard settings. The data acquired are useful in determining suitable targets and control limits for the experiments on process control. These historical lots were simply run using the fixed recipe of our baseline process.

4.4.1 Experimental Setup

(1) Wafers: p-type, 4-inch silicon wafers with 1000 Å of oxide on the top

(2) Equipment: SVG8626/36 photoresist coater track, GCA 6200 wafer stepper, SVG 8632 developer track, SC PR Inspector (modified with Xe light source), and Nanoline (for CD measurement)

(3) Photoresist : OCG820 G-line positive photoresist

(4) Recipe : same as the recipe used for the baseline process at the microlab

SPS(rpm)	SPT(sec)	BTE(°C)	BTI(sec)	DOSE	DevT(sec)
4600	30	90	60	**	60

Note(1): "SPS" = spin speed, "SPT" = spin time, "BTE" = bake temperature, "BTI" = baketime, "DOSE" = exposure dose (in milijoule/cm^2), "DevT" = develop time.

**Note(2): Exposure dose was not always fixed. Since the lamp of the stepper was changed about every three weeks, a focus-exposure test was done after each lamp change to obtain an optimal set of focus and exposure values.

(5) Duration, frequency of runs, and amount of lots and wafers: The historical lots were run for about two months. The lapsed time between two consecutive lots ranged from three to six days, due to some practical constraints. There were a total of twelve lots. Each lot had five wafers from lot#1 to lot#7. Three wafers were included in each lot from lot#8 to lot#12.

(6) Measurements: Three readings of four output parameters were made on each wafer. The readings were taken at three different sites near the center of a wafer (as shown in Figure 4.13), and their average was treated as one measurement.

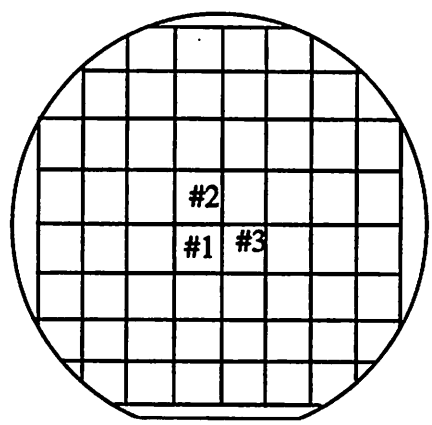


Figure 4.13 Location of Measurements Used in the Historical Runs

4.4.2 Trendcharts of the Four Output Parameters

Shown in Figures 4.14, 4.15, 4.17, and 4.18 are the trendcharts of the four output parameters: thickness, PAC, PACxp, and CD. Each point represents the average of three readings for one wafer.

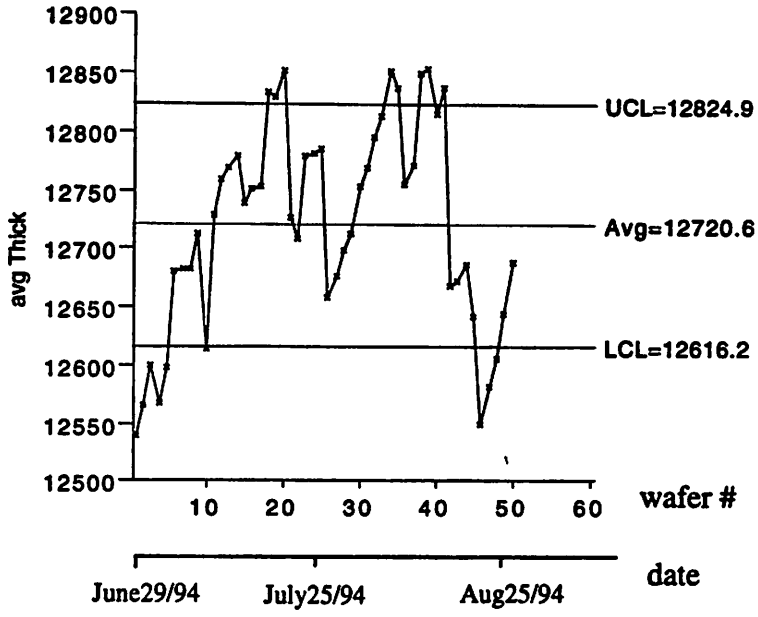


Figure 4.14 Trendchart of Photoresist Thickness

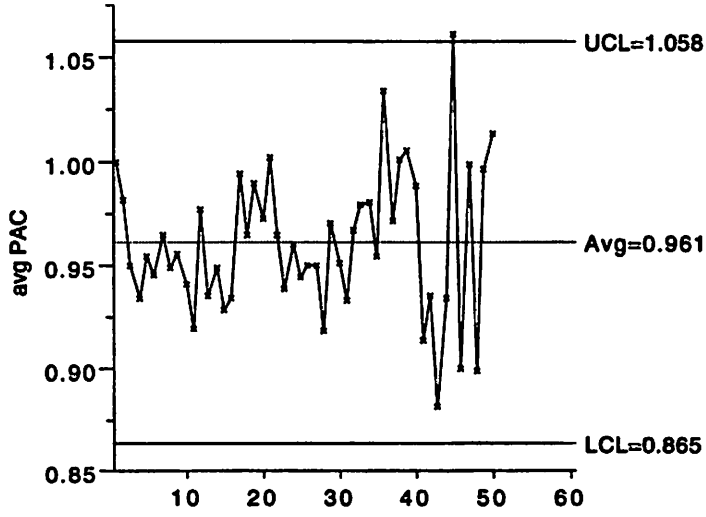


Figure 4.15 Trendchart of PAC of Photoresist

As shown in Figure 4.14, photoresist thickness exhibited an obvious drift even when the recipe of the SVG coater was fixed during the entire period. Two other factors, relative humidity and room temperature were recorded and compared against thickness. The scatter plots in Figure 4.16 show that neither of the two factors had a significant effect on thickness. However, this may be attributed to the relatively small range of variation for both humidity and room temperature since neither of these factors was intentionally varied. Past data showed that, at a relative humidity of 24%, thickness of about 13100 Å was measured, as opposed to thickness around 12700 Å at 38% humidity. A much longer series of historical runs is required to obtain a quantitative relation between humidity, room temperature, and thickness. Since relative humidity and room temperature could hardly explain thickness variation of the historical runs, there must have been some other less obvious factors. The trendchart of thickness well justified the need for additional process control.

Figure 4.15 shows no obvious trend of PAC, though considerable fluctuation can be seen.

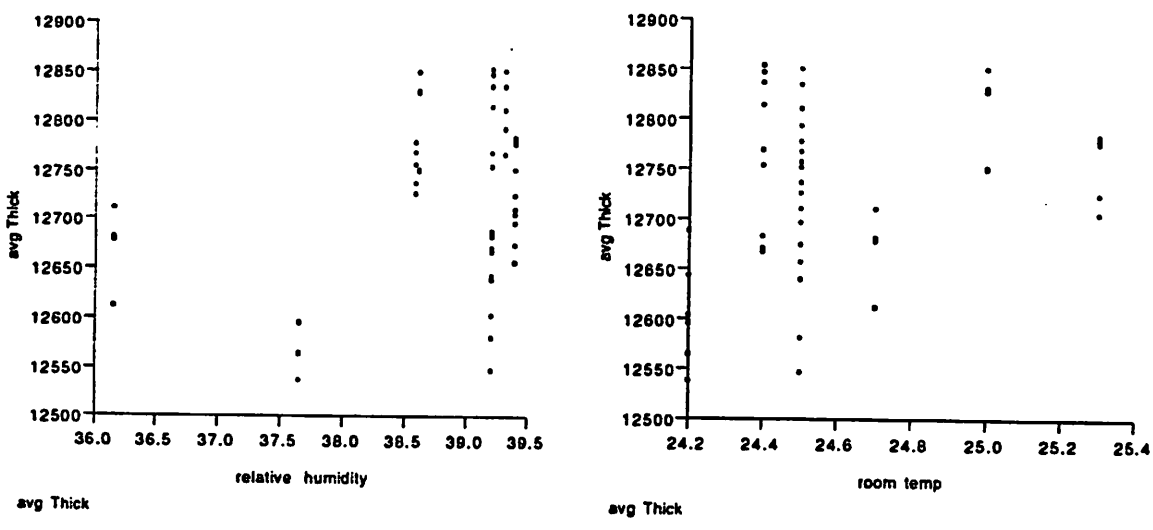


Figure 4.16 Effects of Relative Humidity and Room Temperature on Thickness

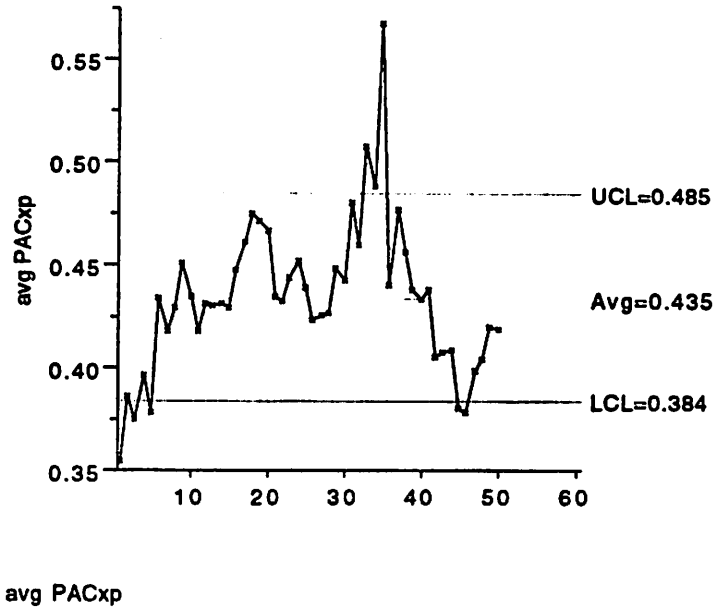


Figure 4.17 Trendchart of PACxp

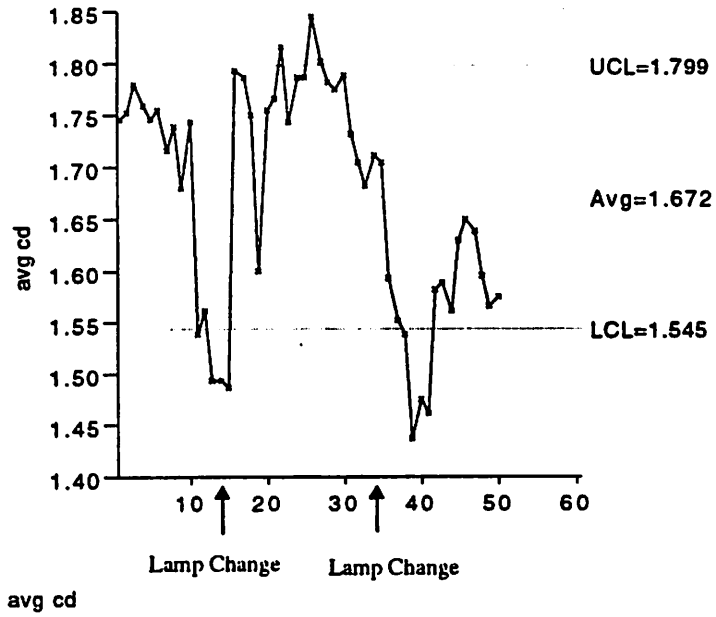


Figure 4.18 Trendchart of CD

Figure 4.17 shows that PACxp had a more obvious drift than PAC did. This may be attributed to two reasons: First, as shown in Figure 4.19, PACxp had a strong correlation with thickness, but PAC did not. When thickness drifted, PACxp followed. Second, the exposure dose was not actually fixed throughout the entire period. The lamp of the stepper was changed twice during the experimental period. Right after each lamp change, a routine focus-exposure test was done to get the optimal focus and exposure values for the baseline process. Since three different lamps were used during the experimental period, three different exposure doses were applied. The degradation due to lamp aging between two lamp changes was automatically compensated by the GCA wafer stepper itself.

This situation is further demonstrated by the trendchart of CD (Figure 4.18). The first lamp change happened between wafer #15 and #16, the second lamp change happened between wafer #35 and #36. The actual exposure dose of the third group (wafer #36 to #50) was a little higher than those of the first two groups and resulted in lower CD for the third group. The variation of CD within the third group could be well explained by thickness fluctuation. The only lot behaving strangely was lot #3 (wafers #11 to #15) which was not supposed to show such low values of CD, given the very similar process conditions to wafers #1 to #10.

4.4.3 Correlation Between the Output Parameters

Figure 4.19 shows that PACxp had a strong correlation ($\rho = 0.7941$) with thickness. In table 4.4, the correlation coefficients between output parameters are listed.

In addition to the correlation between thickness and PACxp, some correlation was observed between thickness and CD. It is interesting to compare this result with the data from another experiment which was designed to explore solely the effect of thickness on CD, with exposure dose and develop time fixed. In that lot, ten wafers were run through photolithographic sequences with ten different spin speeds, thus ten different thickness ranging from 12000 Å to 14000 Å (the thickness of historical runs ranged only from 12500 Å to 12850 Å). All other controllable inputs

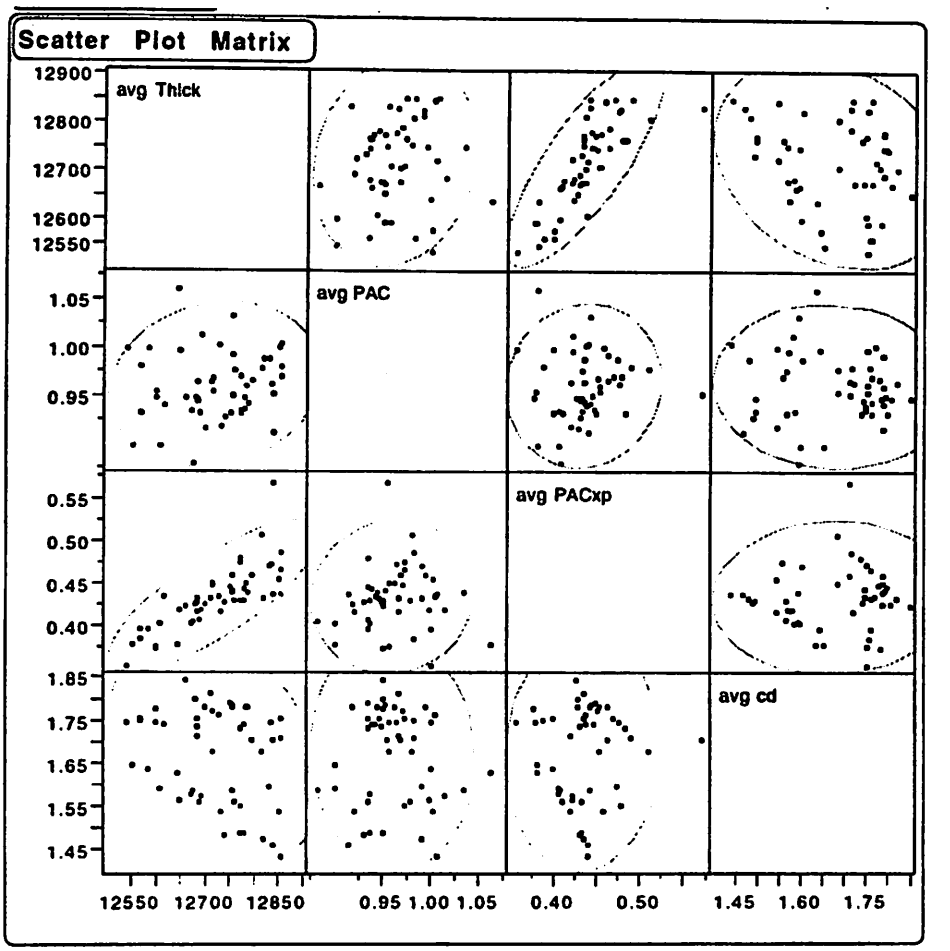


Figure 4.19 Scatter Plot Matrix for the Four Output Parameters

Table 4.4 Correlation Matrix of Output Parameters (Historical Run)

	Thickness	PAC	PACxp	CD
Thickness	1	0.1419	0.7941	-0.2572
PAC	0.1419	1	0.0637	-0.0507
PACxp	0.7941	0.0637	1	0.0426
CD	-0.2572	-0.0507	0.0426	1

were the same for all ten wafers. The result is shown in Figure 4.20. At this figure, CD decreased as thickness increased from 12500 Å to 12850Å, the range of variation of the historical runs. This explained the negative correlation between thickness and CD for historical runs. It also explained the CD variation of wafers #35 to #50 due to thickness fluctuation.

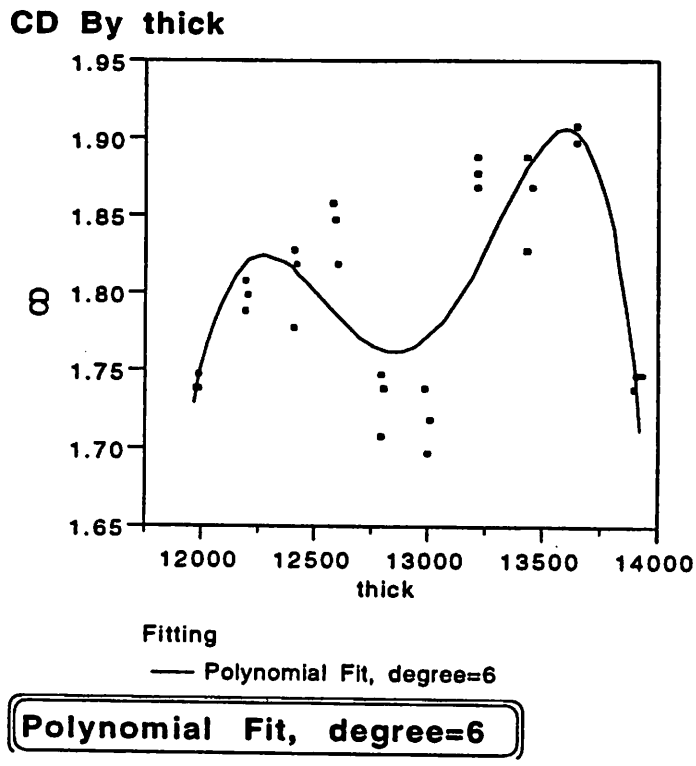


Figure 4.20 Effect of Thickness on CD (One point represents one measurement. Ten wafers were run. On each wafer three readings were made.)

Surprisingly the correlation between PACxp and CD during the historical runs was rather insignificant. It is assumed that PACxp is a good measure of exposure dose and should affect CD significantly. The reason might be as follows: During the historical runs, the exposure dose was relatively stable. Thus the limited variation of PACxp came mainly from thickness fluctuation, not a change of dose. For the factorial experiments, the exposure dose was changed over a very wide range; thus the distribution of PACxp was much wider. Thus the correlation between PACxp and CD was more apparent during the factorial experiment, as shown in Table 4.5

Table 4.5 Correlation Matrix of Output Parameters (data from the model-building factorial experiment, not historical runs)

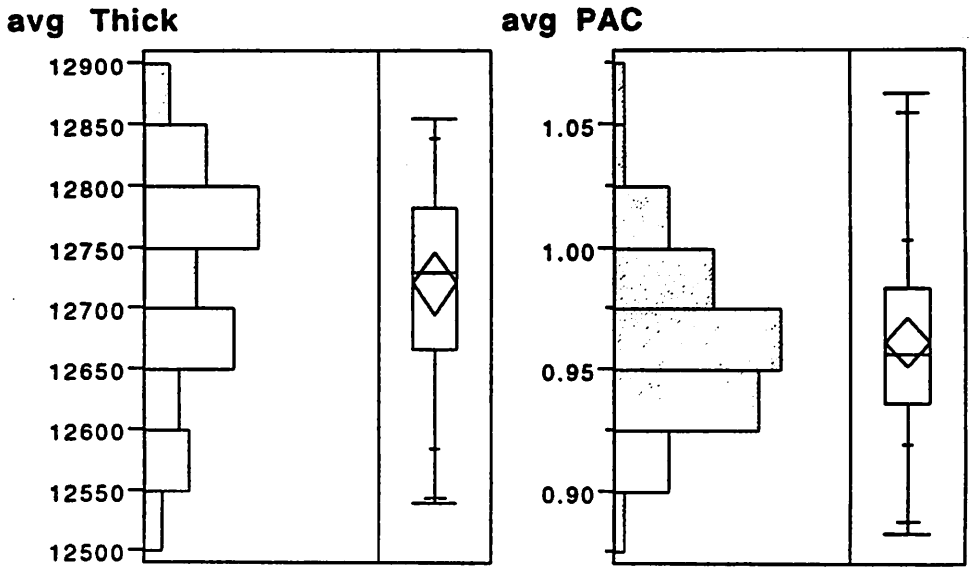
	Thickness	PAC	PACxp	CD
Thickness	1	0.3818	0.3729	-0.1598
PAC	0.3818	1	0.5082	-0.068
PACxp	0.3729	0.5082	1	0.2771
CD	-0.1598	-0.068	0.2771	1

4.4.4 Distribution of Output Parameters

The distribution (histograms), along with statistical moments and quantiles of the four output parameters are shown in Figure 4.21 and Figure 4.22. These data show the process capability of the baseline process at the Berkeley Microfabrication Laboratory. These histograms form a useful reference for setting specifications on process control and alarm systems.

4.5 Summary

In this chapter the experimental results of the PR Inspector calibration, the photolithographic equipment model generation, and the baseline historical runs are presented and discussed. These experiments are essential to construct an infrastructure for the supervisory control on a photolithographic workcell. In the next chapter experimental results of comparing different control methods to the one that uses dynamic specifications will be discussed.



Quantiles

maximum	100.0%	12855
	99.5%	12855
	97.5%	12855
	90.0%	12839
quartile	75.0%	12783
median	50.0%	12728
quartile	25.0%	12667
	10.0%	12584
	2.5%	12542
	0.5%	12540
minimum	0.0%	12540

Moments

Mean	12720.56
Std Dev	88.88
Std Err Mean	12.57
upper 95% Mean	12745.83
lower 95% Mean	12695.30
N	50.00
Sum Wgts	50.00

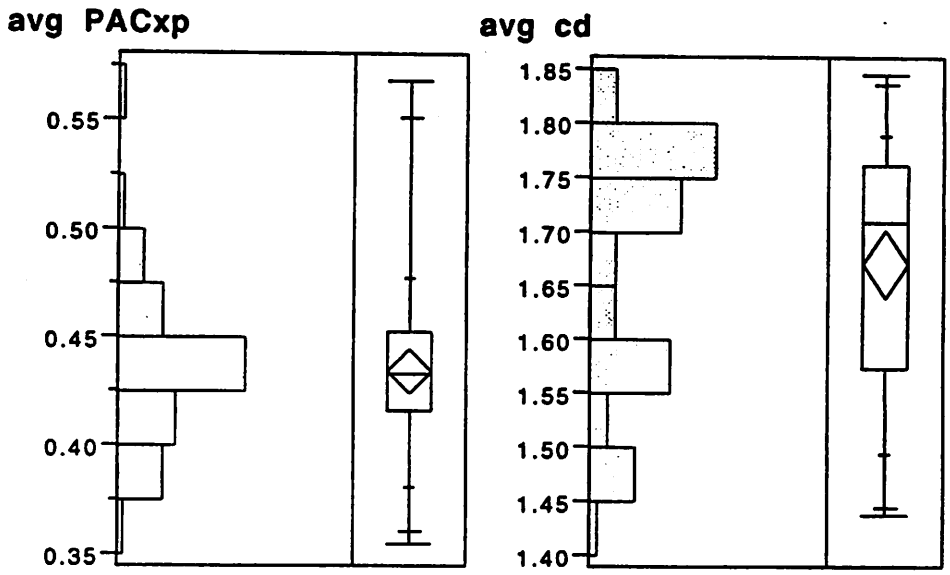
Quantiles

maximum	100.0%	1.0623
	99.5%	1.0623
	97.5%	1.0546
	90.0%	1.0031
quartile	75.0%	0.9838
median	50.0%	0.9558
quartile	25.0%	0.9360
	10.0%	0.9187
	2.5%	0.8872
	0.5%	0.8823
minimum	0.0%	0.8823

Moments

Mean	0.96139
Std Dev	0.03486
Std Err Mean	0.00493
upper 95% Mean	0.97130
lower 95% Mean	0.95149
N	50.00000
Sum Wgts	50.00000

Figure 4.21 Distribution of Thickness and PAC (of Historical Runs)



Quantiles

maximum	100.0%	0.56833
	99.5%	0.56833
	97.5%	0.55174
	90.0%	0.47777
quartile	75.0%	0.45233
median	50.0%	0.43383
quartile	25.0%	0.41608
	10.0%	0.38157
	2.5%	0.36102
	0.5%	0.35533
minimum	0.0%	0.35533

Moments

Mean	0.43483
Std Dev	0.03664
Std Err Mean	0.00518
upper 95% Mean	0.44525
lower 95% Mean	0.42442
N	50.00000
Sum Wgts	50.00000

Quantiles

maximum	100.0%	1.8467
	99.5%	1.8467
	97.5%	1.8384
	90.0%	1.7897
quartile	75.0%	1.7617
median	50.0%	1.7100
quartile	25.0%	1.5742
	10.0%	1.4933
	2.5%	1.4440
	0.5%	1.4367
minimum	0.0%	1.4367

Moments

Mean	1.67180
Std Dev	0.11198
Std Err Mean	0.01584
upper 95% Mean	1.70362
lower 95% Mean	1.63998
N	50.00000
Sum Wgts	50.00000

Figure 4.22 Distribution of PACxp and CD (historical runs)

Chapter 5 Experimental Results and Discussion: Part II

In the preceding chapter, results of several experiments used to build the infrastructure of a supervisory control system have been presented and discussed. In this chapter we present the experimental results of a comparison of different control methods, including the use of dynamic specifications.

5.1 Experiments to Compare Different Control Methods

In order to demonstrate the benefits of applying dynamic specifications to supervisory control, three groups of wafers were run through the photolithographic sequence. These three groups were:

(1) Baseline historical runs, without additional process control (described in the previous chapter).

(2) Runs with local feedback control to each piece of equipment. Fixed specifications were used for all outputs, intermediate as well as final.

(3) Runs with local feedback control to each machine. In addition, the intermediate specifications of the system were dynamically adjusted. Only the final specification was fixed.

5.1.1. Experimental Setup

(1) Wafers: p-type, 4-inch silicon wafers with 1000 Å of oxide on the top.

(2) Equipment: SVG8626/36 wafer track, GCA 6200 wafer stepper, SVG 8632 develop track, SC PR Inspector INS-800-1, and Nanoline IV critical dimension measurement system.

(3) Photoresist: OCG820 G-line positive photoresist

(4) Control system: The BCAM (Berkeley Computer Aided Manufacturing) system was used as the process alarm system (providing multivariate model-based, statistical process control alarms and malfunction alarms) and the local feedback controller (providing model update and recipe update).

(5) Duration, frequency of runs, and wafer size: The three groups with different control methods were run independently within the same period of about one month. Thirty four-inch wafers (in ten lots) were run for each of the three groups. For each group, the separation between

two consecutive lots was about three days. Control was done on a run-to-run (lot-to-lot) basis; thus no change of recipe, models, or targets took place within one lot. Model updates were based on data from the last six wafers. No feed-forward control was used in any of the experiments.

(6) Measurements: Four readings of all output parameters were taken on each wafer. These readings were taken at four different sites near the center of a wafer (as shown in Figure 5.1). The average of these four readings was used as one data point.

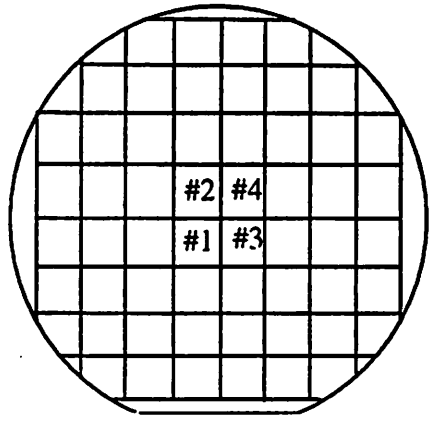


Figure 5.1 Location of Readings During the Control Comparison Experiments

5.1.2. Thickness Trend and Distribution of Control Comparison Experiments

In this section, the thickness trendcharts and distribution of the three groups are presented in Figures 5.2, 5.3, and 5.4. The specifications (targets, upper spec limits, lower spec limits), model predicted values, and actual measured values are included in these charts.

Tables 5.1 and 5.2 list the control actions on thickness of the two controlled groups. From these two tables and Figures 5.2, 5.3, and 5.4, several observations are made:

As seen in Figure 5.3 of the second group (local control only), wafers #7 ~ #9 and wafers #16 ~ #18 both drifted away from the model prediction and target. Therefore, two control alarms were triggered at wafers #9 and #18. Consequently, the thickness models were updated and model prediction was changed twice. In Figure 5.3, the second such change brought the model prediction away from the target. Usually this is undesirable. The reason was that the first model update also resulted in a recipe update, but the second one did not. The second model update was based on measured data of wafers #13 ~ #18; thus, it predicted a lower thickness value if the

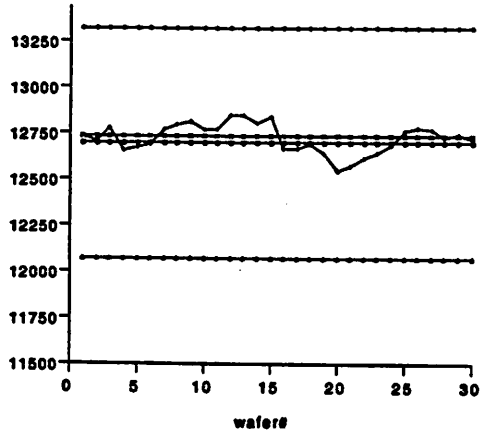


Figure 5.2 Trendchart of Thickness of Baseline Runs (Fixed Recipe)

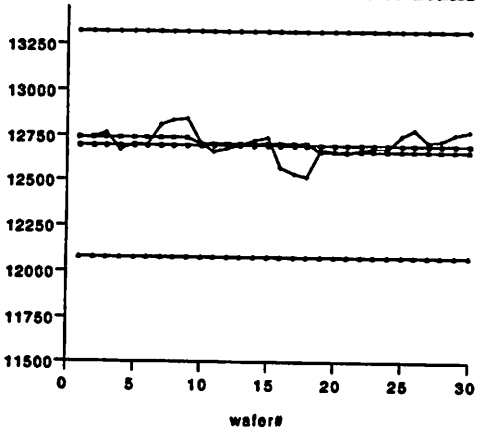


Figure 5.3 Trendchart of Thickness of the Second Group (Local Feedback Control)

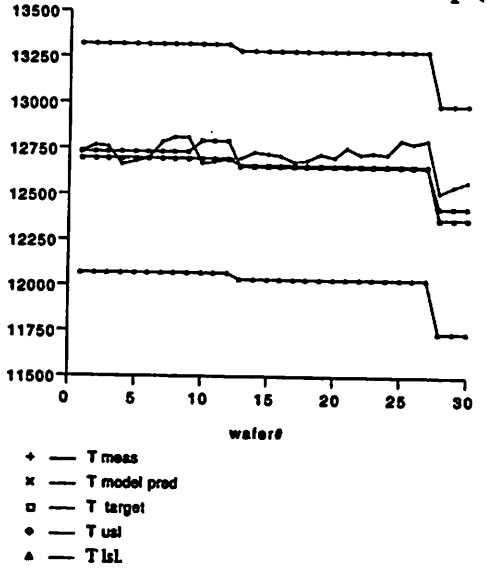


Figure 5.4 Trendchart of Thickness of the Third Group (Dynamic Specifications)

recipe was not updated accordingly. Usually, after each model update, the recipe was updated accordingly to produce an output closer to the target. However, in this case, the recipe was not updated since the thickness prediction was "close enough" (as judged by the BCAM controller) to the target. Tighter thickness specifications would have led to greater sensitivity.

Table 5.1 Summary of Alarms and Control Actions of Thickness of the Local-Control Group

wafer ID*	Target Change	Control Alarm**	Malfunction Alarm	Model Update	Recipe Update
#3	NONE				
#6					
#9		√		√	√
#12					
#15					
#18		√		√	√
#21					
#24					
#27					
#30		√	√		

Table 5.2 Summary of Alarms and Control Actions of Thickness of the Dynamic-Spec Group

wafer ID*	Target Change	Control Alarm**	Malfunction Alarm	Model Update	Recipe Update
#3					
#6					
#9		√		√	
#12	√				√
#15					
#18					
#21					
#24					
#27	√	√		√	√
#30					

* The wafers listed here were all the last (third) wafer in a lot. Control alarms were valid only if they occurred at these wafers since only lot-to-lot control was employed.

** Thickness and PAC jointly determined the control and malfunction alarms of the SVG coater.

As seen in Figure 5.4 of the dynamic-specification group, wafers #7~ #9 drifted away from the model prediction and a control alarm was triggered (as shown in Figure 5.6). The model was updated but the recipe was not (for the same reason mentioned above). The specification was changed at wafer #13 and #28 due to CD drift and subsequent spec propagation. In general, a specification change did not update the model (unless a control alarm and a model update were also triggered at the same time), but it updated the recipe to fit the new target. In the case of wafer#13, the thickness specification was changed, but the model was not. For wafer #28, both the specification and the model were changed. The recipe was changed to bring model prediction closer to the new target, and the actual measured values followed.

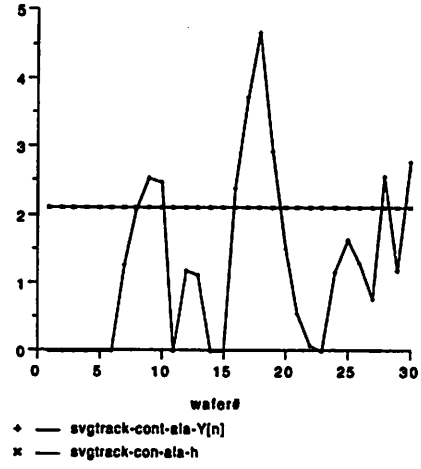


Figure 5.5 SVG Track Control Alarm of the Locally Controlled Group

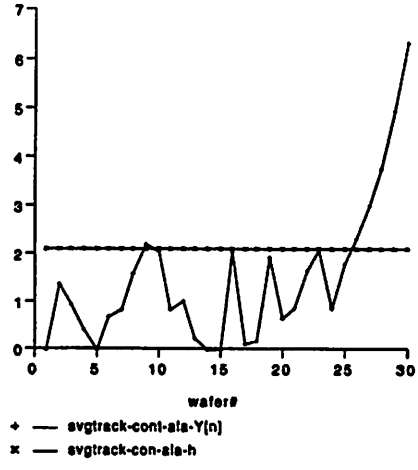


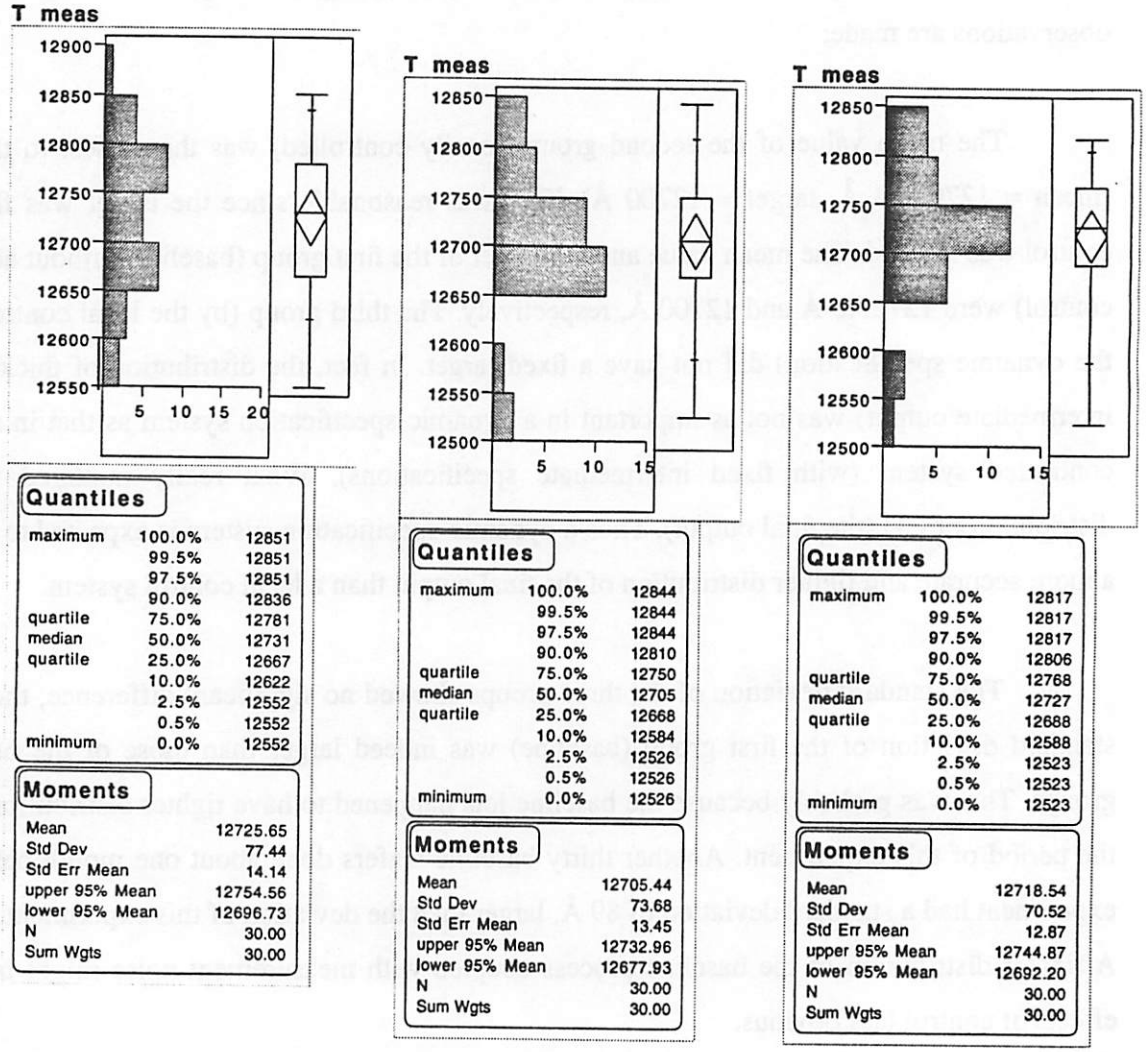
Figure 5.6 SVG Track Control Alarm of the Dynamic-Spec Group

From the thickness distribution of the three groups in Figure 5.7, the following observations are made:

The mean value of the second group (locally controlled) was the closest to the target (mean = 12705.44 Å, target = 12700 Å). This was reasonable since the target was fixed and control was applied. The mean value and the target of the first group (baseline without additional control) were 12725.6 Å and 12700 Å, respectively. The third group (by the local controller and the dynamic specification) did not have a fixed target. In fact, the distribution of thickness (an intermediate output) was not as important in a dynamic-specification system as that in a locally controlled system (with fixed intermediate specifications). What really mattered was the distribution of CD (the final output). Thus a dynamic-specification system is expected to produce a more accurate and tighter distribution of the final output than a local control system.

The standard deviation of the three groups showed no significant difference, though the standard deviation of the first group (baseline) was indeed larger than those of the other two groups. This was probably because the baseline lots happened to have tighter distribution during the period of this experiment. Another thirty baseline wafers done about one month before this experiment had a standard deviation of 89 Å, larger than the deviation of this experiment, 77.4 Å. A tighter distribution of the baseline process coupled with measurement noise might make the effects of control less obvious.

In retrospect, both controlled groups were controlled on a run-to-run, not wafer-to-wafer basis. The model used for the current lot were actually checked and updated (as directed by the control alarm) based on the measured data from the preceding lots. An tighter distribution can be expected with wafer-to-wafer control and/or with feed-forward control activated.



1. Baseline 2. Local Feedback Control 3. Dynamic Specification

Figure 5.7 Comparison of Thickness Distribution of the Three Groups

5.1.3 Trends and Distribution of PAC and PACxp of Control Comparison Experiments

PAC and thickness are both outputs of the SVG coater. Therefore, the malfunction and control alarms were calculated for the SVG coater in a multivariate way. The variance-covariance matrix between thickness and PAC was taken into account. Therefore, a control alarm on the SVG coater could be triggered by process drift of either thickness or PAC, or both. After a control alarm was triggered, the feedback controller was enabled. Model update did not necessarily happen even if a control alarm was triggered. The model update worked independently for the thickness model and the PAC model even though the control alarm test was done jointly.

The specifications, model prediction and measured values of PAC from the three groups are shown in Figures 5.8, 5.9, and 5.10. As in the case of thickness, the target of PAC of the third group (dynamic specification) was changed twice due to specification propagation. The recipe was also modified twice to bring the model prediction to the new target (as shown in Figure 5.10). Although the control alarm was triggered twice on the SVG coater for the third group (Figure 5.6), no model update was done by the feedback controller, as explained in the preceding paragraph. This was reasonable since PAC exhibited more oscillation around the model prediction rather than drifting to one direction.

For the locally controlled group, the target was always fixed. The PAC model of the locally controlled group was also not updated even though three control alarms were triggered on the SVG coater (Figure 5.5). However, the thickness model of the local-control group was indeed updated and the recipe was updated accordingly. This explained the change of PAC model prediction in Figure 5.9. Although the PAC models of both the local-control and the dynamic-spec groups were not updated throughout the experiment, the PAC distributions of both groups were better than that of baseline lots (Figure 5.11). This might be attributed to better thickness control and to recipe updates of the SVG coater due to the thickness model update.

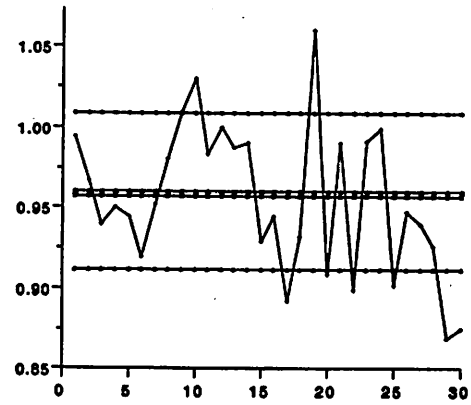


Figure 5.8 Trendchart of PAC of Baseline Runs (without additional control)

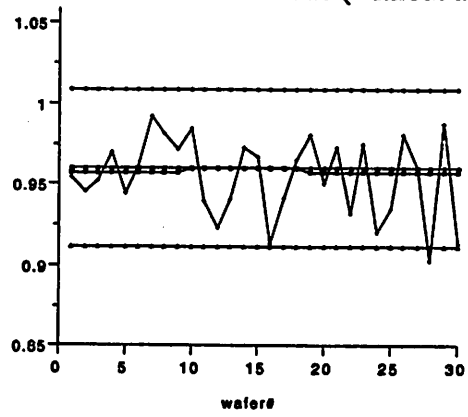


Figure 5.9 Trendchart of PAC of the Second Group (Local Feedback Control)

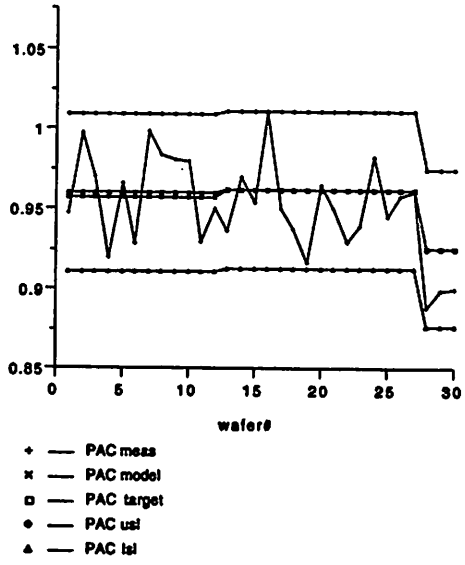


Figure 5.10 Trendchart of PAC of the Third Group (Dynamic Specifications)

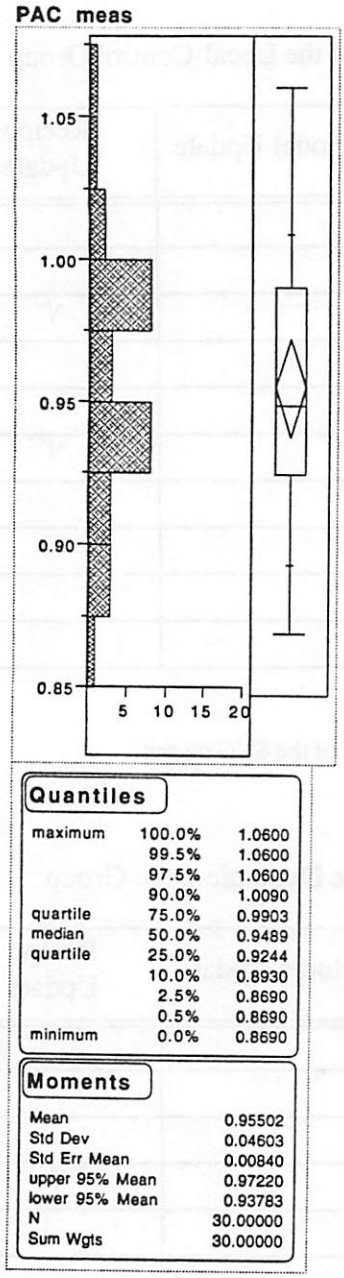
Table 5.3 Summary of Alarms and Control Actions of PAC of the Local-Control Group

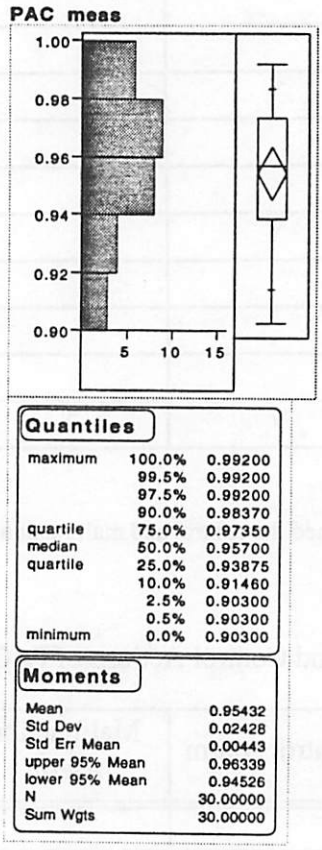
wafer ID	Target Change	Control Alarm**	Malfunction Alarm	Model Update	Recipe Update
#3	NONE				
#6					
#9		√			√
#12					
#15					
#18		√			√
#21					
#24					
#27					
#30		√	√		

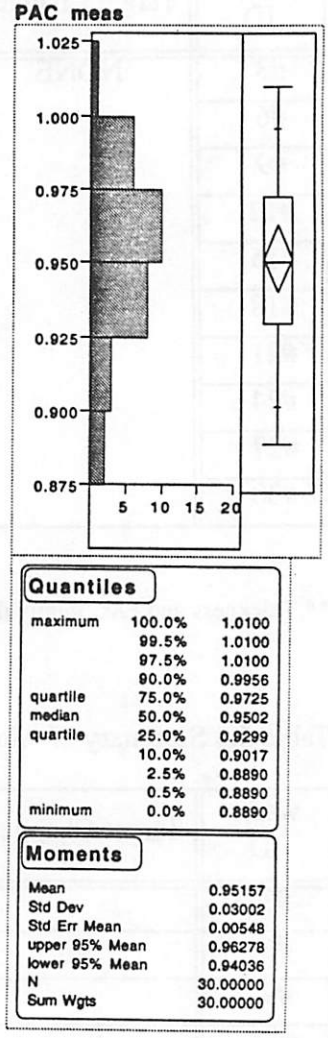
** Thickness and PAC jointly determined the control and malfunction alarms of the SVG coater.

Table 5.4 Summary of Alarms and Control Actions of PAC of the Dynamic-Spec Group

wafer ID	Target Change	Control Alarm	Malfunction Alarm	Model Update	Recipe Update
#3					
#6					
#9		√			
#12	√				√
#15					
#18					
#21					
#24					
#27	√	√			√
#30					







1. Baseline 2. Local Feedback Control 3. Dynamic Specification

Figure 5.11 Comparison of PAC Distribution of the Three Groups

Shown in Figures 5.12, 5.13, and 5.14 are the PACxp specifications, model prediction and measured values of the three groups.

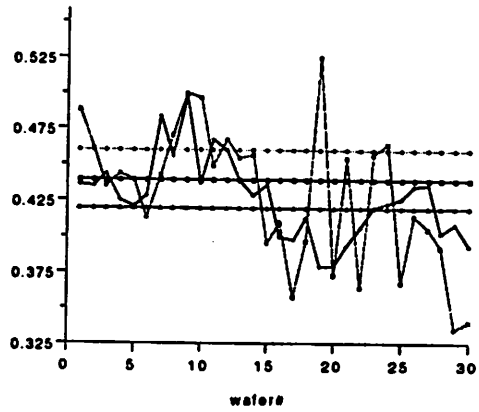


Figure 5.12 Trendchart of PACxp of Baseline Runs

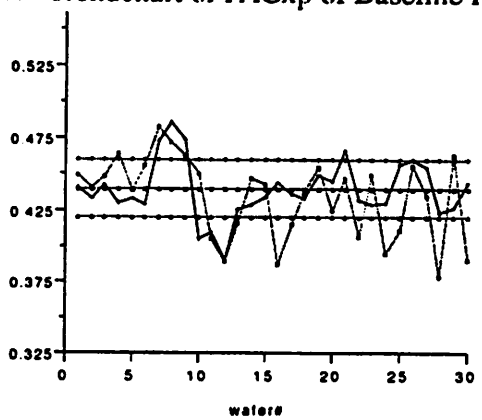


Figure 5.13 Trendchart of PACxp of the Second Group (Local Feedback Control)

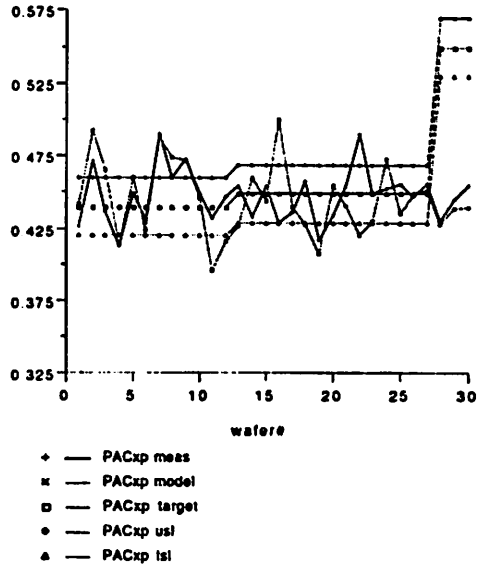


Figure 5.14 Trendchart of PACxp of the Third Group (Dynamic Specifications)

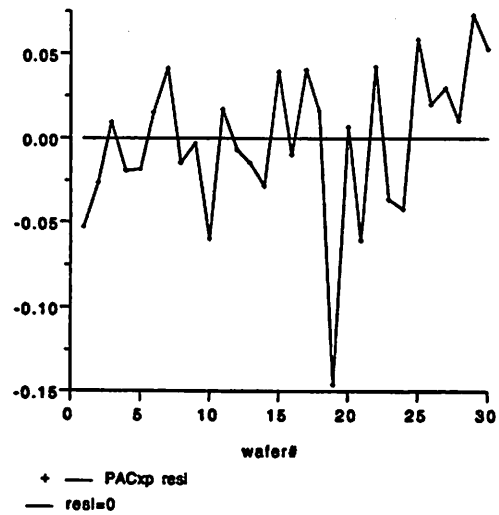


Figure 5.15 Residual Values (Measurement - Model Prediction) of PACxp of Baseline Runs

Shown in Figure 5.15 is the residual values of PACxp (= Measured value -Model prediction) of the baseline group. Except for wafers #19 and #29, the residual values were distributed between -0.05 to +0.05 or so. This distribution indicates that the model of PACxp was not a very precise model. Measurement noise and process variation both contributed to it. However, if we observe the distribution of real measured PACxp values in Figure 5.12, we notice that the values were roughly distributed between 0.35 to 0.5 (a wider range than that of the residuals) and showed a drifting-down trend. These residual values did not show any obvious drift. This means that the model could still predict the real values, and that it would have been useful in the application of feed-forward control.

Tables 5.5 and 5.6 show the summary of alarms and control actions of PACxp (on the GCA stepper). It is noteworthy that the model prediction and measured values of wafer# 28 ~ #30 of the dynamic-spec group did not follow the target. This was because the new target (0.55, a result of specification propagation caused by an extraordinary CD drift) was too high to be reached by any allowed exposure dose with two other "uncontrollable" inputs (thickness and PAC) from the upstream machine. In fact, the exposure dose was reduced to the minimal value (80 mj/cm^2) in order to produce a high PACxp value. However, given also the unusually low PAC

value, neither the model prediction of PACxp nor the measured value could match the target.

Table 5.5 Summary of Alarms and Control Actions of PACxp of the Local-Control Group

wafer ID	Target Change	Control Alarm:high	Control Alarm:low	Malfunction Alarm	Model Update	Recipe Update*
#3	NONE					
#6						√
#9						√
#12						√
#15						
#18						
#21						
#24						
#27						
#30		√		√		

*All recipe updates were due to baseline adjustments, not control actions.

Table 5.6 Summary of Alarms and Control Actions of PACxp of the Dynamic-Spec Group

wafer ID	Target Change	Control Alarm:high	Control Alarm:low	Malfunction Alarm	Model Update	Recipe Update
#3						
#6						√*
#9						√*
#12	√					√
#15						
#18						
#21						
#24						
#27	√					√
#30		√		√		

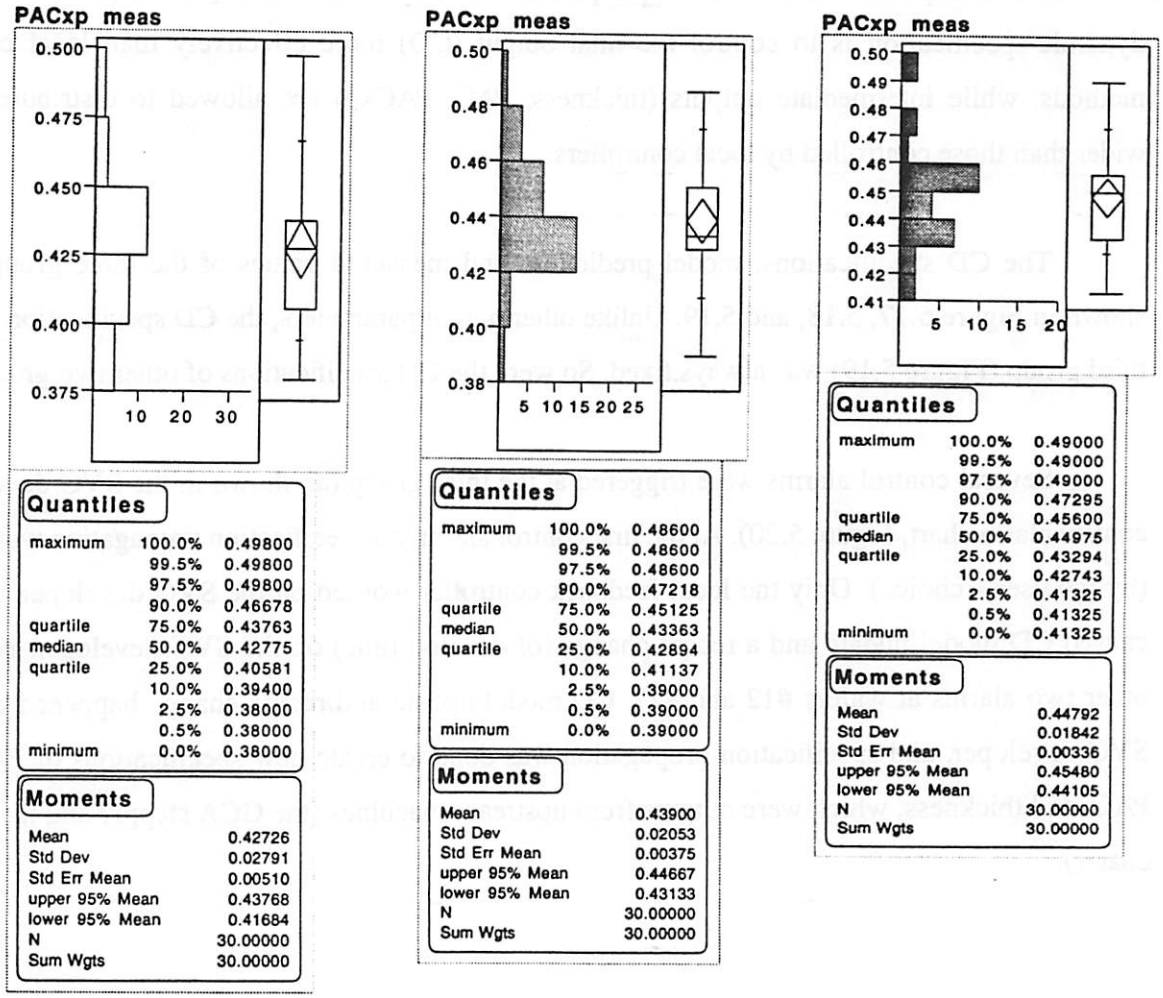
*These recipe updates were due to baseline adjustments, not control actions.

If we compare the trendcharts of PAC and PACxp of the dynamic-specification group (Figure 5.10 and Figure 5.14), we can see that the PAC of wafers #28 ~ #30 dropped significantly while the PACxp of the same group did not. Therefore, ΔPAC (PAC-PACxp) decreased as predicted by the exposure dose decrease. This was also consistent with the fact that the CD of wafer #28 ~ #30 was brought up toward the target. In this example, CD is a strong function of exposure dose, which correlates significantly with ΔPAC rather than PACxp, as demonstrated in Chapter 4.3 (equipment model generation). However, PACxp is still useful in a dynamic-specification system in that it can serve as a target by which the exposure dose is computed by the local controller on the GCA stepper. In other words, PACxp effectively links CD to exposure dose.

In collecting the measured data of PAC and PACxp, significant measurement noise was evident. In addition, some computed PACxp values were discarded because they were too far off the normal values (for PACxp, about 0.3 to 0.55). Of the four measurements made on each wafer, only those with normal values were taken and averaged. Overall, nearly fifteen to twenty percent of PACxp average values were computed from less than four valid measurements.

This was not the case for thickness. Almost all measured thickness values were valid and tightly distributed on any given wafer. The noisy measurement of PAC and PACxp might result from the relatively complex algorithm of computation, combined with limits set by our in-situ automatic measuring instrument and incomplete information on properties of photoresist. However, from the experimental results, PAC and PACxp were still useful parameters for control purposes, even if they were not as predictable as thickness (due to measurement noise).

The distribution of PACxp of the three groups were shown in Figure 5.16. Again, both controlled groups showed tighter distribution than the baseline group. The second group (local feedback control with a fixed PACxp target) had a mean value of 0.439, which was very close to the target (0.44). The third group (dynamic specification) did not have a fixed PACxp target.



1. Baseline 2. Local Feedback Control 3. Dynamic Specification

Figure 5.16 Comparison of PACxp Distribution of the Three Groups

5.1.4 CD Trends and Distribution of Control Comparison Experiments

CD (Critical Dimension) of photoresist is the final output of a photolithographic workcell. It is the most important of the four output parameters. The essence of supervisory control using dynamic-specification is to control the final output (CD) more effectively than local control methods, while intermediate outputs (thickness, PAC, PACxp) are allowed to distribute even wider than those controlled by local controllers.

The CD specifications, model prediction, and measured values of the three groups are shown in Figure 5.17, 5.18, and 5.19. Unlike other output parameters, the CD specification of the third group (Figure 5.19) was always fixed. So were the CD specifications of other two groups.

Several control alarms were triggered at the third group (as shown in the SVG developer control alarm chart, Figure 5.20). At the first control alarm, no specification propagation was done (by the user's choice). Only the local feedback controller worked on the SVG developer, which caused CD model update and a recipe change (of develop time) on the SVG developer. For the other two alarms at wafers #12 and #27, CD model update and recipe change happened on the SVG developer, and specification propagation was done to create new specifications of PACxp, PAC, and thickness, which were outputs from upstream machines (the GCA stepper and the SVG coater).

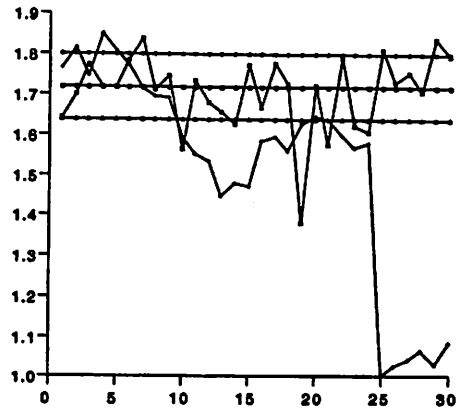


Figure 5.17 CD Trendchart of Baseline Runs (without additional control)

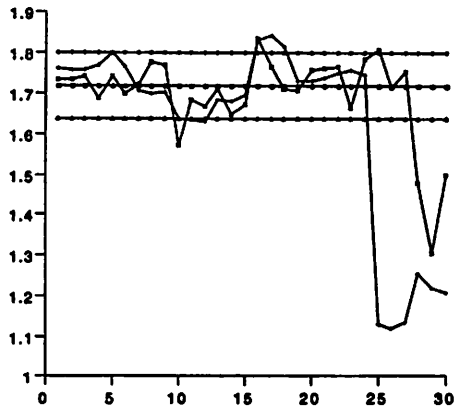


Figure 5.18 CD Trendchart of the Second Group (Local Feedback Control)

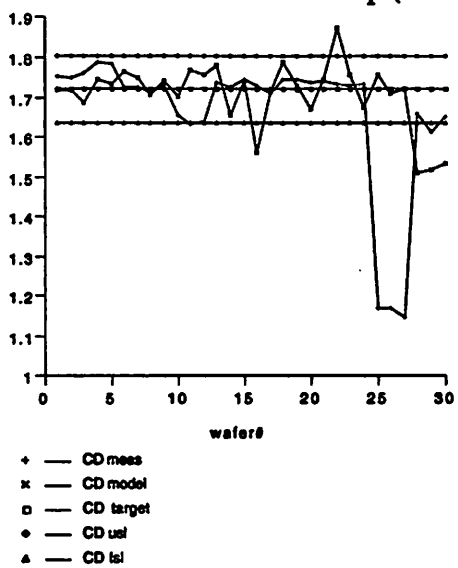


Figure 5.19 CD Trendchart of the Third Group (Dynamic Specifications)

Table 5.7 Summary of Alarms and Control Actions of CD of the Local-Control Group

wafer ID	Target Change	Control Alarm high	Control Alarm low	Malfunction Alarm**	Model Update	Recipe Update*
#3	NONE					
#6		√		√	√	√
#9			√	√	√	√
#12						
#15						
#18		√		√	√	√
#21						
#24						
#27			√	√	√	√
#30			√	√		

** The rate of malfunction alarm is high due to the larger-than-usual CD drift and shift during this experiment.

Table 5.8 Summary of Alarms and Control Actions of CD of the Dynamic-Spec Group

wafer ID	Target Change	Control Alarm:high	Control Alarm:low	Malfunction Alarm	Model Update	Recipe Update*
#3	NONE	√		√	√	√
#6						
#9						
#12			√	√	√	√
#15			√			
#18		√				
#21						
#24						
#27			√	√	√	√
#30		√	√	√		

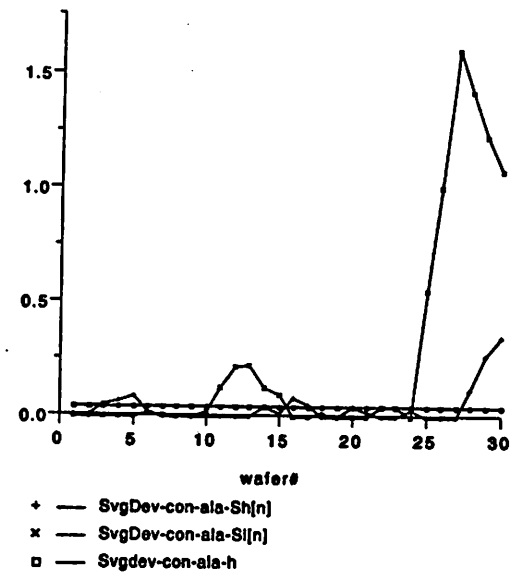


Figure 5.20 Control Alarm Chart of SVG Developer (Dynamic Specification Group)

Summary and Discussion

The superior performance of dynamic-specification control over local control was obvious when the CD trends and control actions of both methods were compared. At wafers #10 ~ #12, both groups had a CD drop that resulted to control alarms. In the local-control group, the CD model was updated and a new recipe was calculated. The only controllable input at the SVG developer was the develop time. It was updated in order to raise CD back to the target. However, even with the lowest allowable develop time (50.2 seconds) combined with the fixed-target thickness, PAC, and PACxp, the CD model prediction still could not be brought back to the target. As a result, the next lot (wafers #13 ~ #15) was improved but still below-target with a CD of about 1.66 μm, while the target was 1.72μm).

On the other hand, in the dynamic-specification group, not only the CD model and develop time (recipe of SVG developer) were updated, but also the specifications of PACxp, PAC, and thickness were changed. Consequently, new recipes of the GCA stepper and the SVG coater were also calculated. This way the CD model prediction could be easily brought back to the target. As a result, the measured CD values of about 1.74 μm of the next lot (wafers #13 ~ #15)

were very close to the 1.72 μm target.

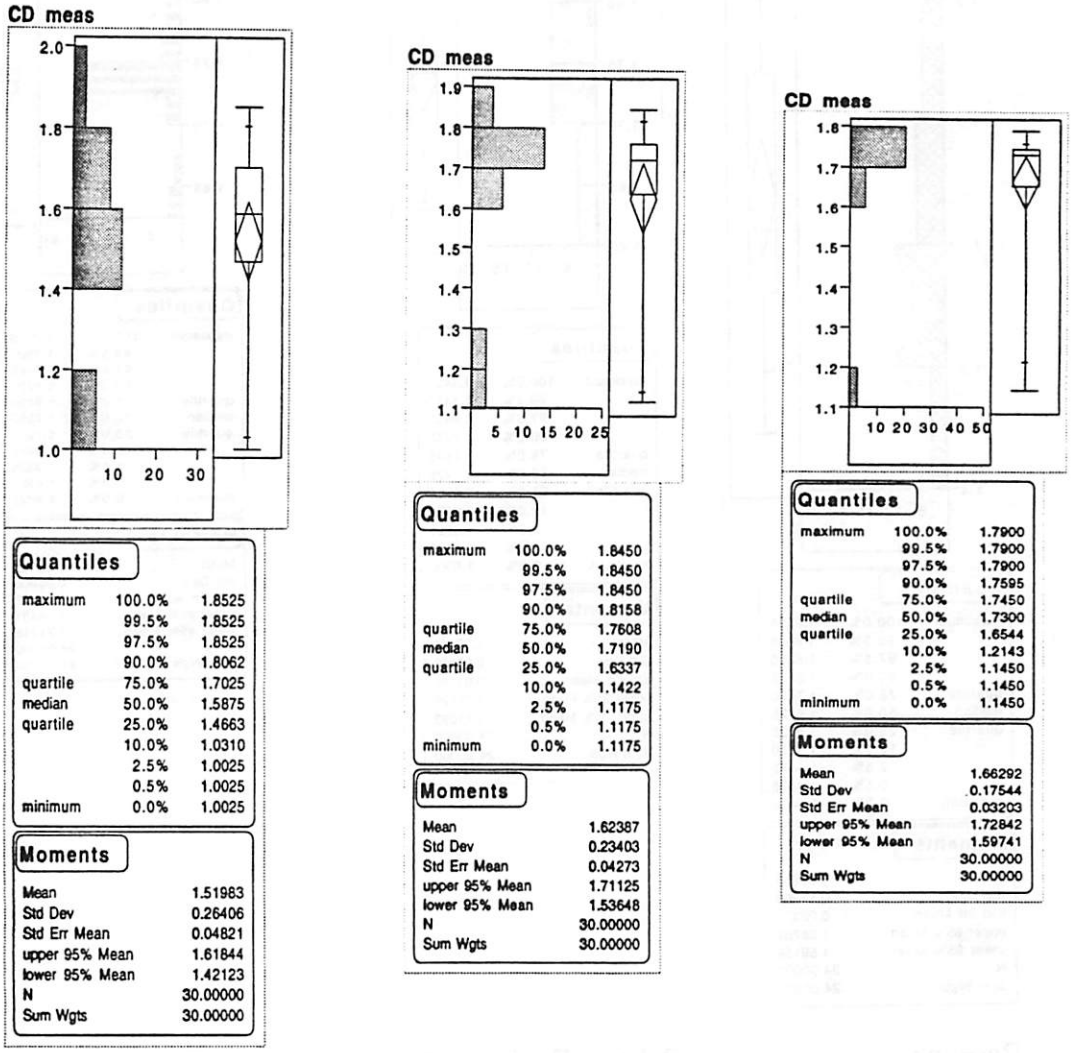
The contrast was even more dramatic at wafers #27 to #28. A serious drop of CD was seen in all three groups at wafers #25 ~ #27. Malfunction and control alarms were triggered on both controlled groups. In the local control group, only limited corrections could be made on CD model prediction and measured values of wafers #28 ~ #30 (only brought the CD from 1.12 μm up to 1.25 μm). However, in the dynamic-specification group, the CD model prediction and measured values were brought much closer to the target (CD values brought back from 1.15μm to 1.63μm). The reason was the same as stated in the preceding paragraphs.

It was known that exposure dose had a stronger effect on CD than develop time did. However, in the local control group, only develop time was directly updated in response to a CD control alarm and model update. Not being a controllable input of the SVG developer, exposure dose could not be modified due to any CD control alarm or model update. The exposure dose could only be updated on the GCA stepper after a control alarm and model update of PACxp were triggered.

Considering a situation where exposure dose went wrong and both PACxp and CD drifted away in the same direction and created control alarms, the local-control system should work well to bring CD back to the target. However, in our experiments, PACxp did not show obvious drift when CD drifted significantly. Therefore, CD control in the local-control group was obviously less effective than that in the dynamic-specification group. The question remains, why the CD drift was not accompanied by a PACxp drift? It was suspected that other less obvious factors, which had little effect on PACxp, appeared and affected CD significantly for these wafers. Since PACxp (or more relevant to exposure dose, ΔPAC) did not drift away from the target, no control actions on exposure dose could be taken by the locally controlled sequence.

On the other hand, the dynamic-specification control worked very well to bring CD back to its target since all preceding parameters (exposure dose, PACxp, thickness, PAC) as well as develop time were "linked" through the dynamic specifications and allowed to be modified. As explained in Chapter 5.1.3, PACxp served as a good "media" to translate a CD model update to

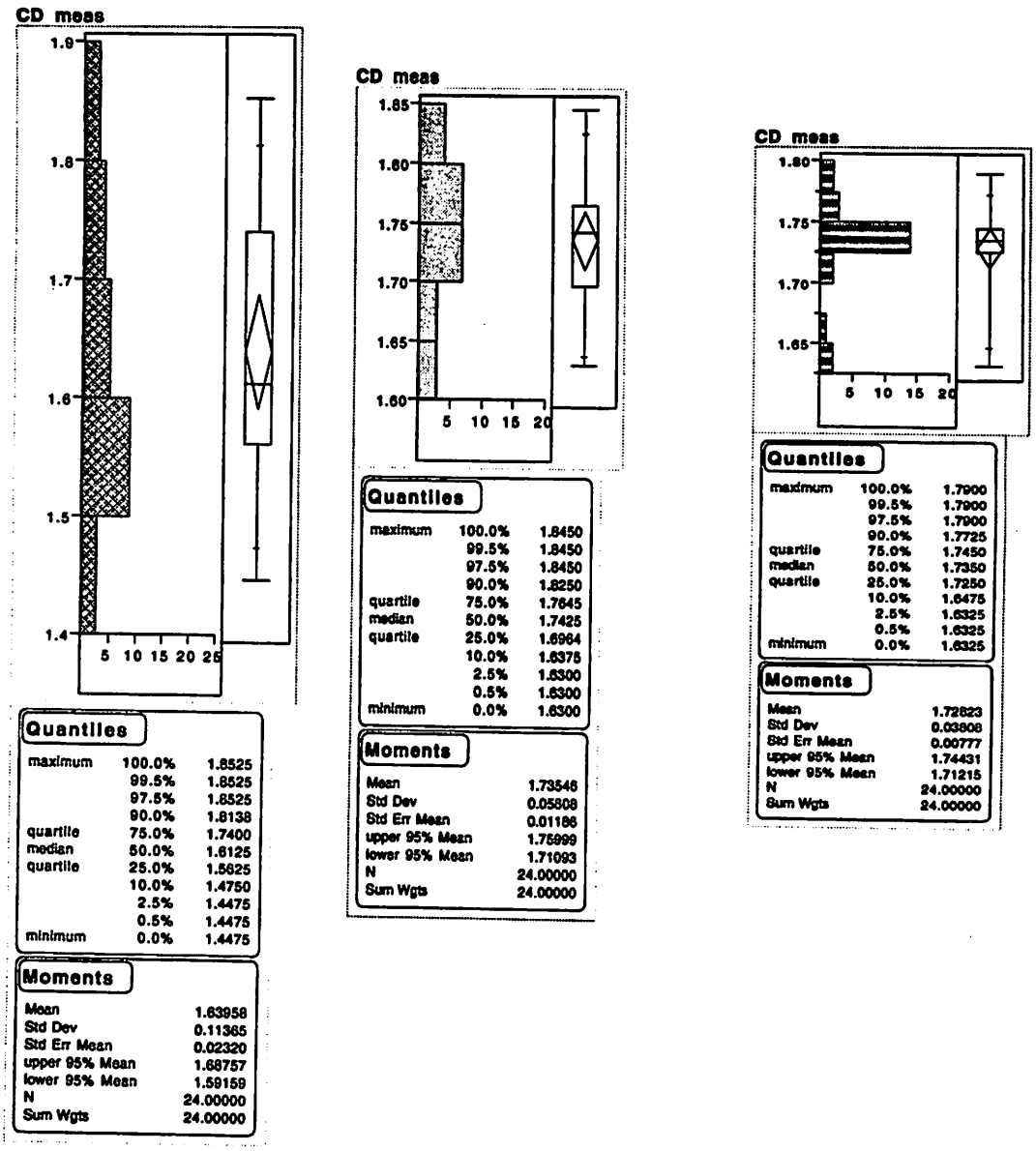
exposure dose update when PACxp itself might not be on target. Also the dynamic-specification control exhibited the capability to correct things, even when some factors not included in the equipment models might have affected the process. Shown in Figure 5.21 is the CD distribution of the three groups. It is obvious that the dynamic-specification group had the tightest distribution and a mean value closest to the 1.72 μm target. The performance of the locally controlled group was worse than that of the dynamic-specification group but better than that of the baseline group.



1. Baseline 2. Local Feedback Control 3. Dynamic Specification

Figure 5.21 Comparison of CD Distribution of the Three Experimental Groups

If we treat data from the last two lots (wafers #25 ~ #30) as outliers (because they were too far away from normal values and triggered serious malfunction alarms) and exclude them in calculating the distribution, the CD distribution is shown in Figure 5.22. In Figure 5.22 the CD



1. Baseline 2. Local Feedback Control 3. Dynamic Specification

Figure 5.22 Comparison of CD Distribution of the Three Group (Excluding the last two lots)

distribution improvement of the local-controlled group over the baseline group was more obvious

than that in Figure 5.21. In Figure 5.22 the dynamic-specification group still had the tightest and most accurate distribution. This means that under usual process drift and variation the local controller worked reasonably well (though not as well as the dynamic-specification group), but under serious process variation, like those of wafers #25 ~#30, the control capability of the local-control group was quite limited while the dynamic-specification control still worked well. The results of Figure 5.22 can be quantified as follows. Assuming that the process capability (Cpk) of the historical baseline was 1.00, then the relative Cpk of the locally controlled process is 1.90 and the relative Cpk of the process using dynamic specifications is 2.90.

To sum up, we have demonstrated that the dynamic specifications can be used to produce a tighter and more on-target distribution, which means higher Cpk, of the final output (e.g. CD) of a workcell consisting of multiple process steps. If feed-forward control is applied, we can expect better distributions of PACxp and CD. If wafer-to-wafer control and a tighter thickness spec are used, even better distributions of all outputs can be expected.

References

[1] *SAMPLE User Guide*, The SAMPLE group, ERL/Deparment of EECS, University of California at Berkeley, June, 1991

[2] John Thomson, MS Thesis, UCB/ERL M92/118, University of California at Berkeley, October, 1992

[3] J. Edward Jackson, *A User's Guide to Principal Component*, John Wiley & Sons, 1991

[4] Frederick H. Dill et. al, "Characterization of Positive Photoresist", *IEEE Transactions on Electron Devices,* July, 1975

[5] Frederick H. Dill, "Optical Lithography", *IEEE Transactions on Electron Devices*, July, 1975

[6] Sovarong Leang, Ph.D. Dissertation, University of California at Berkeley

[7] *Inspector INS-800-1 Instruction Manual*, SC Technology Inc., 1991

[8] Bart Bombay, MS Thesis, UCB/ERL M92/113, University of California at Berkeley, September 1992

[9] Ronald Crosier, "Multivariate Generalization of Cumulative Sum Quality Control Schemes" *Technometrics*, Vol 30 No.3, August, 1988