

Copyright © 1991, 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.

**SPECIAL ISSUES IN SEMICONDUCTOR
MANUFACTURING**

II

EECS 290W Class Projects Reports, Fall 1990

Professor:

Costas J. Spanos

Students:

Raymond Chen, Sovarong Leang, Nelson Tam

Memorandum No. UCB/ERL M91/8

30 January 1991

COVER PAGE

**SPECIAL ISSUES IN SEMICONDUCTOR
MANUFACTURING**

II

EECS 290W Class Projects Reports, Fall 1990

Professor:

Costas J. Spanos

Students:

Raymond Chen, Sovarong Leang, Nelson Tam

Memorandum No. UCB/ERL M91/8

30 January 1991

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

Preface

This is the second annual edition of the 290W report. This edition includes descriptions of projects completed during the Fall semester of 1990. These projects were completed in the context of the graduate course "Special Issues in Semiconductor Manufacturing". Three students have participated, and according to the course requirements, these students worked with me on their projects during the last six weeks of the semester. Although all of the three projects have something to do with lithography, they are quite diverse, as they treat three very different subjects.

The first project, deals with the application of economic criteria towards the implementation of viable control strategies in semiconductor manufacturing. This study was motivated by the recent introduction of expensive wafer mapping machines that can be used to monitor the quality of lithographic patterning. The cost of such a quality control scheme dictates an optimum design, where metrology and analytical resources are allocated in a fashion that balances the cost of inspection with the expected loss due to misprocessing. The approach taken in this project is based on classical work on the Economical Design of control charts, extended to cases where multiple process steps need to be controlled using one finite pool of metrology and diagnostic resources.

The second project also deals with photolithography. The objective here is to define a "run-by-run" supervisory control scheme, that, by means of feedback and feed-forward adjustments, will reduce the variability of the critical dimensions of the developed photoresist patterns. In this report we show early experimental evidence that such a scheme is possible. This is shown by means of an actual implementation of a feedback loop around the spin/coat & bake equipment used in the Berkeley microfabrication laboratory. This feedback loop operates using model-based, cumulative sum control charts of the resist thickness and reflectance, and also on optically measured critical dimensions.

The third project focuses on the application of formal experimental design techniques towards the derivation of novel models that describe the dissolution of chemically amplified resist. Two such models have been developed, and they have been used to predict two critical parameters of the final dissolution rate model. These models show excellent agreement with the experimental results, and they are going to be used within existing process simulators.

It is my hope that these reports will be a useful addition to our understanding of semiconductor manufacturing. My thanks go to Raymond, Sov and Nelson, the 290W students whose work made this document possible.

Costas J. Spanos

January, 1991

The first part of the report deals with the general situation of the country and the progress of the work. It is followed by a detailed account of the work done during the year, and a summary of the results. The report is divided into several sections, each dealing with a different aspect of the work.

The first section deals with the general situation of the country and the progress of the work. It is followed by a detailed account of the work done during the year, and a summary of the results. The report is divided into several sections, each dealing with a different aspect of the work.

The second section deals with the work done during the year. It is followed by a detailed account of the work done during the year, and a summary of the results. The report is divided into several sections, each dealing with a different aspect of the work.

The third section deals with the results of the work. It is followed by a detailed account of the work done during the year, and a summary of the results. The report is divided into several sections, each dealing with a different aspect of the work.

The fourth section deals with the conclusions of the work. It is followed by a detailed account of the work done during the year, and a summary of the results. The report is divided into several sections, each dealing with a different aspect of the work.

Table of Contents

1. Economic Design of Control Charts	<i>Page 7</i>
2. Feedback Control for a Lithography Workcell	<i>Page 29</i>
3. Modeling a Chemically Amplified Resist	<i>Page 45</i>

January 1991

The following information was obtained from the records of the Department of the Interior, Bureau of Land Management, regarding the acquisition of the land described herein:

Economic Design of Control Charts

Raymond L. Chen

Due to the limited resources available for the implementation of a statistical quality control program, as well as due to the loss of profit caused by non-conforming products, it is logical and feasible to design a control chart by economical considerations. This design is carried out by optimizing the net income per unit time of a process, taking into account the cost invested in the operation of the control system. Numerical examples indicate that our theoretical model and computer programs are applicable into modern wafer production.

1.0 Introduction

In this report we will present a detailed analysis on economic design of the charts for the statistical quality control of semiconductor wafer production. First, we will state our motivation and present a simple background review, followed by a detailed derivation of the formulas. Computations will be carried out, followed by some numerical results and conclusions. The computer programs, written in FORTRAN-77, are included and explained in the Appendix.

1.1 Motivation

Modern wafer inspection systems can give incredible amount of detail. For example, we can now inspect a patterned wafer after each patterning step and record the number, location and size (down to about $0.2 \mu\text{m}$) of defects. The problem is that such systems are very costly to acquire (\$1-2 mil) and to maintain (\$0.5mil/year). In addition, complete mapping a single wafer takes about 30 minutes. In a mid-volume production facility that produces 500 wafers/week, and each of them goes through ten patterning steps, we cannot inspect every single wafer. In order to balance the cost of inspection and the cost of faulty production, we must do an "economic design" of the statistical control procedure.

In other words, we need to generate guidelines on how many wafers will be sampled and how often, depending on the probability that each patterning operation has to go out of control, and given the cost of producing faulty wafers.

1.2 Background

\bar{X} control charts are widely used to establish and maintain the statistical control of a production sequence. To design a control chart, one needs to know the sample size of each group of measurements (n), and the sampling frequency or, equivalently, the interval between samples (h). Assuming that the production process generates defects whose numbers are randomly distributed according to known mean and sigma, we also need to determine the control limits of the chart defined as $(\mu \pm k\sigma)$.

Although in the wafer inspection example the practical sample size n is one, due to the relative long time to investigate a single wafer, all the formulas are derived in a general way so that n could larger than one.

The cost to maintain a control chart of a process is mainly associated with the sampling, testing, investigating true out-of-control signals, possibly correcting the assignable causes, as well as investigating the occasional false alarm signals that occur when the process is actually in control. In order to design the control chart from an economic viewpoint, one must optimize the total net income of the process. One must take into account both the cost of maintaining the control charts and the loss due to the non-conforming wafers that are produced while the process is out of control.

Early in 1956, Duncan [1] proposed a simple production model to design the \bar{X} -chart. Later, Goel et al. [2] used an algorithm to find the exact optimum solution of Duncan's problem by computer, which is academically interesting, but might be practically difficult. Chiu et al. [3] developed a simple, approximate procedure to optimizing Duncan's model. Finally, Montgomery [4] offered a simple FORTRAN program to solve the problem based on these previous works.

A more historical review of the research done in the area of economic design of control charts is presented by Montgomery [5], and we will not repeat it in this report. Rather, based on Montgomery's review, we will develop our own model for the specific situation in the wafer production control system.

In our model, however, we will add a total cost constraint on maximizing the income in Duncan's original model. This is dictated by the fact that the resources we can apply towards control are finite. Furthermore, since the total control cost is not a fixed but a random number due to the randomness in the occurrence of the out-of-control status as well as the randomness of the false alarms (type I error), we will treat the total control demand as a random variable obeying some distribution inferred from the convolution of Binomial and Poisson distributions.

We will not only derive a mathematical model, but we will also provide a practical computer program, which will make the optimization problem numerically feasible.

2.0 Methodology

The production, monitoring, and adjustment process may be thought of as a series of repetitive cycles over time. Each cycle begins with the production process in the in-control state. By definition, when the production process is in control, the measured variable is normally distributed with a mean μ and a standard deviation σ . When the process goes out of control, it is assumed that there is a single assignable cause, which takes the form of a shift in the process mean from μ to $\mu+\delta\sigma$ or $\mu-\delta\sigma$, where δ is known and σ is unchanged.

The process is monitored by an \bar{X} -chart with center line μ and upper and lower control limits $\mu\pm k(\sigma/\sqrt{n})$, where n is the sample size (which is equal to one in our case). Samples are to be taken at intervals of h hours. After an out-of-control signal is produced by the control chart monitoring system, an investigation of the alarm is carried out and possibly an adjustment is made in order to bring the process back into the in-control state. Then a new cycle begins.

Let T be the length of a cycle, and let $E(T)$ be the expected length or the long-term average length (mean length) of a cycle, and let $E(P)$ be the expected net income (profit) incurred during a cycle. Then the expected net income per unit time is:

$$E(A_P) = \frac{E(P)}{E(T)}$$

Our goal is to maximize $E(A_P)$. This can be accomplished by standard optimization techniques, given the maximum total control cost $E(Cc)$ in one cycle. The objective is to determine the optimized design parameters n , k , and h . Though $n = 1$ in our derivation, we will still keep n in the notation for the future extension of our model.

In addition, since we have more than one process, we will have to determine the values of k and h for all of the processes (or "multiple photolithographic steps" in wafer production). We will use the subscript i to distinguish among these steps. Wherever there is no i subscript, the equations and formulas are assumed to apply unaltered for all the steps.

2.1 Expected Time of a Process Control Cycle

Each cycle consists of four periods: (1) the in-control period G , (2) the out-of-control period B_β , (3) the time to take a sample and interpret the results, and (4) the time to find the assignable cause and bring the process back to the controlled state. The total time of (3) and (4) is called B' . Then total time the process is out-of-control is $B=B_\beta+B'$, and the total cycle time is $T=G+B$.

The assignable cause of the shift is assumed to occur according to a Poisson process with an intensity of λ occurrences per hour. That is, if the length of the in-control period

is G , then the expected value of G , $E(G) = 1/\lambda$. Therefore, given the occurrence of the assignable cause between the j th and $(j+1)$ st samples, the expected time of occurrence within this sampling interval is given by:

$$\tau = \frac{\int_{jh}^{(j+1)h} e^{-\lambda t} \lambda dt}{\int_{jh}^{(j+1)h} e^{-\lambda t} dt} = \frac{1 - (1 + \lambda h) e^{-\lambda h}}{\lambda (1 - e^{-\lambda h})} \equiv f(x) \cdot h \quad (2-2)$$

where $x = \lambda h$, and

$$f(x) = \frac{1 - (1 + x) e^{-x}}{x (1 - e^{-x})} \quad (2-3)$$

The number of samples M_β that is required in order to produce an out-of-control signal, given that the process is actually out of control, is a geometric random variable with mean $E(M_\beta) = 1/(1-\beta)$. From this we conclude that the expected length of the out-of-control period B_β is $E(B_\beta) = h/(1-\beta) - \tau$. The time required to take a sample and interpret the results is a constant g proportional to the sample size, so gn is the length of this segment of the control cycle. The time required to find the assignable cause following a true alarm signal is a constant g' . So $B' = gn + g'$ is a constant time. Therefore, the expected length of a cycle is:

$$E(T) = \frac{1}{\lambda} + \frac{h}{1-\beta} - \tau + gn + g' \quad (2-4)$$

where the "power" of detecting an assignable cause (true alarm) is:

$$\begin{aligned} 1 - \beta &= \Phi(-k + \delta\sqrt{n}) + \Phi(-k - \delta\sqrt{n}) \\ &\approx \Phi(-k + \delta\sqrt{n}) \end{aligned} \quad (2-5)$$

since usually $\Phi(-k - \delta\sqrt{n}) \approx 0$, where:

$$\Phi(z) = \int_{-\infty}^z \phi(z') dz' \quad (2-6)$$

and $\phi(z) = (2\pi)^{-1/2} \exp(-z^2/2)$ is the density of the standard normal distribution.

2.2 Control Cost of a Process Cycle

The net income per unit time of operation in the in-control state is v_0 , and the net income per unit time in the out-of-control state is v_1 . In other words, the loss per unit time due to the production of non-conforming units is $d=v_0-v_1$. The cost of taking samples in one interval (h) is $a+bn$, where n is the sample size. The cost of finding an assignable cause (true alarm) is c , and the cost of investigating a false alarm is c' .

Without loss of generality, we define the time of one sampling interval h as our "time unit". Thus, the total sample number in a cycle is M_T with $E(M_T) = E(T)/h$. The total number of samplings before the shift is M_G with $E(M_G) = [E(G) - \tau]/h = [1/\lambda - \tau]/h$, and the rest of M_T is $M_B = M_T - M_G$ with $E(M_B) = E(M_\beta) + E(B') = 1/(1-\beta) + (gn+g')/h$.

Let us use the average cycle time $E(T)$ to replace $T = M_T h$, and also use the average time for the monitoring system (control chart) needed to detect a shift, i.e., $h/(1-\beta)$, to replace the actual time $B_\beta = M_\beta h$ as our first approximation (more detailed analysis will be included in our future plan). Namely, we will concentrate on the distribution for the number of false alarms (type I error) to get our estimation of total control cost, which is principally a random variable whose distribution is decided by the distributions of random variables T , number of missed alarms ($M_\beta - 1$) (type II error), and the number of false alarms M_α (type I error). So we will replace M_β , M_G and M_T (or T) by their average values.

By definition, we can only have one true alarm per cycle. After that, an investigation and correction follows, so that the assignable cause of the shift is fixed and a new cycle starts. The probability of the process staying in control for a time period of G' is $P(G') = E(M_G)/E(M_T) = (1/\lambda - \tau)/E(T)$. Here we imply that $E(G') = E(G) - \tau$. The probability of a false alarm in a cycle T is thus given by:

$$P_\alpha = \alpha \times \left(\frac{1}{\lambda} - \tau\right) / E(T) \tag{2-7}$$

where $\alpha = 2\Phi(-k)$ as the type I error. The probability of giving M_α false alarms during the entire cycle T is thus given by the binomial distribution:

$$P(M_\alpha) = \frac{M_T!}{M_\alpha! (M_T - M_\alpha)!} P_\alpha^{M_\alpha} (1 - P_\alpha)^{M_T - M_\alpha} \tag{2-8}$$

with the average $E(M_\alpha) = M_T P_\alpha \approx E(M_T) P_\alpha = \alpha(1/\lambda - \tau)/h$.

So the “control cost” of a cycle with M_α false alarms and s single true alarm is given by:

$$C_c = c'M_\alpha + c + (a + bn)M_T \quad (2-9)$$

with the long-run average value:

$$E(C_c) = c' \frac{\alpha(1/\lambda - h)}{h} + c + (a + bn) \frac{E(T)}{h} \quad (2-10)$$

We will use this average value as our initial estimation of the control cost. This cost will also be used against the constraint that dictates that we have finite control resources. After getting (k, h) , it is straightforward to determine the C_c as a distribution of M_α . More detailed work will be included in future plan.

2.3 Net Income Optimization

After defining all the parameters and variables above, we know that the net income per cycle T is given by: $P = v_0 G + v_1 B - C_c = v_0 T - d B - C_c$. Therefore, the expected net income per cycle is given by:

$$E(P) = v_0 E(T) - d E(B) - E(C_c) \quad (2-11)$$

where,

$$E(B) = \frac{h}{1 - \beta} - \tau + gn + g' \quad (2-12)$$

The expected net income per unit time $E(A_p)$ is found by dividing the $E(P)$ by $E(T)$ as indicated in Eq.(2-1), hence:

$$E(A_p) = \frac{E(P)}{E(T)} = v_0 - E(L) \quad (2-13)$$

where the average total cost and loss per unit time is:

$$E(L) = \frac{E(C_c) + d \times E(B)}{E(T)} \quad (2-14)$$

and the control cost per unit time is:

$$E(A_c) = \frac{E(C_c)}{E(T)} = \frac{\alpha c' \frac{(1/\lambda - \tau)}{h}}{E(T)} + \frac{a + bn}{h} \quad (2-15)$$

Now our goal is to minimize $E(L)$ under the constraint of $E(A_c)$. Namely, after considering also the number of total "steps" s running at the same time:

$$\begin{aligned} \text{Min}_{(n_i, k_i, h_i)} E_T(L) &\equiv \text{Min}_{(n_i, k_i, h_i)} \sum_{i=1}^s E(L_i) \\ &= \text{Min}_{(n_i, k_i, h_i)} \sum_{i=1}^s \frac{E(C_{ci}) + d_i E(B_i)}{E(T_i)} \end{aligned} \quad (2-16a)$$

under the constraint:

$$E_T(Ac_i) = \sum_{i=1}^s E(Ac_i) = \sum_{i=1}^s \frac{E(Cc_i)}{E(T_i)} = \text{constant} \quad (2-16b)$$

The "worst case" control cost $W_T(Ac)$ would be estimated from Eqs. (2-8) and (2-9) after we have determined all (n_i, k_i, h_i) (with $n_i=1$ in our case). The detailed derivations will be given in the next section.

2.4 Numerical Implementation

Using the standard formulation of the minimization problem, we have the equivalent mathematical expression from which to determine $(k_i, h_i \mid i=1,2,\dots,s)$:

$$E_T(L) = \sum_{i=1}^s \frac{E(C_{ci}) + d_i \times E(B_i)}{E(T_i)} \quad (2-17a)$$

where,

$$E(T_i) = \frac{1}{\lambda_i} - \frac{h_i}{1-\beta_i} - \tau_i + g_i n_i + g_i' = \frac{1}{\lambda_i} + E(B_i) \quad (2-17b)$$

$$E(Cc_i) = c_i' \frac{\alpha_i \left(\frac{1}{\lambda_i} - \tau_i \right)}{h_i} + c_i + (a_i + b_i n_i) \frac{E(T_i)}{h_i} \quad (2-17c)$$

$$E(B_i) = \frac{h_i}{1-\beta_i} - \tau_i + g_i n_i + g_i' \quad (2-17d)$$

under the constraint:

$$E_T(Ac) = \sum_{i=1}^s \frac{E(Cc_i)}{E(T_i)} = \text{constant} \quad (2-18)$$

Now we define:

$$E_T(\epsilon) = E_T(L) - \epsilon \times E_T(Ac) \quad (2-19)$$

where ϵ is a dummy variable used here as a "Lagrangian multiplier". Therefore, our optimum equations are given by:

$$\begin{cases} \frac{\partial E_T(\epsilon)}{\partial k_i} = 0 \\ \frac{\partial E_T(\epsilon)}{\partial h_i} = 0 \end{cases} \quad i = 1, 2, \dots, s \quad (2-20)$$

Principally, we should be able to determine the $(2s+1)$ unknowns (k_i, h_i) from the $(2s+1)$ equations depicted in (2-20)'s and (2-18). However, in order to obtain a clearer "physical picture" of this model, as well as to make the numerical implementation feasible, we will find the approximate solutions to these equations. Towards this end, we will assume the out-of-control time in a cycle is much smaller than the in-control time, i.e., we will use:

$$E(T_i) = \frac{1}{\lambda_i} + E(B_i) = \frac{1}{\lambda_i} \quad (2-21)$$

in the denominators of the above equations. We will also consider that the testing (sampling) interval h_i is much smaller than the in-control time T , or, equivalently:

$$x_i \equiv \lambda_i h_i \approx \frac{h_i}{E(T_i)} \ll 1 \quad (2-22)$$

hence,

$$\frac{\tau_i}{h_i} = f(x_i) = \frac{1 - (1 + x_i) e^{-x_i}}{x_i (1 - e^{-x_i})} = \frac{1}{2} - \frac{1}{12} x_i + \dots \approx \frac{1}{2} \quad (2-23)$$

And from Eqs. (2-17) and (2-19), we have:

$$\begin{aligned} E_T(\varepsilon) &\equiv E_T(L) - \varepsilon \times E_T(Ac) = (1 - \varepsilon) E_T(Ac) + \sum_{i=1}^s d_i \frac{E(B_i)}{E(T_i)} \\ &\approx (1 - \varepsilon) \sum_{i=1}^s \left[\frac{c_i' \alpha_i + (a_i + b_i n_i)}{h_i} + \lambda_i c_i \right] \\ &\quad + \sum_{i=1}^s \lambda_i d_i \left[h_i \left(\frac{1 + \beta_i}{1 - \beta_i} \right) + (g_i n_i + g_i') \right] \end{aligned} \quad (2-24)$$

Before we apply Eq. (2-20), let's recall Eqs. (2-5) and (2-6),

$$\left\{ \begin{aligned} 1 - \beta &\equiv \Phi(-k + \delta \sqrt{n}) = \int_{-\infty}^{(-k + \delta \sqrt{n})} \phi(z) dz \\ \alpha &= 2\Phi(-k) = 2 \int_{-\infty}^{-k} \phi(z) dz \end{aligned} \right. \quad (2-25)$$

therefore,

$$\begin{cases} \frac{\partial}{\partial k} \left(\frac{1+\beta}{1-\beta} \right) = \frac{2}{(1-\beta)^2} \phi(-k + \delta\sqrt{n}) \\ \frac{\partial \alpha}{\partial k} = -2\phi(-k) = -2\phi(k) \end{cases} \quad (2-26)$$

Finally, we have the approximate results of Eq. (2-20) as:

$$- (1-\epsilon) \frac{2c_i'}{h_i} \phi(k_i) + \frac{\lambda_i d_i h_i}{1-\beta_i} \phi(-k_i + \delta_i \sqrt{n_i}) \approx 0 \quad (2-27)$$

$$- (1-\epsilon) \frac{c_i' \alpha_i + (a_i + b_i n_i)}{h_i} + \frac{\lambda_i d_i}{2} \left(\frac{1+\beta_i}{1-\beta_i} \right) \approx 0 \quad (2-28)$$

The outcome of all this that we have de-coupled the 2s equations in Eqs. (2-27) and (2-28) for different process steps. Hence, we can solve for the k_i s individually, by combining Eqs. (2-27) and (2-28):

$$[c_i' \alpha_i + (a_i + b_i n_i)] \phi(-k_i + \delta_i \sqrt{n_i}) = (1-\beta_i^2) c_i' \phi(k_i) \quad (2-29)$$

or,

$$\frac{\{1 - [\Phi(-k_i + \delta_i \sqrt{n_i})]^2\} \phi(k_i)}{\phi(-k_i + \delta_i \sqrt{n_i})} - 2\Phi(-k_i) = \frac{(a_i + b_i n_i)}{c_i'} \quad (2-29)'$$

where $i=1,2,\dots,s$. From Eq. (2-28), we also have:

$$h_i = \sqrt{\frac{c_i' \alpha_i + (a_i + b_i n_i)}{\lambda_i d_i \left(\frac{1+\beta_i}{1-\beta_i} \right)}} \cdot \Theta \quad (2-30a)$$

where,

$$\Theta = \frac{\sum_{i=1}^s \sqrt{[c_i' \alpha_i + (a_i + b_i n_i)] \cdot \left[\lambda_i d_i \left(\frac{1 + \beta_i}{1 - \beta_i} \right) \right]}}{E_T(AC) - \sum_{i=1}^s \lambda_i c_i} \quad (2-30b)$$

The computer programs in the Appendix will be based on Eq. (2-29) in order to first determine k_i , and subsequently determine h_i from Eq. (2-30). Finally Eq. (2-17) is used to verify the total average control cost. In the current implementation, the programs solve the optimization problem for $s=2$. It is quite straight-forward to extend our programs to $s>2$ cases.

3.0 Examples and Results

In this chapter we present a few simple examples that demonstrate the application of our model, with the help of the computer programs in the Appendix.

The programs need the following inputs:

- s = number of “multiple photolithographic steps”. We will assume $s=2$ for the time being, which could be any positive number in the future. Unit: none.
- $a+bn$ = cost of the sample (sample size $n=1$ in our case) for one sampling. Unit: \$. for a and \$/sample for b .
- c = cost of investigating a true alarm. Unit: \$.
- c' = cost of investigating a false alarm. Unit: \$.
- d = rate of loss of income due to the out-of-control process. Unit: \$/hour.
- $1/\lambda$ = average time for an assignable cause (of the out-of-control shift) to occur after the cycle starts. Unit: hour.
- δ = out-of-control shift of the process from its “good” mean μ . Unit: σ , where σ is the standard deviation of the in-control process. This value is assumed to be known and constant throughout this investigation.
- gn = time to take a sample and to interpret the results (of control chart). Again $n=1$ in our case. Unit: hour.
- g' = time to correct an assignable cause of shift after the investigation of a true alarm. Unit: hour.
- $E_T(AC)$ = the initial estimated average of the total control cost rate, which is (see Eq. (2-35) constrained by a maximum value. Unit: \$/hour.

The programs supply the following outputs:

- k = the Control limit factor for the upper/lower limit of the control charts. No units.
- h = interval between samples, or the time for one sampling period, which equals the time from the beginning of taking the current sample to the beginning of taking next sample. Unit: hour.
- α = type I error, which is a function of k . No units.
- β = type II error, which is also a function of k (and δ and n). No units.
- $E(B)$ = averaged time of the out-of-control process in one cycle. The total average cycle time $E(T)$ is thus $1/\lambda + E(B)$. Unit: hour.
- $E(A_C)$ = the average control cost for individual step. See Eq. (2-15). Unit: \$/hour.
- $E(L)$ = the average production cost for one step. See Eq. (2-14). Unit: \$/hour.

All the above input and output parameters (except s) are given for each of the s individual steps, even when the subscript i ($i=1,2,\dots, s$) is omitted. The following two outputs are the most important for all the steps:

- $E_T(A_C)$ = the average total control cost (for all the steps) of our design results, which is given by Eqs. (2-17b), (2-17c) and (2-18). Comparing these equations with Eq. (2-35), we know this output $E_T(A_C)$ is slightly smaller than the initial estimated average. Unit: \$/hour.
- $E_T(L)$ = the average total production cost for all steps (control cost, plus loss due to production while out-of-control). This is the parameter that needs to be minimized. Unit: \$/hour.

Other programming details and the programs as well are included in the Appendix.

To demonstrate the numerical implementation and practical use of our model, we have calculated some hypothetical examples. The results are plotted in Figs. 1 to 6 on the following pages.

In Fig. 1, we plot the type-I error $\alpha=2\Phi(-k)$, which can be easily found in any normal distribution table. Similarly, Fig. 2 shows the relationship between the sampling power $1-\beta$ and k (when $\delta=3$), as shown in Eq. (2-25).

In Fig. 5 and Fig. 6 we show a typical procedure to determine the (k_i, h_i) and hence the total production cost $E_T(L)$, as shown in Eqs. (2-29), (2-31) and (2-17). The input data are:

First, we use Eq. (2-29) to solve for $k=1.61$ as shown in Fig. 3, then, we have the relationship between $E_T(L)$ and the initial input total control cost constraint $E_T(A_C)$ changing from \$1/hour to \$100/hour. There is a lower limit for $E_T(A_C)$ otherwise the h_i might be negative (see Eq. (2-30)), which is un-physical. This is because we have to invest some initial amount in order to maintain the "statistical quality control" (SQC) process, or the X-chart.

	<i>Step 1</i>	<i>Step 2</i>
<i>a</i>	\$20	\$20
<i>b</i>	\$10	\$10
<i>n</i>	1	1
<i>c</i>	\$30	\$30
<i>c'</i>	\$50	\$50
<i>d</i>	\$100/hour	\$200/hour
$1/\lambda$	50 hour	50 hour
<i>g</i>	0.5 hour	0.5 hour
<i>g'</i>	1 hour	1 hour
$\mu \pm \delta\sigma$	$\mu \pm 3\sigma$	$\mu \pm 3\sigma$

From the curves in Fig. 4, we can see that the optimized value for $E_T(A_c) \cong \$16/\text{hour}$, which gives $(k_i, h_i) = (5.76 \text{ hour}, 4.07 \text{ hour})$ and the total minimum production cost is $E_T(L) = \$39/\text{hour}$.

In Fig. 5, we use the same data except that now $\delta \equiv \delta_1 = \delta_2$ changes from 2.8 to 3.0. We see the total optimized $E_T(L)$ inversely proportional to δ , as when δ increases, it is easier to detect the out-of-control state, or the power $1-\beta$ increases.

Finally, in Fig. 6, we still use the same input data as for Figs. 3 & 4, except $1/\lambda_1$ increases from 50 hours to 100 hours. By comparing Fig. 4 and Fig. 6, we see the $E_T(L)$ drops when the process stays in-control longer ($1/\lambda_1$ larger).

4.0 Conclusions

We have presented a well defined economical model in designing the X-chart in statistical quality control (SQC) process. The model is best suited for modern quality control inspection system on wafer production with multiple photolithographic steps.

We also demonstrated the feasibility of the numerical implementation of our theoretical model. The computer programs offered in this report give a convenient approach to determine the design parameters of the X-control charts, i.e., (k_i, h_i) , where i identifies the different lithographic steps.

Future extensions of our model will help in creating more substantial and practical applications in the field of Computer Aided Manufacturing.

5.0 Future Plans

There are many things in our model that we may improve in the future. Some of these more detailed considerations are addressed below.

5.1 Out-of-control Time Period

The out-of-control time before a true alarm is $B_\beta = M_\beta h$. When the process is out-of-control, the probability of having $(M_\beta - 1)$ missing alarms before the true alarm at the end of M_β 'th interval is given by the binomial distribution

$$P(M_\beta - 1) = (1 - \beta) \beta^{M_\beta - 1} \quad (M_\beta \geq 1) \quad (5-1)$$

The average number of missing alarms is $1/(1-\beta)$.

5.2 In-control Time Period

The total number of in-control sampling intervals is $G' = M_G h$. Note that G' is slightly smaller than G , the total in-control-time in a cycle: $E(G) = E(G') + \tau$. The probability that the process will stay in control until the end of the $(M_G + 1)$ 'th interval is given by the Poisson distribution:

$$P(M_G) = \lambda' \exp(-\lambda' M_G h) \quad (5-2)$$

The average length of the in-control run is $1/\lambda' = 1/\lambda - \tau$.

5.3 Modifications to $P(M_a)$

We can modify Eqs. (2-7) & (2-8) by considering Eqs. (5-1) & (5-2):

$$\begin{aligned} P(M_\alpha) &= \sum_{M_T=1+m_0}^{\infty} \left[\sum_{M_G=0}^{M_T-1-m_0} P(M_\alpha | M_G) \cdot P(M_G, M_T) \right] \\ &= \sum_{M_T=1+m_0}^{\infty} \left[\sum_{M_G=0}^{M_T-1-m_0} P(M_\alpha | M_G) \cdot P(M_\beta) \cdot P(M_G) \right] \end{aligned} \quad (5-3)$$

where $M_\beta \equiv M_T - M_G - m_0$ with $m_0 = gn + g'$ as a constant. The last equation in Eq. (5-3) comes from:

$$P(M_T) = P(M_G + M_\beta + m_0) = \sum_{M_G, M_\beta} P(M_G) \cdot P(M_\beta) \quad (5-4)$$

where $M_G + M_\beta \equiv M_T - m_0$, and $P(M_G)$ and $P(M_\beta)$ are given by Eqs. (5-1) & (5-2), respectively. Similarly to Eq. (2-8), we have the binomial distribution:

$$P(M_\alpha | M_G) = \frac{(M_G!)}{M_\alpha! (M_G - M_\alpha)!} \alpha^{M_\alpha} (1 - \alpha)^{M_G - M_\alpha} \quad (5-5)$$

hence we may obtain $P(M_\alpha)$ from Eq. (5-3).

Finally, concerning the determination of the extremum, we may inspect the 2nd order derivatives of $E_T(\epsilon) = E_T(L) + E_T(Ac)$ with respect to $(k_i, h_i \mid i=1,2,\dots,s)$ and their combinations. Only if all of those derivatives are not negative then we can formally say that we have found the minimum value of $E_T(L)$.

Also, we may generalize our computer programs for $s > 2$ to accommodate any positive number of "photolithographic steps".

6.0 Figures

FIGURE 1. Type I Error $\alpha = \Phi(-k)$

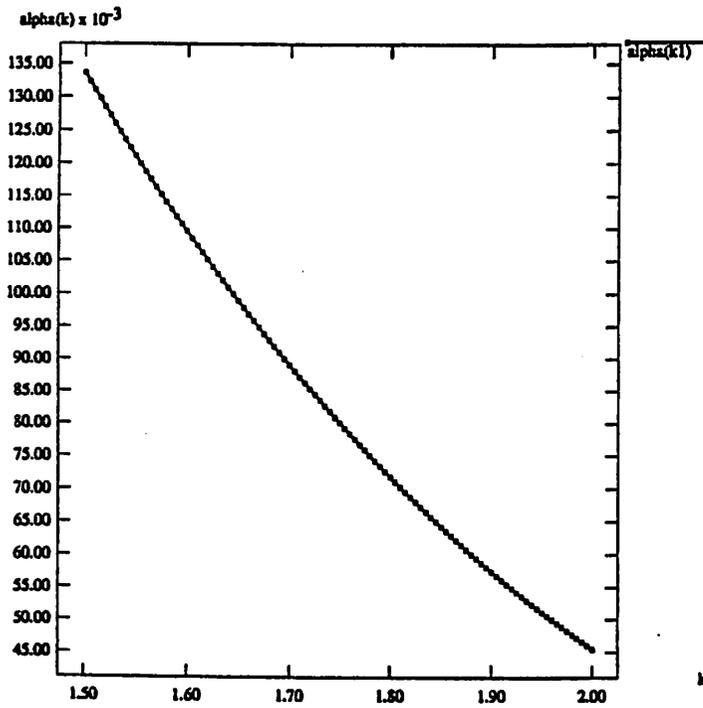


FIGURE 2. Power of Sampling with $\delta = 3$.

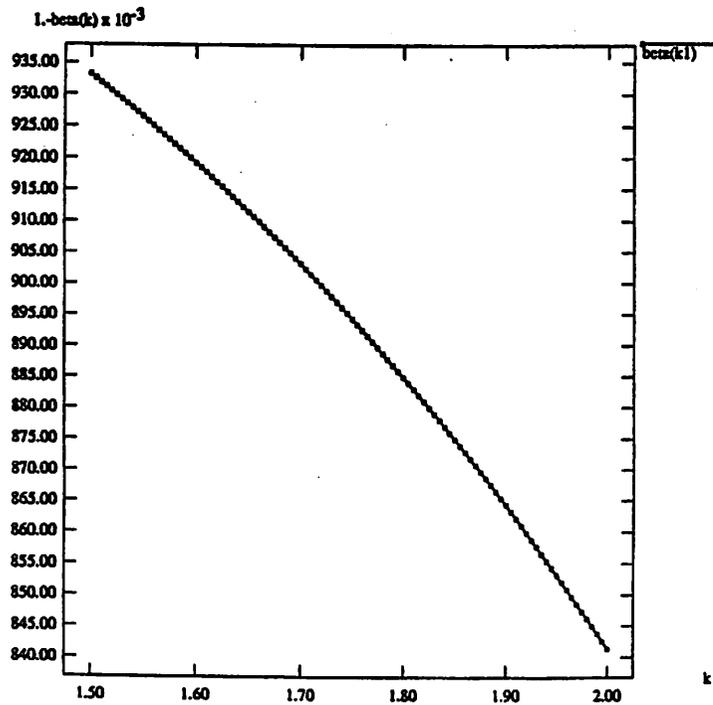


FIGURE 3. Solving for k (Eq. 2-29), $\delta = 3$, $(\alpha + bn) / c' = 0.6$

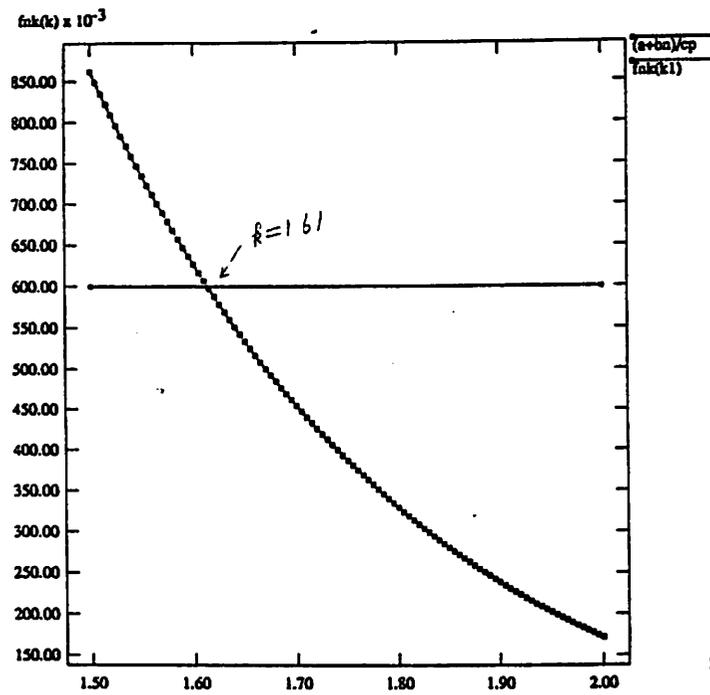


FIGURE 4. Total Production Cost versus Control Cost ($\lambda_1 = \lambda_2$)

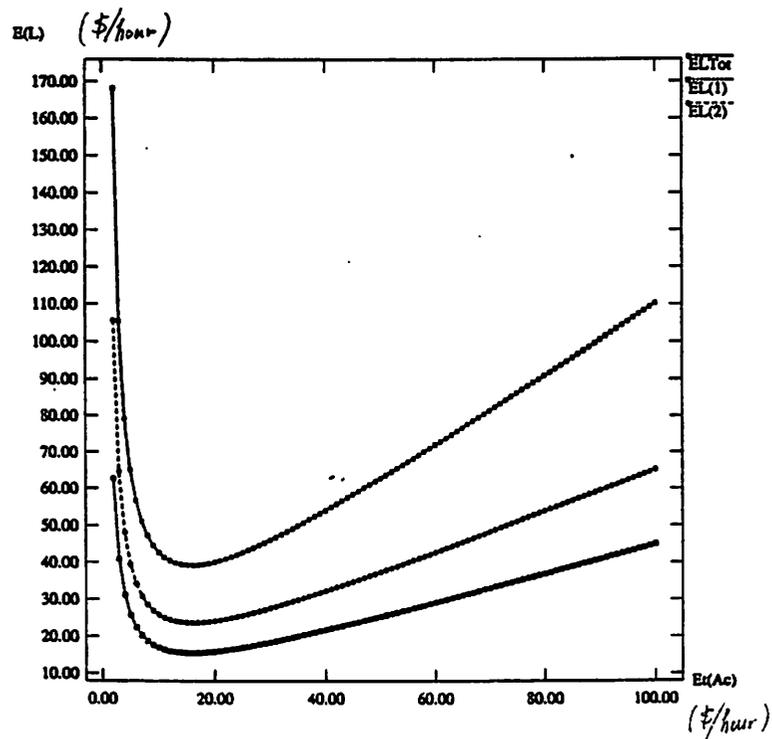


FIGURE 5. Total Control Cost versus Shift δ .

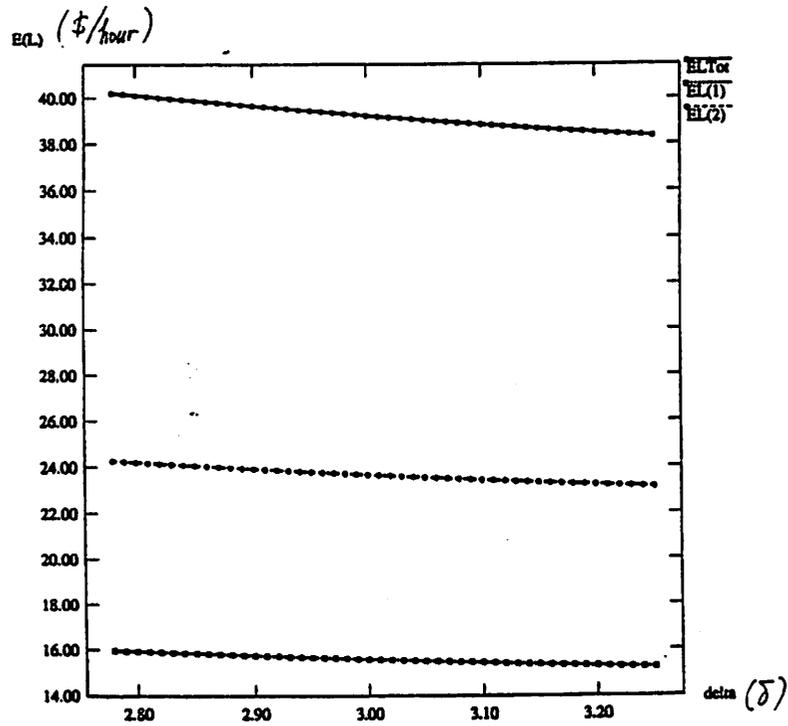
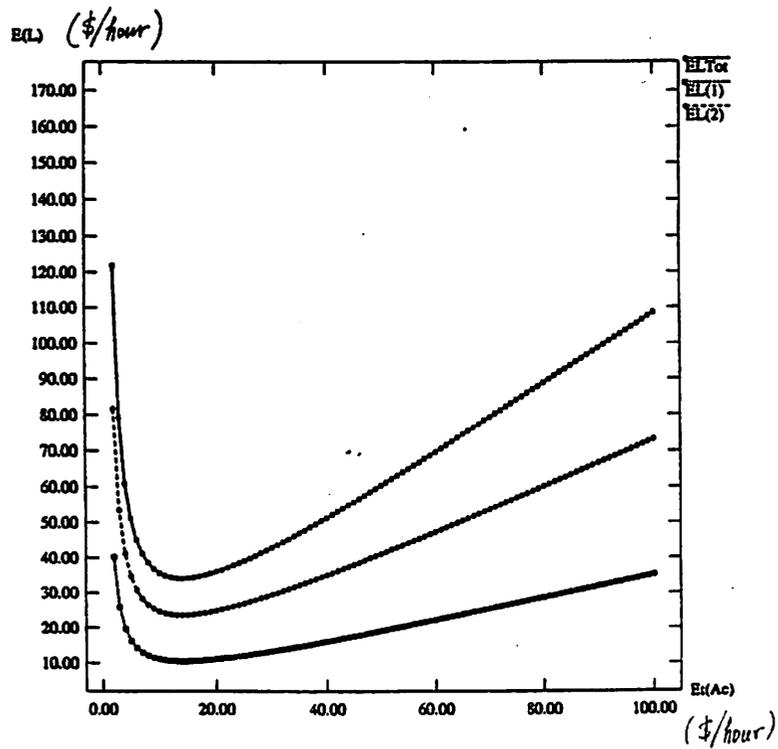


FIGURE 6. Total Production Cost versus Control Cost (λ_1 not equal to λ_2).



pk.f

30/17/75
1004331

3

```

fuk(i) = (1.-beta(i))*beta(i) * ordn(xk(i))
fuk(i) = fuk(i) / ordn(bbb) - alpha(i)
abcn = (a(i)+b(i))*xn(i) / cp(i)
c   fuk(i) = fuk(i) - abcn
10  continue
c   return
    end
c   -----
c   This subroutine gives h, E(Ac), E(L) from the
c   k determined in subroutine k1k2
c   -----
subroutine h1h2(xk, alpha, beta)
real a(2), b(2), c(2), cp(2), d(2)
real xlam(2), del(2), g(2), gp(2)
real cnrl, xn(2)
real xk(2), alpha(2), beta(2)
real fuk(2), xk(2), alpha(2), beta(2)
real h(2), EB(2), EAc(2), EL(2), EAcT, ELT
real ET(2), MT(2)
common /ETMT/ET,MT
common /hEh,EB,EAc,EL,EAcT,ELT
common /fuk/fuk,xk,alpha,beta
common /abcd/a,b,c,cp,d
common /spec/xlam,del,g,gp
common /cnrl/cnrl,xn
rr1 = xlam(1)*d(1)*0.5*(1.+beta(1))/(1.-beta(1))
rr2 = xlam(2)*d(2)*0.5*(1.+beta(2))/(1.-beta(2))
sr = sqrt(rr1/rr2)
abc1 = a(1) + b(1)*xn(1) + cp(1)*alpha(1)
abc2 = a(2) + b(2)*xn(2) + cp(2)*alpha(2)
sabc = sqrt(abc1 * abc2)
c   ctm = cnrl - ( xlam(1)*c(1) + xlam(2)*c(2) )
h(1) = ( abc1 + sabc / sr ) / ctm
h(2) = ( abc2 + sabc * sr ) / ctm
c   do 30 i=1,2
xxx = xlam(i)*h(i)
EB(i) = 1./(-beta(i)) - f(xxx)
EB(i) = EB(i) * h(i) + g(i)*xn(i) + gp(i)
c   ccp = c(i) + cp(i)*alpha(i)* (1.-f(xxx) * xxx) / xxx
ET(i) = 1./xlam(i) + EB(i)
MT(i) = ET(i) / h(i)
EAc(i) = ( a(i)+b(i)*xn(i) ) / h(i) + ccp / ET(i)
EL(i) = EAc(i) + d(i) * EB(i) / ET(i)
30  continue
c   EAcT = EAc(1) + EAc(2)
ELT = EL(1) + EL(2)
return
end

```

pk.f

30/17/75
1004331

```

write(i+10,'0')
write(i+20,'0')
20  continue
800 format(81(1h-)' i, fuk, k-star, alpha, beta ;
& i2, 4f11.5 )
c   -----
c   write(6,*)
do 71 j1=1,j1
do 72 j2=1,j2
xk0(1) = xk1(j1)
alpha0(1) = alpha1(j1)
beta0(1) = beta1(j2)
xk0(2) = xk1(j2)
alpha0(2) = alpha1(j2)
beta0(2) = beta1(j2)
call h1h2(xk0, alpha0, beta0)
write(6,900) j1, j2, h, EB
write(6,901) EAc, EAcT
write(6,902) EL, ELT
write(6,903) ET, MT
72  continue
71  continue
900 format(81(1h-)' j1, j2, h(0), EB(0) ;
& 2i2, 4f11.5 )
901 format( 4x, , EAc(0), EAcT ,
& 4x, 2f11.5, 11x, f11.5 )
902 format( 4x, , EL(0), ELT ,
& 4x, 2f11.5, 11x, f11.5 )
903 format( 4x, , ET(0), MT(0) ,
& 4x, 4f11.5 )
stop
end
c   -----
c   This subroutine is for determining optimized k value
c   -----
subroutine k1k2
real a(2), b(2), c(2), cp(2), d(2)
real xlam(2), del(2), g(2), gp(2)
real cnrl, xn(2)
real fuk(2), xk(2), alpha(2), beta(2)
real h(2), EB(2), EAc(2), EL(2), EAcT, ELT
common /hEh,EB,EAc,EL,EAcT,ELT
common /fuk/fuk,xk,alpha,beta
common /abcd/a,b,c,cp,d
common /spec/xlam,del,g,gp
common /cnrl/cnrl,xn
c   do 10 i=1,2
alpha(i) = 2.*pnorm(-xk(i))
bbb = del(i) * sqrt(xn(i)) - xk(i)
beta(i) = 1.0 - pnorm(bbb)

```

5

pk.f

```

c
c-----
c f(x) is the ratio of 'neu and h
c-----
c      function f(x)
c      f = (1.- (1.+x)* exp(-x) )
c      & / ( x*(1.- exp(-x)) )
c      return
c      end
c
c-----
c Normal distribution density
c-----
c      function ordn(z)
c      ordn=0.39894228* exp(-z*z/2.)
c      return
c      end
c
c-----
c Normal distribution probability function
c-----
c      function pnorm(x)
c      dimension cc(7)
c      data cc/3.19381530, -.356563782, 1.781477937, -1.821255978,
c      & 1.330274429, .2316419, 2.506628725/
c      y=x
c      if(x.lt.0.) y=-x
c      i=1./(1.+cc(6)*y)
c      s=(((cc(5)+cc(4)))*i + cc(3))*i +cc(2))*i + cc(1))*i
c      pnorm=s*exp(-y*y/2.) /cc(7)
c      if (x.gt.0.) pnorm=1.-pnorm
c      return
c      end
c-----

```

7.0 References

- [1] Duncan, A.J., "The Economic Design of X-Charts Used to Maintain Current Control of a Process," *Journal of the American Statistical Association*, Vol.51, pp228-242 (1956)
- [2] Goel, A.L., S.C. Jain, and S.M. Wu, "An Algorithm for the Determination of the Economic Design of X-chart Based on Duncan's Model," *Journal of the American Statistical Association*, Vol.62, pp304-320 (1968)
- [3] Chiu, W.K., and G.B. Wetherill, "A Simplified Scheme for the Economic Design of X-Charts," *Journal of Quality Technology*, Vol.6, No.2, pp63-69 (1974)
- [4] Montgomery, Douglas C., "Economic Design of an X-Control Chart," *Journal of Quality Technology*, Vol.14, No.1, (1982)
- [5] Montgomery, Douglas C., "Introduction to Statistical Quality Control" (2nd Edition), Chapter 10, John Wiley and Sons, (1990)

Feedback Control for a Photolithographic Workcell

Sovarong Leang

Our goal is to apply supervisory control on the definition of critical dimensions during optical lithography. At this stage, only the feedback control loop for the spin-coat and bake station has been established. The methodology is based on the monitoring and modeling the thickness and the reflectance of the photoresist applied on oxide and silicon. The dependence of these parameters on the process settings has been modeled empirically using statistically designed experiments. The resulting models form the core of the control algorithm, since they permit us to detect a change in the process, and allow us to modify the process settings in order to maintain control over the critical dimensions of the transferred pattern.

1.0 Introduction

In a typical CMOS process there are over 10 masks, and each one of the respective photolithographic steps is subject to small changes which can be damaging to the final product. Being able to control each of the photo steps is crucial in a high yield production environment. In the Berkeley Computer Aided Manufacturing group we are developing a supervisory control system for a photolithographic workcell. This is accomplished by using a monitoring scheme that measures the thickness and the reflectance of the photoresist after the deposition, bake and exposure of the photoresist layer. We then employ models of the lithography equipment in order to establish a control system with feedback and feed-forward control capabilities [1], [2] as shown in Fig.1.

Thus, if a wafer is misprocessed by the spin-coat equipment, we can remedy the problem by establishing a feed-forward control loop, so that the settings of the exposure station will be adjusted in order to compensate for the spin-coat problem. If the process change turns out to be a permanent process drift, we can prevent problems on future wafers by establishing permanent compensating corrections on the concerned machine.

In this paper, we describe in detail the feedback control methodology that is used to control the spin-coat and bake station. A similar feedback control will be established for the exposure and the develop stations in the near future.

2.0 Monitoring Methodology

Most existing in-line monitoring schemes record only the thickness of the photoresist after its application. Our methodology includes one additional parameter, the photoactive compound concentration M in the resist. As indicated in the models within SAMPLE [3] and PROLITH [4], M depends strongly on the baking time, baking temperature, exposure time, and other parameters as well. During development M determines the dissolution rate and therefore the critical dimensions (CD) of the developed image. Although M cannot be measured directly in-line, it can be inferred from the reflectance of the film.

Measuring reflectance is difficult, since it depends on the film thickness, the oxide thickness, and the wavelength of the beam that measures it. The film reflectance is also very sensitive to variations in the film thickness, because of the creation of interference patterns within the transparent film. In order to decouple the resist reflectance from its dependence on the thickness of the resist and of the underlying oxide layer, we developed a novel measuring scheme [5]. According to this scheme, we must first identify the wavelength that will result in the maximum measured reflectance. This wavelength is determined by modeling the thin film as a finite combination of parallel plane sub-layers. This set of sub-layers has a known refractive index and absorption coefficient. For the j -th layer, the complex reflection, r_{j-1} , is given by:

$$r_{j-1} = \frac{\exp(-2i\phi_j) [(F_j r_j) - F_j(1 - F_j r_j)]}{F_j [\exp(-2i\phi_j)] [(F_j - r_j) - (1 - F_j r_j)]}$$

where

$$F_j = \frac{n_0 - n_j}{n_0 + n_j}$$

is the classic Fresnel coefficient relative to air, n_j is the complex index of refraction of the j -th layer, n_0 is the refractive index of air, and

$$\phi_j = \frac{2\phi}{\lambda} n_j x_j$$

is the optical phase thickness of each layer. Also, x_j is the thickness of the j -th layer, λ is the wavelength of the light beam, and n_j is the real part of the refractive index of the j -th layer. The apparent reflectance of the entire film, defined as the intensity ratio of the incident and the reflected beams, is given by:

$$R = |r_0|^2$$

where r_0 is the complex reflectance at the air-photoresist interface. Thus, starting from the substrate interface, we calculate the complex refractive index r_j from Eqs. 1, 2 and 3

for each dielectric layer and resist sub-layer until we reach the photoresist surface. This way, we finally obtain the apparent reflectance of the entire film as given by equation 4. In our case, we only have one layer of oxide between the substrate and the photoresist.

In conclusion, the measured film reflectance depends on three parameters: (1) the oxide thickness; (2) the resist thickness; and (3) the wavelength of the beam at which it is measured. In order to make the reflectance measurement depend only on the photoactive compound concentration, a wavelength is chosen that yields the maximum measured reflectance. This way, the reflectance measurement is decoupled from its direct dependence on resist and oxide thickness.

In order to determine this optimum wavelength, we plot the maximum reflectance with respect to resist thickness and wavelength at a fixed oxide thickness. Several such plots are developed for a range of oxide thicknesses, as shown in Fig 2. For each oxide thickness, the maximum reflectance is a linear function of resist thickness and wavelength. At this point, the slope and the y-intercept of these functions is be calculated and tabulated versus the respective oxide thickness. The theoretical wavelength at which the reflectance is maximized can then be found from the measured resist thickness and the linear function that corresponds to the measured oxide thickness. In summary, the monitoring methodology consists of the following steps:

1. Solve for the theoretical maximum reflectance and generate the reflectance versus resist thickness and wavelength plots for a range of oxide thickness.
2. Measure the oxide thickness before we apply the photoresist on the wafer.
3. Measure the film thickness after spin-coating the wafer.
4. Using the previously described reflectance plot for the measured oxide thickness, calculate the wavelength at which we should measure.
5. Measure the reflectance at the prescribed wavelength, and also at two "bracketing" wavelengths set at five nanometers above and below the theoretical maximum. The true maximum reflectance is found by fitting a parabolic interpolation through the three measured values.

This methodology depends on accurate measurements of the film reflectance. Accurate and reliable measurements require however a longer scanning time of the machine, which might result in inadvertent resist exposure. Through experimental analysis, we find that one full second is necessary for the measuring instrument to record an accurate and reliable reflectance measurement. Fortunately, the received dose is such that the resist is not being exposed.

3.0 Modeling Methodology

Empirical models of the spin-coat and bake equipment have been developed using a two-level full factorial design of experiments with three centerpoint replications. The inputs to the models are the process settings, which include the spin speed, the spin time, the baking temperature, and the baking time of the wafer. The outputs of the models are

the photoresist thickness and reflectance. Although empirical models do not add to our physical understanding of the process, they are simple, and they are accurate over the region covered by the experiments.

The models, obtained through step-wise regression, are linear to the estimated coefficients. The reflectance model has ten terms, and the thickness model has six terms. Each model has been analyzed through an Analysis of Variance (ANOVA) table, as shown in Fig. 3. This analysis reveals the replication error of the machine, as well as the prediction error of the model. The thickness model predicts the mean response of the equipment with a one sigma *prediction* error of +/- 38 Å on the average, while the actual response of the equipment varies around its mean value with a one sigma *replication* error of +/- 65 Å. The *F*-distribution test shows that this model is highly significant, since the probability that the residual explained by the model is practically zero. We draw this conclusion by examining the ANOVA table, where we see that the probability of $F(5,70) > 6516.0$ is extremely small.

A scatterplot of the resist thickness values predicted by the model, versus the corresponding experimental values is given in Fig. 4a. A similar plot is given in Fig. 4b for the reflectance model. These models are very important for our feedback control strategy and their role is described next.

4.0 Feedback Control Methodology

Feedback correction is implemented when there is a significant process drift. A “drift” is defined as a consistent process change that persists over a prolonged period of time. Such a prolonged change is identified with the help of the Cumulative Student-t Statistic. Here, we use a threshold of 2.0, at which point the level of confidence that the process has indeed deviated from its past position is 95%. The Student-t statistic is given by:

$$t = \frac{\sum_{i=1}^n (Y_i - y'_i)}{\text{Var} \left[\sum_{i=1}^n (Y_i - y'_i) \right]}$$

where Y'_i is the predicted response from the model, and y'_i is the actual response from experiments for the *i*-th wafer [6]¹.

Once we identify the process change, the model must be updated to reflect the present process, which has drifted from its original operating point. The model is updated by

changing its constant term to a new value, so that the sum of the squared residuals is minimized. All the other terms of the model are kept the same, since the information obtained from a single-recipe process only gives us one degree of freedom². Once the model has been updated, we solve for the new settings of the process using our updated model. This is accomplished by minimizing the sum of the squared differences between the process targets and the model predictions.

5.0 An Example of Feedback Control in the Berkeley Microfab

In this section we describe an actual experiment performed in the Berkeley Microfabrication Laboratory. Twelve wafers were coated with 1000Å thermal oxide. We used the flat of the wafer as the reference, and located four test points on the vertical and horizontal axis of the wafer, at the same radius from the center. Each wafer was given a unique number, so that we could follow the process precisely at each test area. The oxide was then measured for each test area and recorded.

Next, we ran the twelve wafers as a batch through the spin-coat and bake equipment with the following standard recipe: The photoresist used was the positive resist KTI 820. We used a spin speed of 4600 +/- 30 RPM, a spin time of 30 sec, a baking temperature of 120 +/- 2°C, and a baking time of 60 sec. We then measured the thickness and the reflectance of each of the four test sites on each wafer, using the monitoring method described above. Afterwards, the wafers were exposed in a GCA I-line stepper, and developed using an MTI Omnichuck. Finally we optically measured the critical dimensions of the 2µm resist lines on each wafer test cite.

Once the data was recorded, we filtered out any points that were significantly different from the others. This was done by using a standard R-chart, as shown in Fig. 5. Data which showed a range that was greater than the upper control limit, set at the 95% level of confidence, were examined in more detail. To lower the range of the measurements on those wafers, the data point which lay the furthest away from the group was eliminated, and a new R-chart was produced. This procedure was repeated until all outliers have been eliminated. In some cases, when there was n obvious outlier point, the data was kept, even when the range of the measurements was higher than the upper control limit. The average response of each wafer was then calculated, and a control chart using the cumulative Student-t statistic was developed for the film thickness and its reflectance. These charts are shown in Fig. 6, along with a Student-t statistic chart that was developed for the final response, the critical dimension of the photoresist pattern.

1. At this point, the reader should note that this control scheme can be applied over a range of equipment settings, since it effectively controls the residuals of the process response. These residuals can be calculated as the difference between the process measurement and the prediction of the equipment model. This scheme of "model-based S PC" was originally introduced in [6]. Such a scheme is highly appropriate for a supervisory control environment, where the recipes are adjusted automatically and continuously in order to accommodate known equipment changes.

2. After multiple process corrections, several degrees of freedom become available. Therefore, the re-evaluation of the process model can include updated values for several of the model parameters, i addition to the constant of the model. This issue will be addressed in the future and it is outside the scope of this report.

In the second part of the experiment, we introduced some faults, in order to simulate a process drift: the baking time was increased to 80 sec, and the spin time was increased to 90 sec. We ran an additional twelve wafers and we monitored them as before. The same method of filtering out data outliers was used to reject 10 to 20% of all the data³. The responses of these wafers were then analyzed using cumulative Student-t control charts. As it can be seen in Fig. 6a, the thickness chart, by exceeding its 2.0 control limit, generated an alarm on wafer #18.

In response to this alarm, in the third part of the experiment, we solved for the feedback recipe using the models and the methodology described above. Since we introduced the fault by changing some of the settings, we worked with recipe adjustments (i.e setting corrections) rather than absolute values: In our experiment, the feedback correction recommended a spin time reduction by 25 sec, and this resulted in a 65 sec absolute setting for the feedback recipe. For the baking time, the feedback correction came out to -10 sec, which resulted in a 70 sec absolute setting in the feedback recipe.

Twelve wafers were then run after the correction was made. The average critical dimension was observed to have changed back toward its old value, but it overshot it (Fig. 7a-7c). The reason was because the correction that the model predicted would be needed to remedy the situation was quite drastic: a baking temperature of 105°C was recommended and used. That baking temperature was the lower limit of the temperature when we designed our experiments to model the spin-coat and bake machine. Therefore, the model would not be very accurate at such an extreme point. The methodology proved however that the critical dimensions can be brought back towards their old values, and with a second iteration, we are confident that we can get closer to our target. This experiment is currently under way and it will be reported in the near future.

6.0 Conclusion

In conclusion, a feedback control methodology has been developed, and it has been shown to correct inadvertent drifts in the critical dimension during optical lithography. The precision of the feedback control depends however on the accuracy of the model used. In some cases, more than one iteration might be necessary. Although this method is currently based on correcting long term drifts on the spin-coat & bake equipment, it can be extended to a complete supervisory control scheme with feed-forward and feedback capabilities over a number of steps within a photolithographic workcell.

7.0 Acknowledgment

The author wishes to thank the personnel and management of the Berkeley Microfabrication Laboratory, for their help in preparing the samples used in this project.

3. This relatively high percentage of outliers is a result of our measurement techniques and equipment. It is expected, that as our reflectance measuring techniques mature, and as we move into more automated testing equipment, the percentage of the outliers will drop.

8.0 Figures

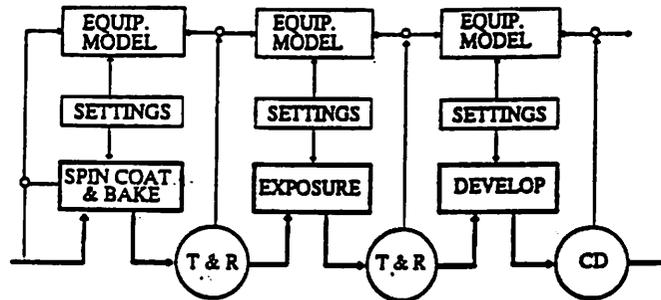


Figure 1. Lithography Control Using In-Line Measurements of Resist Thickness and Reflectance.

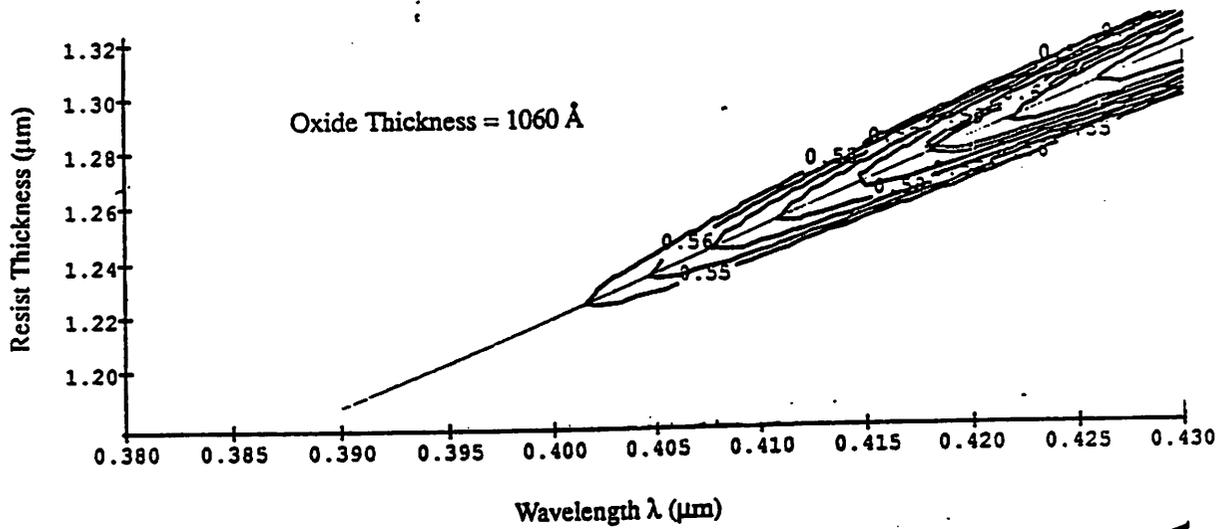


Fig. 2a.

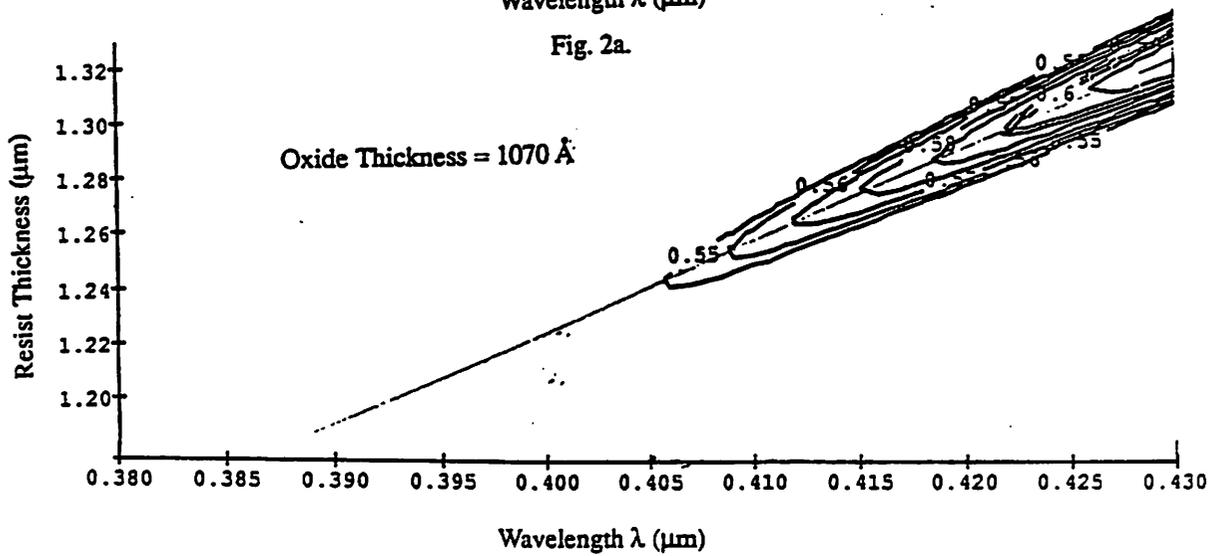


Fig. 2b. Plot of Reflectance versus Resist Thickness and Wavelength

0 Source	1 df	2 Sum Sq.	3 Mean Sq.	4 F-Ratio	5 Sig.Lev.
1 Total (Corr)	75	1e+08	1814605		
2 Regression	5	1e+08	27160721	6516.000	0.000
3 Residual	70	291764	4168.060		
4 Lack of Fit	12	247572	20630.988	27.080	0.000
5 Pure Error	58	44192.333	761.937		

R-SQUARE = 0.9979 Adjusted R-SQUARE = 0.9977

F(5,70) as large as 6516 is a very rare event =>
 highly unlikely that all coefficients are zero.
 F(12,58) as large as 27.08 is a very rare event =>
 highly unlikely that model is correct.
 Estimate of Pure Error from 17 groups of replicates.

Fig. 3a. Analysis of Variance Table for the Thickness Model

0 Source	1 df	2 Sum Sq.	3 Mean Sq.	4 F-Ratio	5 Sig.Lev.
1 Total (Corr)	55	12384251	225168.197		
2 Regression	9	11312551	1256950	53.950	0.000
3 Residual	46	1071700	23297.820		
4 Lack of Fit	5	721134	144226.804	16.870	0.000
5 Pure Error	41	350566	8550.382		

R-SQUARE = 0.9135 Adjusted R-SQUARE = 0.8965

F(9,46) as large as 53.95 is a very rare event =>
 highly unlikely that all coefficients are zero.
 F(5,41) as large as 16.87 is a very rare event =>
 highly unlikely that model is correct.
 Estimate of Pure Error from 13 groups of replicates.

Fig. 3b. Analysis of Variance Table for the Reflectance Model

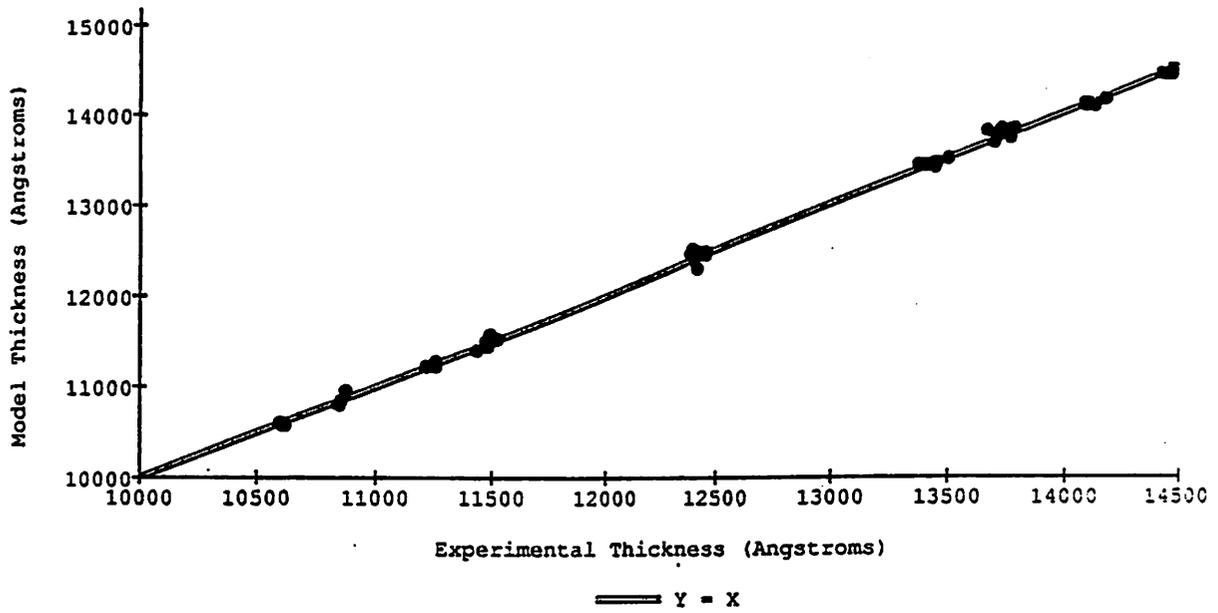


Fig. 4a. Scatterplot of the Predicted Resist Thickness versus the Experimental Thickness

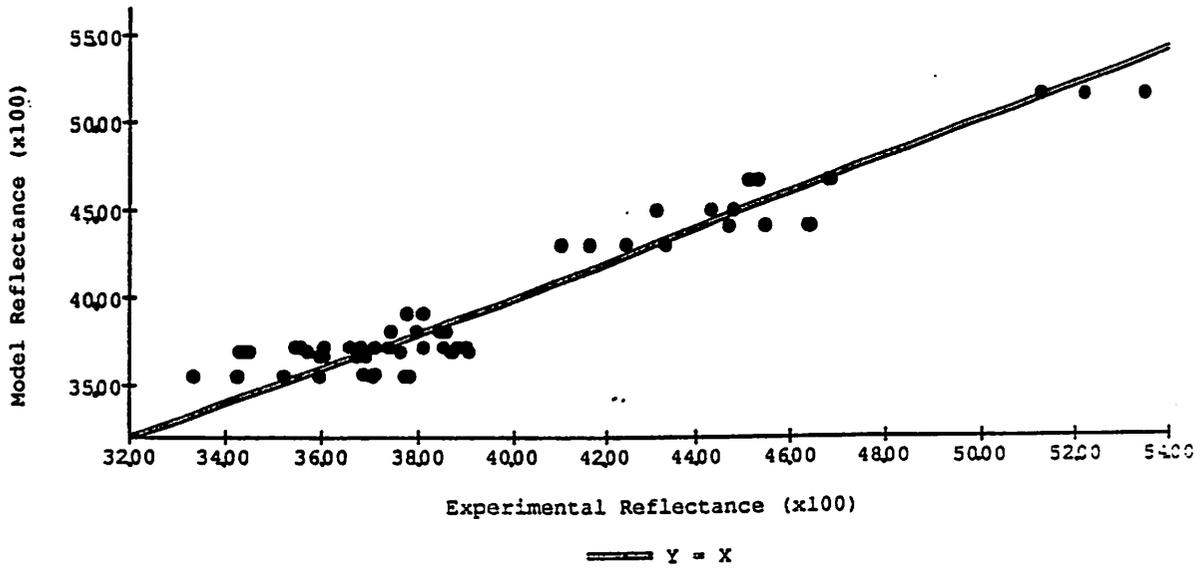


Fig. 4b. Scatterplot of the Predicted Reflectance versus the Experimental Reflectance

Fig. 5a. R-Chart for the Resist Thickness

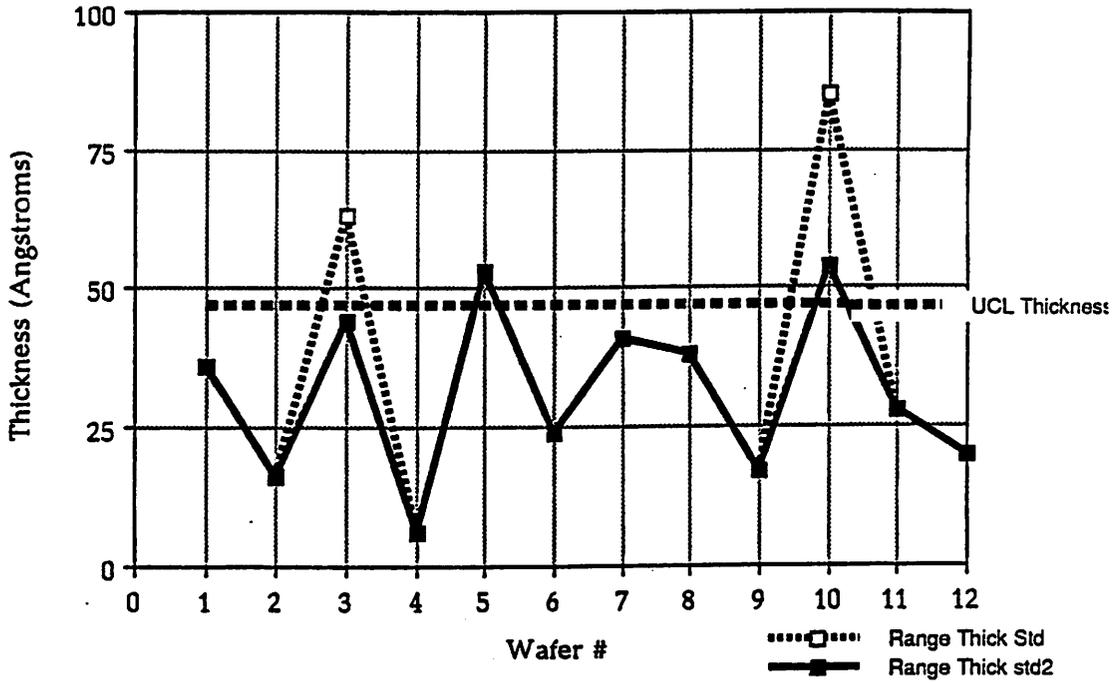


Fig. 5b. R-Chart for the Reflectance

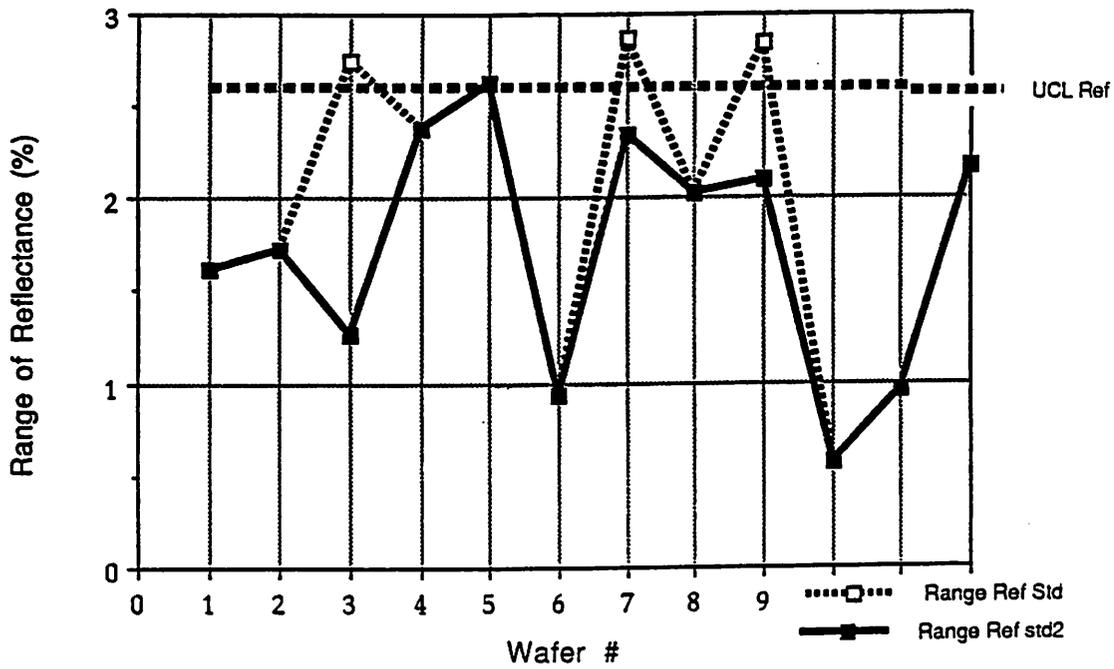


Fig. 5c. R-Chart for CD

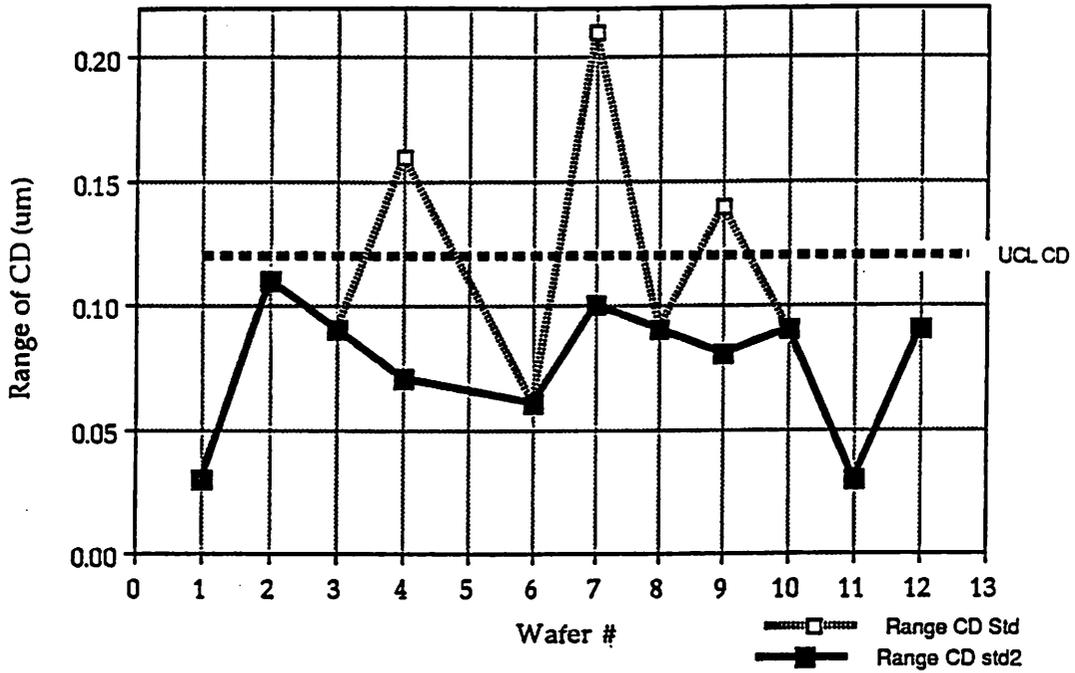


Fig. 6a. Cumulative Student-t Chart for Photoresist Thickness

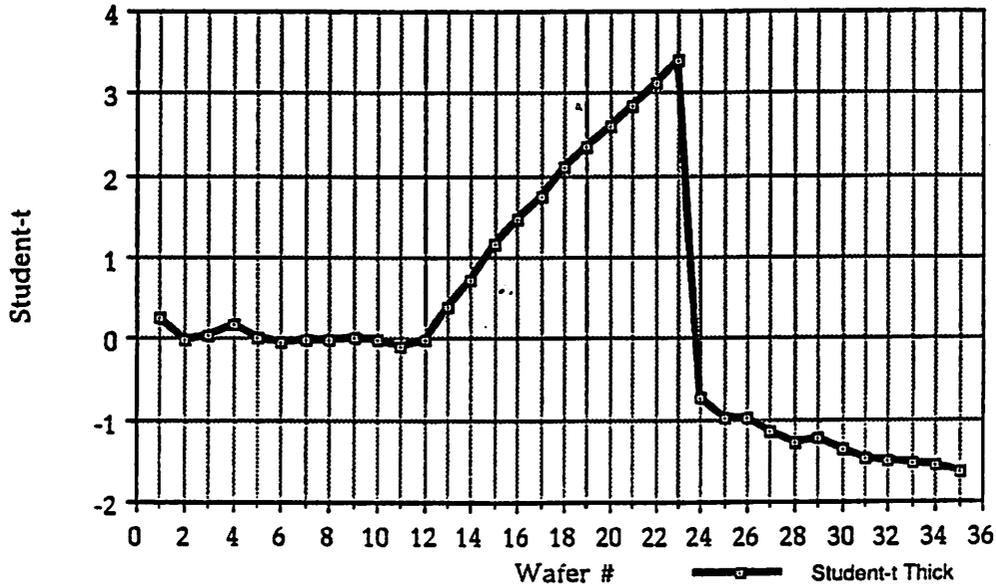


Fig. 6b. Cumulative Student-t Chart for Reflectance

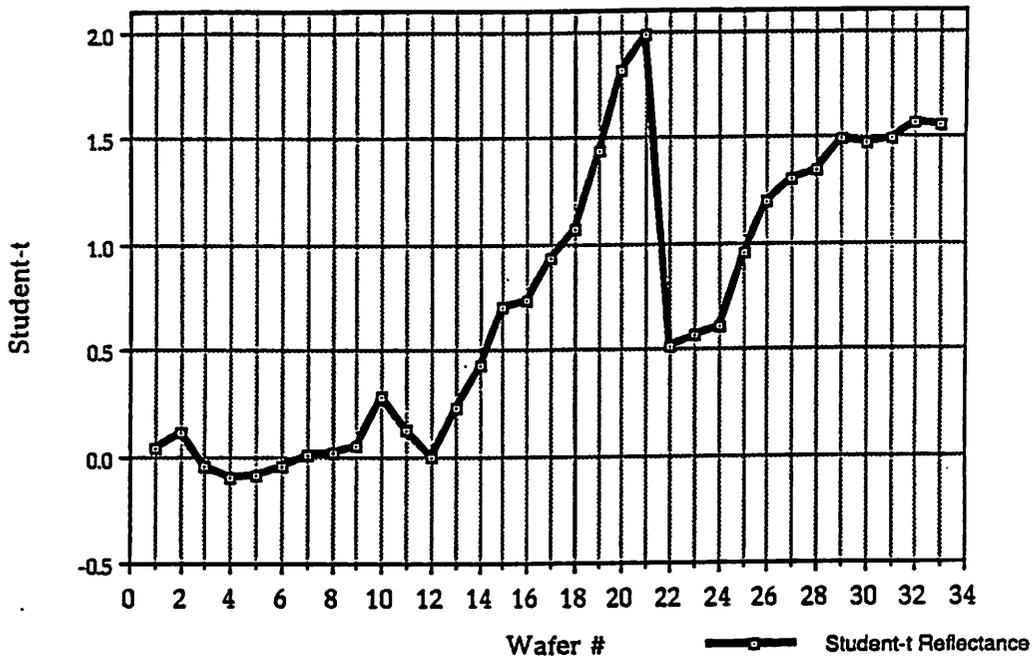


Fig. 6c. Cumulative Student-t for Critical Dimensions

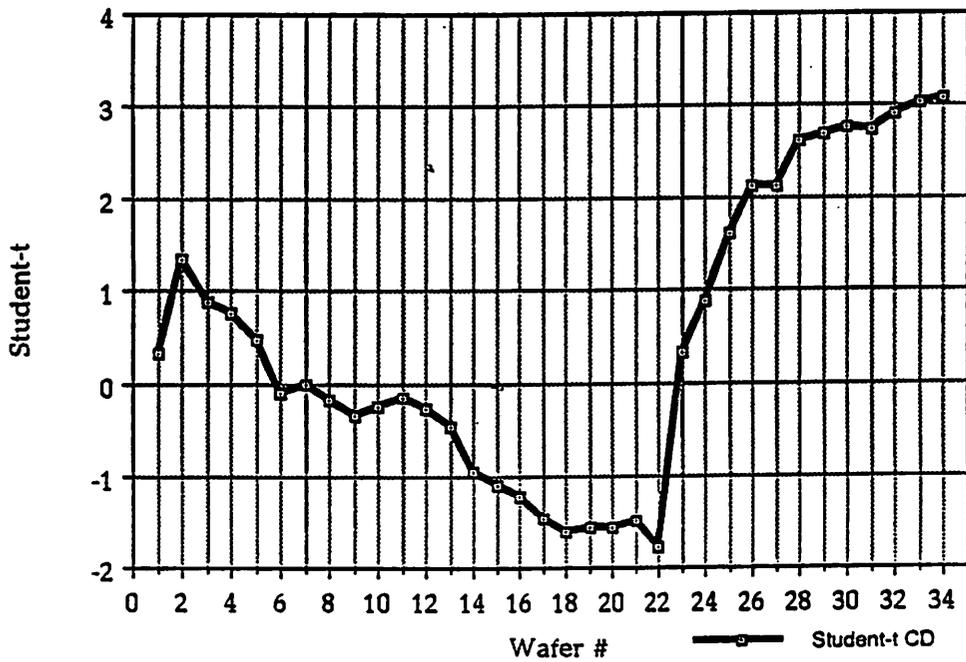


Fig. 7a. Thickness Chart

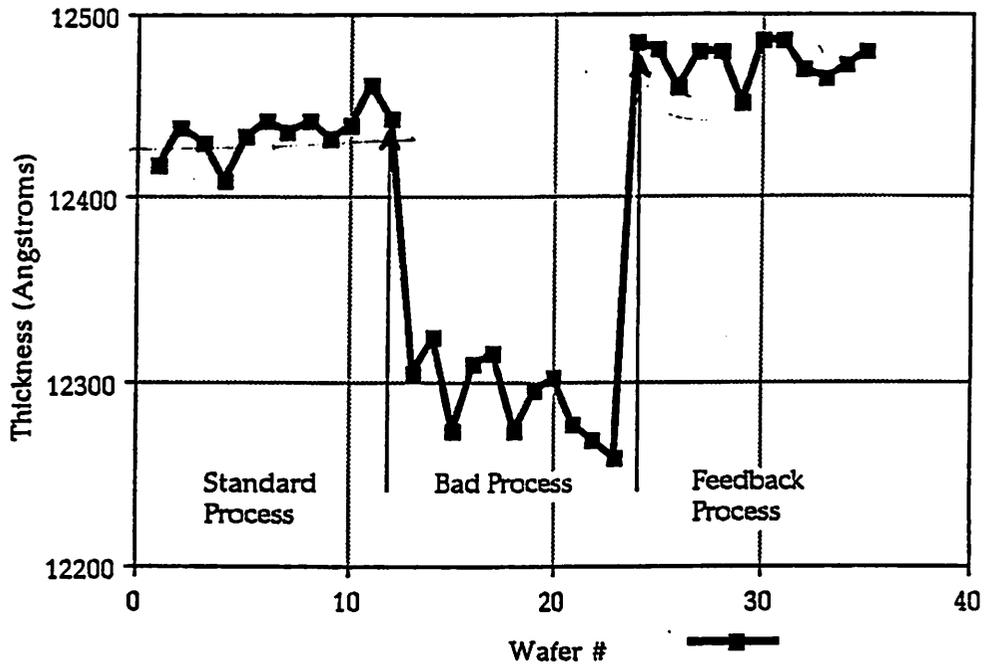


Fig. 7b. Reflectance Chart

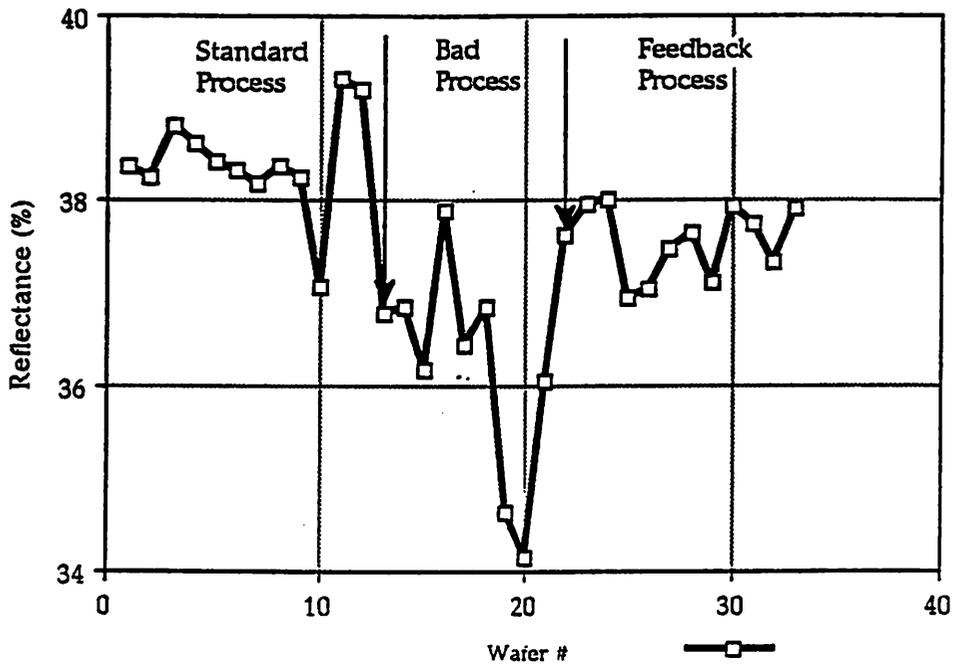
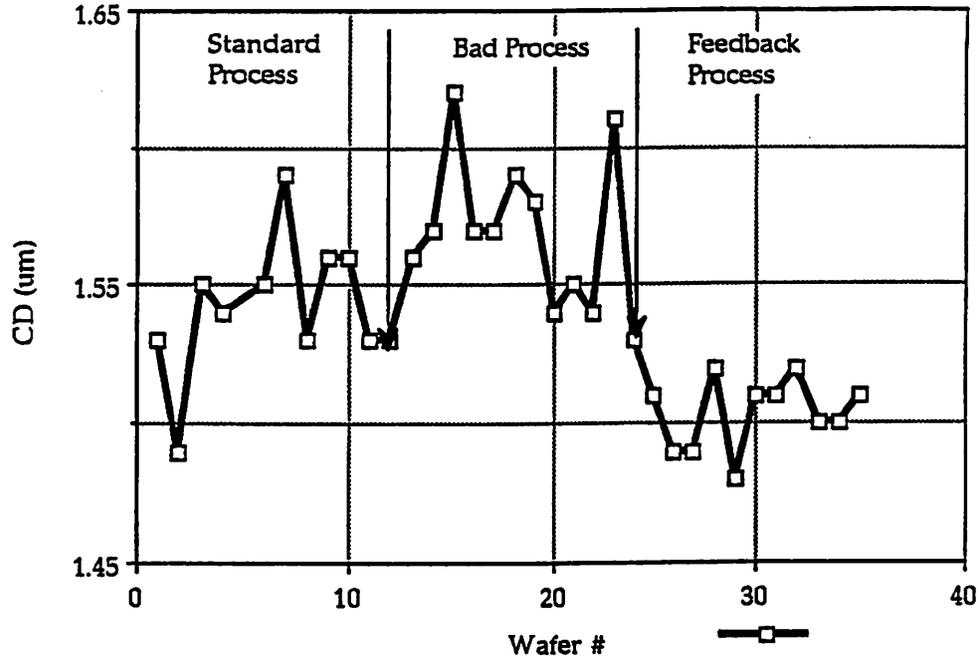


Fig. 7c. Critical Dimension Chart



9.0 References

- [1] Zhi-Min Ling, Sovarong Leang, and Costas J. Spanos, "A Lithography Workcell Monitoring and Modeling Scheme", ME'90, Leuven, Belgium, Sept. 1990.
- [2] Zhi-min Ling, Sovarong Leang, and Costas J. Spanos, "In-Line Supervisory Control in a Photolithographic Workcell", SPIE, Symposium on Microelectronic Processing Integration, Santa Clara, 30th Sept. - 5th Oct. 1990.
- [3] W.G. Oldham, et al, "A General Simulator for VLSI Lithography and Etching Process", IEEE trans. Electronic Devices, vol. ED-27, p717, 1979.
- [4] C.A. Mack, "PROLITH: A Comprehensive Optical Lithography Model", SPIE vol. 538, p.207, 1985.
- [5] Zhi-min Ling and Costas J. Spanos, "A Novel Approach for Film Reflectance Measurement and its Application for the Control of a Photolithography Workcell", ECS, Washington, Sept. 1990.
- [6] B.J. Mandel, "The Regression Control Chart", Journal of Quality Technology, vol. 1, No. 1, pp. 1-9, Jan. 1969.

Modeling a Chemically Amplified Resist with Factorial Experiments

Nelson Tam

A complete model for an electron-beam exposed chemically amplified resist, Shipley's SAL-601-ER7, was obtained by using factorial experiments to study the effects of processing conditions on the dissolution rate of the resist. The model consists of a two-parameter non-linear function relating the dissolution rate to the absorbed electron energy, and of two linear functions relating the two parameters to the post-exposure bake temperature, time, and developer concentration.

1.0 Introduction

The optimization of the resist process traditionally involves finding the best dose for a resist, from its dissolution rate as a function of exposure dose. With the advent of in-situ dissolution rate measurement equipment, a methodology to characterize resist with a dissolution rate model was developed [1]. These models when used in SAMPLE [2], can let the user investigate the trade-offs between exposure dose and development time efficiently and accurately. In the case of electron beam lithography, dissolution rate models were obtained for both positive and negative resists such as PMMA and Hitachi RD2000N [2], [3] by combining development rate data with Monte Carlo simulation of electron energy deposition. However, these models are not general enough to accommodate developer concentration as one of the parameters. As a result, if a different developer concentration is used, a new characterization experiment is needed.

Recently, a new class of highly sensitive resists promises to improve the performance of lithography using the concept of chemical amplification. Upon exposure, acid molecules are generated, which subsequently act as catalysts in a thermal driven reaction to change the resist dissolution rate. Several approaches to model these new resists have been developed. Ferguson et al. used reaction kinetics to describe the exposure and the post-exposure bake of these systems [4]. The disadvantage of this technique is that the chemical reactions have to be characterized with Fourier Transform IR spectroscopy, and the chemical changes of some resists are almost impossible to measure. Furthermore, the model can only allow for optimization in exposure, post-exposure bake temperature, time, and development time.

Liu et al. used factorial experiments to study the effects of post-exposure bake temperature, time, and developer concentration on the resist sensitivity and contrast of electron beam exposed resist [5]. That study however, did not provide dissolution rate models for simulation, and thus had limited use in optimizing the resist processing.

In this paper, a novel approach to model an electron-beam exposed, chemically amplified resist is described. A factorial experiment was performed to study the effects of post-exposure bake temperature, time, and developer concentration on the parameters of a dissolution rate model. Simple models of the parameters were obtained through linear regression. By using these parameter models, the dissolution rates of the resist can be calculated for wide ranges of processing conditions. This way, we obtain a complete model for the simulation and the optimization of the resist processing.

2.0 Experiment

2.1 Resist Preparation

The resist used in this study was Shipley's SAL-601-ER7 negative electron-beam resist. We used 4 inch wafers spin-coated with resist at 4500 rpm for 45 seconds to a thickness of about 0.6 μ m. The wafers were then soft-baked in an oven at 80°C for 30 sec. Subsequently, the wafers were exposed with a pattern containing 12 2mm by 8mm rectangles on a Joel system at H.P. Labs in Palo Alto with 20 keV accelerating voltage and 0.25 A/cm² current density. Each rectangle received a different exposures dose ranging from 0.3 to 3.5 C/cm². After post-exposure bake, the resists were developed in the Perkin Elmer Development Rate Monitor (DRM[®]) as the dissolution rates of the exposed areas were being measured. Average total development time was about 10 min. The developer used was the MF-312 developer from Shipley, diluted with DI water, and the developer temperature was set at 21°C for all the runs. The post-exposure bake conditions and the developer concentration used are discussed in the next section.

2.2 Factorial Design

The experiments were conducted in two stages. The first stage of experiments followed a 2³ factorial design, enhanced with 3 center point replications. This experiment was used to estimate main and interaction effects of the factors and the experimental error. Since the result from the first stage of experiments indicated significant curvature in the parameter models, an augmented design [6] involving 6 "star" runs, in addition to 2 center points was completed. The order of the runs was randomized. The three factors and their levels used in the two stages are listed in Table 1.

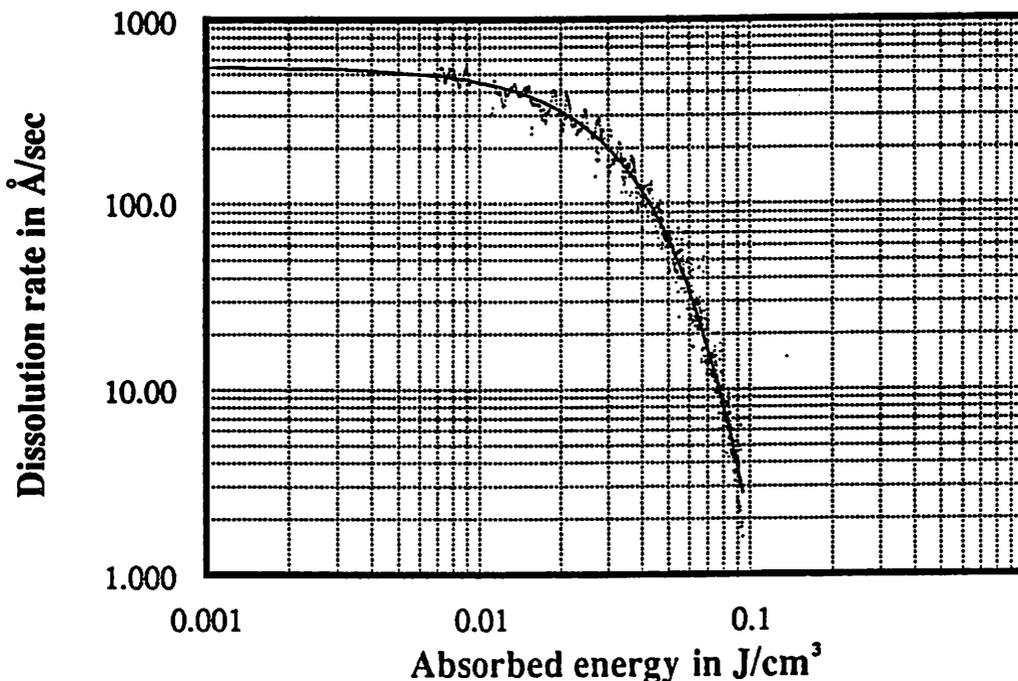
TABLE 1. Factors and levels in the two stages of the designed factorial experiment

Level	Bake Temperature	Bake time	Developer Concentration
-1.4	108°C	60 sec ⁴	0.229N ⁵
-1	110°C	75 sec	0.257N
Center	115°C	90 sec	0.297N ⁶
+1	120°C	105 sec	0.351N
1.4	122°C	150 sec ⁷	0.370N

2.3 Dissolution Rate Data and Model

Before any meaningful presentation of the results can be given, a detailed description of the dissolution rate model is needed. The dissolution rate model is a semi-empirical function fitted to the dissolution rate versus absorbed energy data. These data were generated by combining dissolution rate data from DRM[®] measurement and Monte Carlo simulation of absorbed energy. Since both sets of data are functions of depth in the resist, each rate point can be associated with an absorbed energy by matching their locations in the resist. A typical dissolution rate versus absorbed energy is shown in Figure 1.

FIGURE 1. Dissolution rate versus absorbed energy of SAL-601 from run #5.



- 4. The actual level for this time is -2.
- 5. The actual level of this concentration is -1.6.
- 6. The actual level of this concentration is -0.149.
- 7. The actual level of this time is 4.

These data were then fitted to Equation (1) by minimizing the least square errors of the log of the dissolution rate.

$$R = \frac{R_o}{\left(1 + \left(\frac{E}{E_o}\right)^\beta\right)^\alpha} \quad (\text{EQ 1})$$

Of the four parameters in the above function, only R_o is obtained from physical measurement. For negative resist, R_o is the dissolution rate of unexposed resist and it is directly measured on the DRM[®]. E_o represents a sensitivity indicator. β and α are parameters which determine the contrast of the resist. In negative resists with no chemical amplification, β is usually found to be 1 and α is usually between 2 to 10. Since the dissolution rate data for chemically amplified resist were highly nonlinear, different values β and α were tried in fitting the rate equation.

Although it would be possible to extract all four parameter values for each experimental run, this would lead to an undetermined problem, as there will be multiple solutions for the same least square residual. In order to have consistency in extracting the parameters for the dissolution rate functions from all the runs, we set β to 1.5 and α to 31. Once β and α were set, R_o and E_o were extracted by minimizing the ratio of the predicted to the experimental sum of squares for the residuals and their dependence on the three processing parameters were studied.

3.0 Results

3.1 First Experimental Stage and Linear Effects

The experimental average and the effects of R_o and E_o were calculated from the 2^3 factorial experiment using Yate's algorithm. These values are listed in Table 2.

TABLE 2. Extracted R_o , E_o and their effects.

Run	T	t	C	$R_o(\text{\AA}/\text{sec})$	R_o Effect	$E_o(\text{J}/\text{cm}^3)$	E_o Effect	Effect name
.1	-	-	-	77	294.8	253	218.6	AVG
2	+	-	-	81	-30.4⁸	174	-76.3	T
3	-	+	-	86	-25.8	224	-21.3	t
4	+	+	-	61	-16.5	166	6.3	Tt
5	-	-	+	552	437.5	288	28.8	C
6	+	-	+	521	-19.7	202	-7.8	TC
7	-	+	+	525	-20.2	262	-2.8	tC
8	+	+	+	456	-2.5	180	-4.3	TtC
Standard Error =					± 6.8		± 4.1	

The standard errors for the effects were calculated from the estimate of the experimental error obtained from the 5 center runs⁹. If effects within $\pm 2\sigma$ were considered significant, then for E_o , only the main effects were significant. However, for R_o , only the three factor interaction effects can be considered insignificant. In order to obtain a simpler model for R_o , several transformations were tried. Eventually, the square root transformation was found to lead to a simpler model. This square root transformation was adopted and the effects on $\sqrt{R_o}$ are listed in Table 3. After the transformation, the significant effects were found to be the bake temperature, bake time, developer concentration, and temperature-time interaction.

TABLE 3. Effects of square root of R_o

Run	T	t	C	$\sqrt{R_o}(\text{\AA}/\text{sec})^{1/2}$	$\sqrt{R_o}$ Effect	Effect name
1	-	-	-	8.79	15.67	AVG
2	+	-	-	8.97	-0.87	T
3	-	+	-	9.25	-0.69	t
4	+	+	-	7.80	-0.63	Tt
5	-	-	+	23.49	13.95	C
6	+	-	+	22.83	-0.24	TC
7	-	-	+	22.91	-0.33	tC
8	+	+	+	21.35	0.19	TtC
Standard Error =					± 0.26	

8. Significant effects are highlighted in bold face.

9. The data for the center runs are presented in Table 4.

3.2 Second Experimental Stage and Quadratic Effects

By comparing the average values of both parameters to that of the center runs, it was found that significant curvature existed. The second stage of the experiments was designed to provide good estimates of higher order terms for the parameter models. The results from the second stage of the experiments and all the center runs are listed in Table 4.

TABLE 4. Results from the second stage of the experiment and all the center runs.

Stage	Run	T	t	C	$\sqrt{R_o}(\text{\AA}/\text{sec})^{1/2}$	Avg.	$E_o(\text{J}/\text{cm}^3)$	Avg.
1	9	c	c	c	14.87		206	
1	10	c	c	c	14.97	14.74	196	199
1	11	c	c	c	14.39		195	
2	7	c	c	c	17.44		185	
2	8	c	c	c	17.00	17.22	178	181.5
2	1	-1.4	c	c	17.80		241	
2	2	1.4	c	c	16.43		157	
2	3	c	-2	c	18.52		202	
2	4	c	4	c	16.49		161	
2	5	c	c	-1.6	6.09		177	
2	6	c	c	1.4	28.09		209	

Despite all efforts to keep the other non-significant processing parameters under control, there was substantial discrepancy in $\sqrt{R_o}$ and E_o between the two stages. A linear model for $\sqrt{R_o}$ derived from the first stage of experiments was used to predict the result from the second stage of experiments. The plots of residuals for that model are shown in Figure 2 and a constant shift was observed. Though not as obvious as the case for $\sqrt{R_o}$, the plots of residuals for E_o (Figure 3) also displayed similar shift, but in the opposite direction. In addition, residuals from the two stages of experiments had similar curvilinear relationships with the predicted E_o , further illustrating the inadequacy of the linear model.

FIGURE 2. Plots of residuals for linear R_o model.

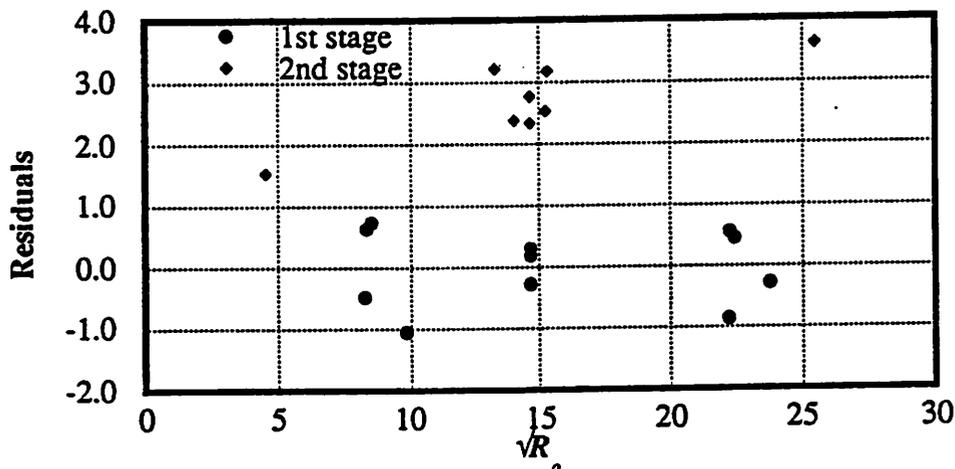
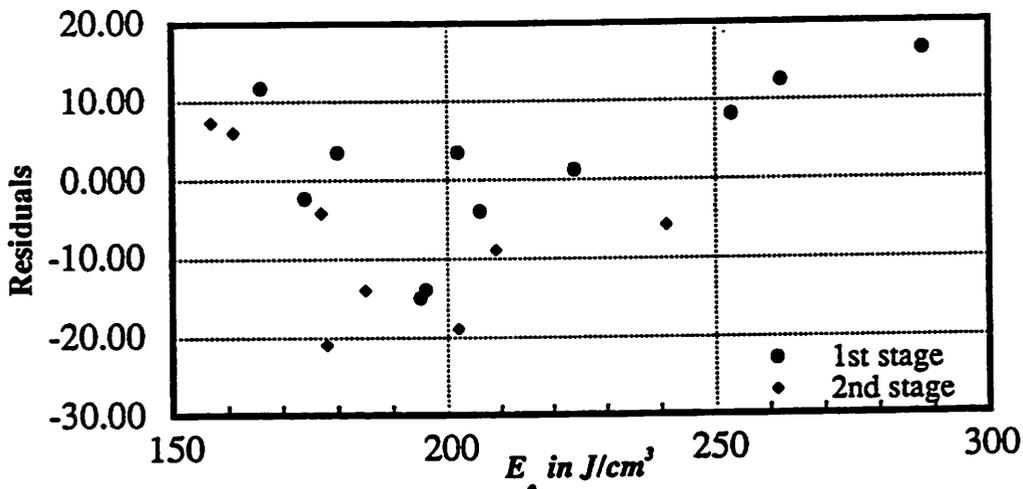


FIGURE 3. Plots of residuals for linear E_o model.



3.3 Effect between First and Second Experimental Stage

There were only a few possible sources of variations in the processing that could have contributed to this blocking effect. The most likely ones were aging of the resist and/or the developer because of the uniformity of the shift in the observed R_o 's. On the other hand, since all but two of the runs in the second stage had the same bake temperature, drift in the oven temperature could also be responsible for this discrepancy. More experiments are needed to identify this blocking effect.

Fortunately, the overall experimental design was such, that despite the blocking effect, significant quadratic models of the process can be derived. These are described next.

4.0 Quadratic Models for R_o and E_o

4.1 Models

Based on the curvature check and the analysis of residual, it was concluded that quadratic models were required to describe $\sqrt{R_o}$ and E_o accurately. Moreover, an extra linear term was added to each model to account for the blocking effect. The coefficients for the two models were then determined using least squares technique with all the runs included. The models for $\sqrt{R_o}$ and E_o are as follows:

$$\sqrt{R_o} = 15.9126 - 0.4537T - 0.3069t + 7.0858C - 0.2794C^2 - 0.3163Tt + 2.5197B \quad (\text{EQ } 2)$$

$$E_o = 207.2 - 35.4T + 9.0T^2 - 7.7t + 13.3C - 19.2B \quad (\text{EQ } 3)$$

In the above equations, T , t , and C are normalized processing parameters with respect to the levels used in the factorial experiments. B is the blocking effect parameter which takes the values of either 0 or 1 corresponding to 1st and 2nd stage.

4.2 Analysis of the Residuals

Before the goodness of the fit can be determined, it is necessary to inspect the residuals for the possible indications of model inadequacy. The residuals for the quadratic models of $\sqrt{R_o}$ and E_o were calculated and are plotted in Figure 4 and 5. For $\sqrt{R_o}$, the plots of residuals did not show any trend and the residuals appeared to be randomly distributed. On the other hand, the plot of residuals versus the predicted values for E_o still shows a slight curvilinear trend. Moreover, the range of the residuals, on the order of a few percent of the predicted E_o , was quite large.

FIGURE 4. Plots of residuals for $\sqrt{R_o}$ quadratic model

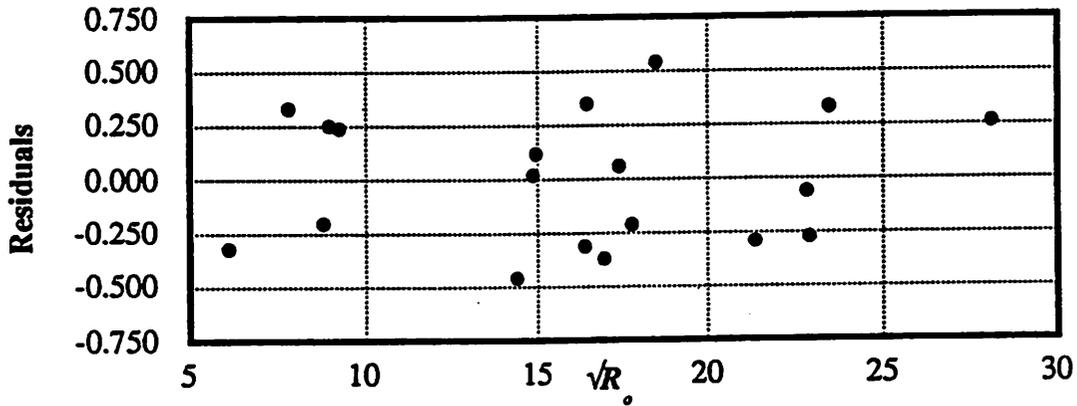
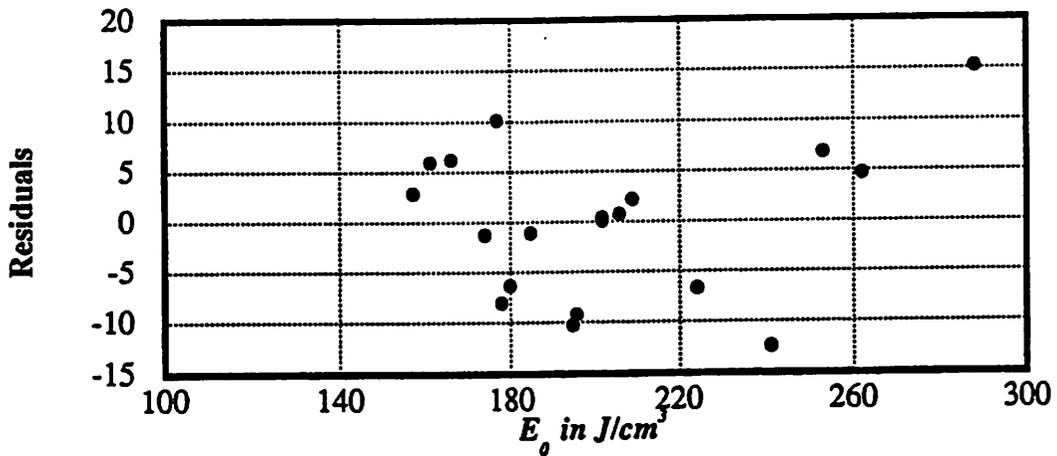
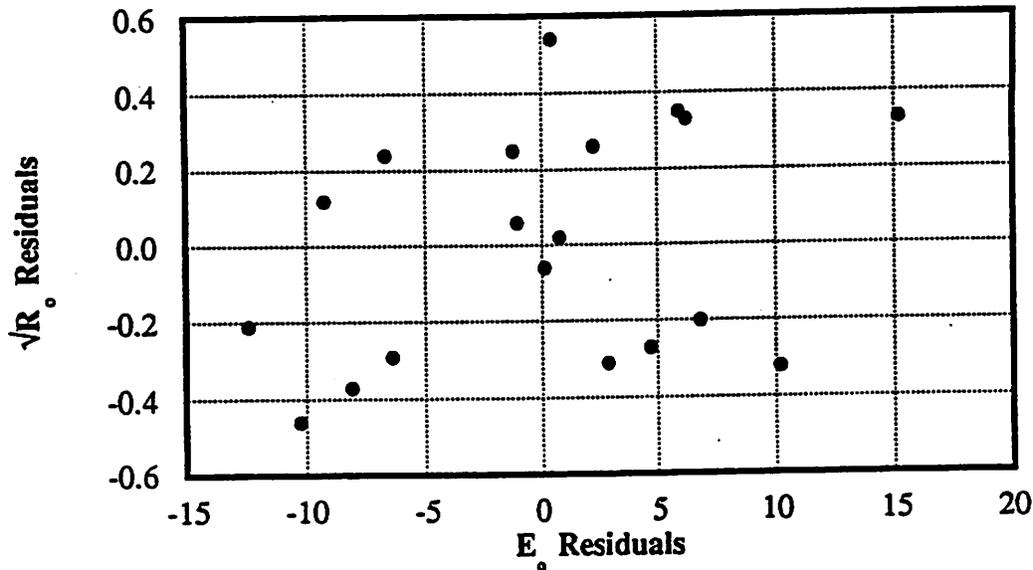


FIGURE 5. Plots of residuals for E_o quadratic model



From the plots of residuals, we can see that the model for $\sqrt{R_o}$ is more accurate than the model for E_o . This difference in the accuracy of the models perhaps can be explained by the fact that R_o is a physical parameter (dissolution rate of unexposed resist), whereas E_o is a parameter extracted with nonlinear regression. In the parameter extraction procedure for the dissolution rate model, the values of α and β had no influence on the determination of R_o . On the other hand, the best extracted value for E_o depended strongly on both α and β . The unknown blocking effect could have also interfered with the assumption that α and β were constant. Nonetheless, the resulting model still matched most of the runs, and only in few of the cases was the discrepancy more than a few percent. The final test for the models was to inspect the correlations between the two residuals. Figure 6 shows no evidence of such correlations.

FIGURE 6. Plots of residual correlation.



4.3 Analysis of Variance

Since the two models did not show inadequacy, their goodness of fit can be determined by the analysis of variance. In Table 5 and 6, the sums of squares for the observed, estimated and residuals of $\sqrt{R_o}$ and E_o are shown. The sums of squares for the residuals are further broken down into a lack of fit part and a pure error part. The ratio of the lack of fit to pure error indicated whether the sum of squares for the residuals are caused by the lack of fit or pure error. In both cases, these ratios were very small suggesting there was no reason to suspect lack of fit. To formally determined the significance level, the F-distribution was used. Ratio for $\sqrt{R_o}$ as great or greater than 0.23 can be expected about 96% of the time. The larger experimental errors in the E_o data push confidence level even higher (99% of the time).

TABLE 5. Analysis of variance for $\sqrt{R_o}$ data

source	sum of squares	degrees of freedom	mean square	
model	$S_M = 5629.4$	7		
lack of fit	$S_L = 0.36$	7	0.0514	
residual	$S_R = 1.5$	12	0.125	ratio = 0.23
pure error	$S_E = 1.14$	5	0.228	
total	$S_T = 5630.9$	19		

TABLE 6. Analysis of variance for E_o data

source	sum of squares	degrees of freedom	mean square	
model	$S_M = 805070$	6		
lack of fit	$S_L = 194$	8	24.3	
residual	$S_R = 990$	13	76.2	ratio = 0.15
pure error	$S_E = 796$	5	159	
total	$S_T = 806060$	19		

5.0 Estimation of Confidence Interval for the Main Model

5.1 Variance of fitted $\sqrt{R_o}$ and E_o

Since direct calculation of the variance of the fitted parameters was very complicated, an average variance was computed instead using Equation (4) [7]

$$\bar{V}(\hat{y}) = \frac{1}{n} \sum_{i=1}^n V(\hat{y}_i) = \frac{p\sigma^2}{n} \quad (\text{EQ 4})$$

On the assumption that the model was adequate, an estimate of the error variance σ^2 for $\sqrt{R_o}$ were

$$s^2 = \frac{S_R}{n-p} = \frac{1.5}{19-7} = 0.125 \quad (\text{EQ 5})$$

and for E_o

$$s^2 = \frac{S_R}{n-p} = \frac{990.2}{19-6} = 76.2 \quad (\text{EQ 6})$$

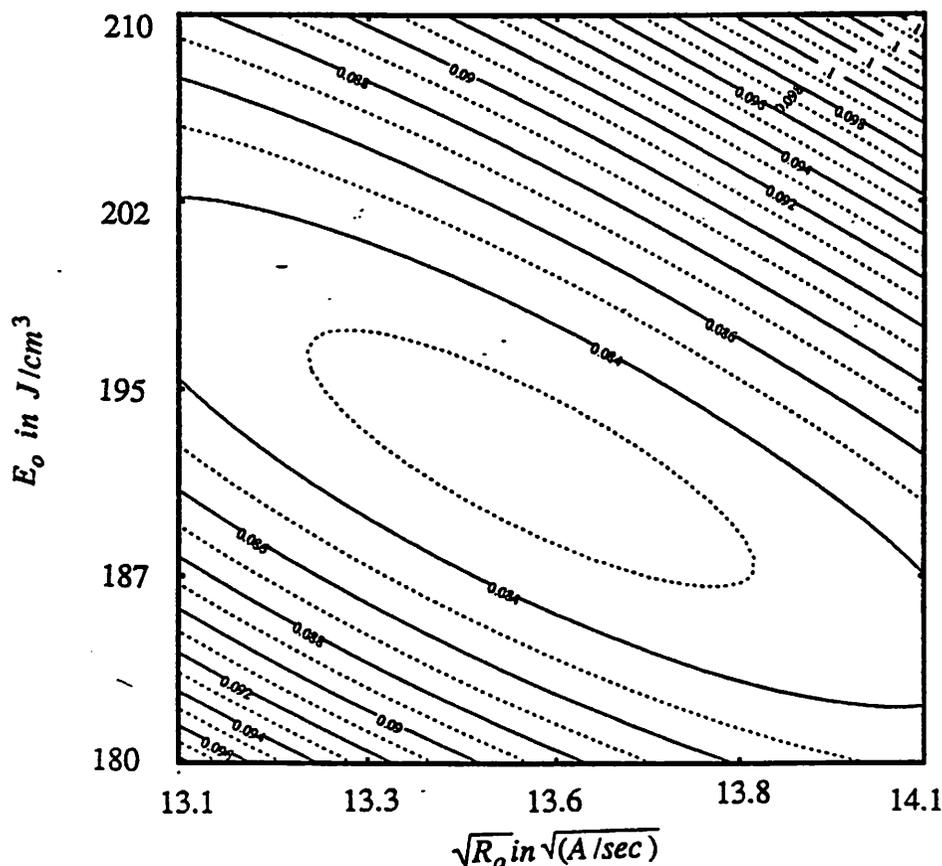
Substituting the result from Equation (5) and (6) into Equation (4), we found $\bar{V}(\sqrt{R_o}) = 0.046$ and $\bar{V}(E_o) = 24.1$. The confidence limits for the two estimates can be easily calculated from the student t -distribution with their respective degrees of freedom. For example, an extra wafer (wafer #20) was included in the 2nd batch of wafer with post-exposure bake at 115°C for 60 sec and developer concentration of 0.27N. The observed and estimated values from the quadratic models are compared in Table 7 along with the 95% confidence limits.

TABLE 7. Comparison of observed and estimated $\sqrt{R_o}$ and E_o for wafer #20

Parameter	Observed	Quadratic model	95% Confidence limits
$\sqrt{R_o}$	13.67	13.77	0.47
E_o	185	194	10.6

The results in Table 7 indicate there was good agreement between the estimated and observed values. However, the credibility of a hypothesized pair of values for $(\sqrt{R_o}, E_o)$, with the joint confidence regions formed by the two limits, is questionable. In Figure 7 we show the contours of the sum of squares surface calculated from the ratios of estimated to experimental dissolution rate values for wafer #20. The values of the surface were indicators of the goodness of fit of the rate model, and the smaller the value, the better is the fit. The point with coordinates [13.77, 194] lies well within the region of the minimum value. However, consider another point with coordinates [13.45, 188]; although it is within the individual limits of the joint confidence region, it has a larger sum of squares ratio. As a result, any estimated values should be checked by referring to the contours of sum of squares.

FIGURE 7. Contours of ratio of sum of squares surface, wafer #20.



6.0 Conclusion

A complete model for Shipley's SAL-601-ER7 resist has been obtained by using a 2^3 factorial experiment, later augmented with 6 "star" runs and 5 center runs in a Box-Wilson fashion. The results of these experiments were used to study the effects of post-exposure processing on the dissolution rate of the resist. The model is a nonlinear dissolution rate function characterized by two parameters $\sqrt{R_o}$ and E_o . These two parameters are in turn related to the post-exposure bake temperature, bake time, and developer concentration through two linear functions, which were determined by linear regression. The experiments were conducted in two blocks. An unknown blocking effect was observed, which had to be taken into account, by incorporating an extra blocking term in the linear models for the dissolution rate parameters.

The overall dissolution model compares favorably with the experimental result within the range of the processing parameter studied. These ranges are sufficiently large to enable optimization of the resist processing, although future work should include experiments in order to determine the unknown blocking effect, and to further refine the model. Finally, the robustness of the model should be tested by using it in the SAMPLE program. This will be accomplished by simulating the thickness of the remaining resist versus the exposure dose, and by comparing them to additional experimental readings.

7.0 Acknowledgment

The author wishes to thank Hua-Yu Liu at the H.P. Palo Alto Research Laboratory for the preparation of the samples used in this project.

8.0 References

- [1] W. R. Bell, P. D. Flanner, C. Zee, N. Tam and A. R. Neureuther, "Determination of Quantitative Resist Models from Experiment", *Proceedings of SPIE, Advances in Resist Technology and Processing V*, Vol. 920, 1988.
- [2] A. R. Neureuther, D. F. Kyser, and C. H. Ting, "Electron Beam Resist Edge Profile Simulation", *IEEE Trans. on Electron Devices*, ED-26, No.4, pp. 686-692, April 1979.
- [3] N. N. Tam, R. Coyne, and A. R. Neureuther, "Characterization of Electron-Beam Exposed Optical Resist", *J. Vac. Sci. and Technol.*, B, pp.361-365, 1988.
- [4] R. A. Ferguson, Y. Shacham-Diamand, C. A. Spence, and A. R. Neureuther, "Modeling and Simulating of Multiple Chemical States in Photoresist Materials", *Proceedings of SPIE, Advances in Resist Technology and Processing VI*, Vol. 1086, pp. 262-273, 1989.
- [5] A L. Blum, M.E. Perkins and H.-Y. Liu, "A Study of the Effect of Key Processing Variables on the Lithographic Performance of Microposit SAL601-ER7 Resist", *J. Vac. Sci. Technol.*, B6, pp. 2280, 1988.
- [6] G. P. E. Box, W. G. Hunter, and J. S. Hunter, *Statistics for Experimenters*, John Wiley & Sons, pp. 524, New York, 1978.
- [7] G. P. E. Box, W. G. Hunter, and J. S. Hunter, *Statistics for Experimenters*, John Wiley & Sons, pp. 513-523, New York, 1978.