SPECIAL ISSUES IN SEMICONDUCTOR MANUFACTURING

IV

EECS 290W Class Projects Reports, Spring 1993

Professor:

Costas J. Spanos

Students:

John Hutchinson, Pamela Tsai, Hsing-Jen Wann, and Marco Zuniga

SPECIAL ISSUES IN SEMICONDUCTOR MANUFACTURING

IV

EECS 290W Class Projects Reports, Spring 1993

Professor:

Costas J. Spanos

Students:

John Hutchinson, Pamela Tsai, Hsing-Jen Wann,
and Marco Zuniga

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

.

Preface

This is the fourth annual edition of the 290W report. This edition includes descriptions of projects completed during the Spring semester of 1993, in the context of the graduate course "Special Issues in Semiconductor Manufacturing". Four 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.

Each of the presented projects covers at least one novel aspect of semiconductor manufacturing. The first project discusses the application of an automated probe station for implementing SPC on a CMOS baseline process. creation of a novel multivariate exponentially weighted moving average scheme suitable both for closed loop control as well as SPC. The second project deals with computer-based experiments. More specifically, a process simulator was connected to a statistical analysis package in order to create easy to use response surface models of process behavior. The third project deals with the application of SPC on controlling a process that has both batch as well as sequential characteristics. The technique described here employs novel data grouping and can be applied to many similarly configured microfabrication processes. The last project deal with the application of statistical techniques in modeling the behavior of plasma etching of photoresist. This project is an example on how physical understanding can be coupled with empirical observations to create a compact model of a complex process.

It is my hope that these reports will add to our understanding of semiconductor manufacturing. My thanks go to the 290W students and auditors whose work made this document possible. I am also grateful to the personnel and management of the Berkeley Microfabrication laboratory for their help with the experimental part of the projects presented here.

Costas J. Spanos

January, 1993

Table of Contents

SPC of a Baseline CMOS Process Using an Automated Probe Station

Hsing-jen Wann

In this work we present an automated software/hardware configuration for performing Statistical process control on electrical process data. This system has been deployed for controlling the CMOS baseline process in the Berkeley Microfabrication Laboratory.

1.0 Introduction

An automatic testing and analyzing system is important for statistical process control in the computer-integrated manufacturing environment of the semiconductor industry. This report is to demonstrate the feasibility of a system developed in the device characterization laboratory (407 Cory) for monitoring the baseline process of the microfabrication laboratory.

To use the system for statistical process control, the work is divided into three levels. They are: the general-purpose hardware interface, the parameter-specific measurement routine, and the user (e.g., the quality control engineer). The organization of this report will follow these three levels to provide a comprehensive view. Examples are provided for future users.

2.0 System Implementation

2.1 Hardware Interface

The basic requirement of the system are the automatic probe station, the switching matrix, a control workstation, and the characterization instrument. Since in this work only DC characterization is used, the instrument is HP 4145B. The interface program/driver is "sunbase.c," which is currently in /export/mnt/argon3/lab/staff/gutnik/dcl/code/eric on argon (lead). The measurement starts by initiating "sunbase" at the command line. More information can be found in "docs" (in the same directory). This program also supports the parameter-specific measurement routine which will be explained in section 2.2.

The user has to specify where and how many devices, as well as what device parameters (Vt, Leff, ...), are to be measured. The information is stored in "prober.text." This program is shown in listing 1. In "probe.text," the 9x9 matrix is the wafer map. The 9x9 matrix leaves more flexibility for wafer stepper exposure programs. As can be seen in listing 1, we only use the left-top 6x6 out of the 9x9 matrix in "prober.text" for making measurement. These dies correspond to the center portion of the 8x8 baseline process. This is simply done by physically choosing appropriate origin when starting alignment. The location marked with "x" is the first die to start the measurement. The probe card should be manually aligned to this die in the beginning. The probe card will then go through all the die locations labeled with "1." In the following paragraphs of prober.text, sunbase was told to perform Vt measurement on transistor M1, another Vt measurement on transistor M2, and Leff measurement based on transistor M1 and M2. More detailed information such as translation distance is stored in "die.map." The measurement routine, vt.c and leff.c, are shown in listing 2 and 3. As can be seen, the measurement routine does not communicate with the instru-

---

ment directly, but instead uses the command supported by sunbase, such as "HOLD," "SWEEP," so that the focus of the measurement routine is on data (from the instrument) manipulations. This makes it easier for the user to write his/her own routine if necessary.

2.2 Measurement Routine

In general the measurement routine is not provided by the users but by device engineers/technicians. We will briefly go through the measurement routines in listing 2 and 3.

2.2.1 Vt routine

There are several definitions of the MOSFET threshold voltage for different characterization purpose. In this work we will use the conventional one, that is, to define Vt by extrapolating the linear drain current down to zero. Due to the mobility degradation, the slope of the linear drain current is not a constant for the whole range of V_G, thus the linear extrapolation is not well defined. We will define the slope at maximum g_m so that the extrapolated Vt is closest to the onset of strong inversion [1]. The threshold voltage is defined as:

$$V_T = V_G - I_D / g_m, \qquad (1)$$

where V_G, I_D and g_m are evaluated at the maximum g_m.

2.2.2 Leff routine

The methods to measure ΔL are even more diverse than that for Vt measurements. Examples like Paul & Sushi [2], pair Vg [3], etc., each has its own advantage. The Paul & Sushi method can exclude the effect of series resistances. The pair Vg method is good for handling LDD MOSFETs. However the simplest scheme to extract ΔL by measuring the channel resistance in the linear region of one long channel device and one short channel device still has advantages [4]. The ΔL is extracted by finding the interception of the two straight lines passing through the channel resistances for two channel lengths at two different Vg. That is, ΔL is introduced as the intercept of the two straight lines below:

$$Rm1 = \frac{L - \Delta L}{\mu CoxW(Vg1 - Vt)} \qquad (2)$$

$$Rm2 = \frac{L - \Delta L}{\mu CoxW(Vg2 - Vt)} \qquad (3)$$

We can choose large Vg to reduce the effect of the variation in Vt. The two channel lengths chosen in this report are 25μm and 1.5μm. We have to use the short channel device by noting that the extrapolation error will be too large if both devices have long channel lengths (say, 25μm and 5μm).

3.0 Statistical Process Control of Baseline CMOS Process

3.1 Control Chart for Vt

As can be seen in listing 1 there are a total of six data points per wafer, two points measured at the center and four points measured at the circumference. With the highly automated probe station available, measuring more data points per wafer seems justified since the bottleneck in the measuring process is in aligning the wafer. In the baseline CMOS process the threshold voltage is fluctuating due to frequent adjustment of the threshold implant dosage. The measured data is shown in Fig.1. Since the process is evolutionary in nature, an appropriate control chart is evolutionary control chart shown in Fig.2. Some of the occasional out-of-control data points have been identified as the misalignment of the automatic probe station due to accumulation of the error with increasing number of steps. This error can be reduced by starting the measurement at the location as close to the center as possible.

3.2 Control Chart for Leff

The $\bar{X}$-R control chart for variable is used for Leff. The R chart is used to estimate the experimental error. In some occasions the measured parameters are functions of the location in a wafer. A "gradient test" is used to check if there is significant geometrical parameter drift. First we take the average of the sum of the squares of the differences of the parameters located at the top/bottom and at the left/right pairs:

$$G = ((\text{Top} - \text{Bottom})^2 / 8 + (\text{Left} - \text{Right})^2 / 8)^{0.5} \tag{4}$$

If the value G is significant compared with the experimental error (which is estimated by the R chart), the gradient is significant and we have to exclude the effect of the parameter gradient in estimating the experimental error.

The measured data points are shown in Fig. 3. The devices on a wafer are grouped together, the resulting R control chart is shown in Fig.4. Apparently the last data point (from CMOS 22-6) was out-of-control. This data point was excluded to redraw the control chart shown in Fig.5. $\bar{R}$ is 0.216μm. The estimated experimental error is about 0.08 μm. Compared this experimental error with the gradient test, we conclude that the parameter gradient is not significant.

The $\bar{X}$ chart is shown in Fig. 6 using $\bar{R}$ to draw the UCL (-0.767μm) and LCL(-0.559μm). It is seen that the measured ΔL is not under statistical process control. Moreover, the measured ΔL are all negative, while in a single drain device the ΔL is expected to be positive due to source/drain lateral diffusion. The measured ΔL is a result of several process steps, the photolithography, the anisotropic etch, the implantation, and the drive in. Since the device characteristics are quite typical, it is not likely the implantation and the drive-in are causing the problems. The negative ΔL is more likely due to the etching and the lithography. One possibility is the gate etching is not vertical. The other possibility was due to inadequate exposure in the lithographic step. More experiments are need to find the assignable cause.

4.0 Conclusions and Future Work

A complete flow of the statistical process control of baseline CMOS process using the automatic probe station is demonstrated. The measurement time is minimized. The measured data are already on-line. A shell script can read the measurement output file and create the data file with

the data format that can be read by statistical software such as JMP. We have shown two examples of measurement routines and control chart designs.

The PMOSFET are not studied yet. The threshold voltage control of the NMOSFET of the baseline CMOS process is reasonably good. Though up to CMOS 22 the process capability has not been satisfactory if Vt(NMOSFET) = 0.7 is the target, this can be tuned by adjusting the implant dose. The control of the channel length in both magnitude and in variation leaves some room to be desired. Further works should be done to find out the assignable causes.

Acknowledgment: Eric Boskin and Dave Rodriguez helped to complete this project.

[1]   K. Hui, private communications.

[2]   MOS devices, physics, reliability, and characterization, an intensive two-day course, April 17-18 1990.

[3]   "Gate-voltage-dependent effective channel length and series resistance of LDD MOSFETs," IEEE Trans. on Electron Devices, vol. ED-34, No. 12, Dec 1987.

[4]   M. Rodder, private communications.
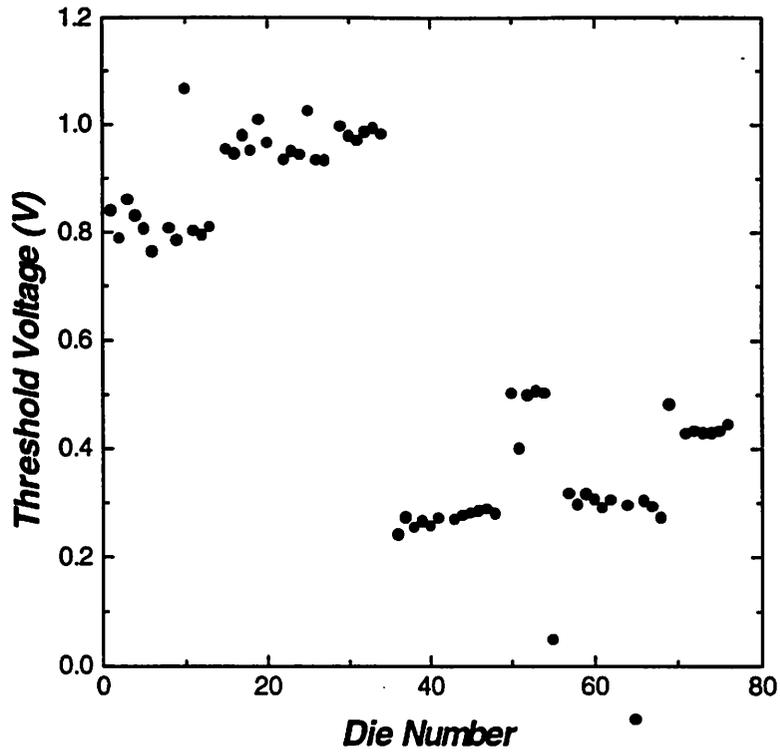
Fig.1 Raw data of Vt measurement from CMOS 19, 20, and 22.
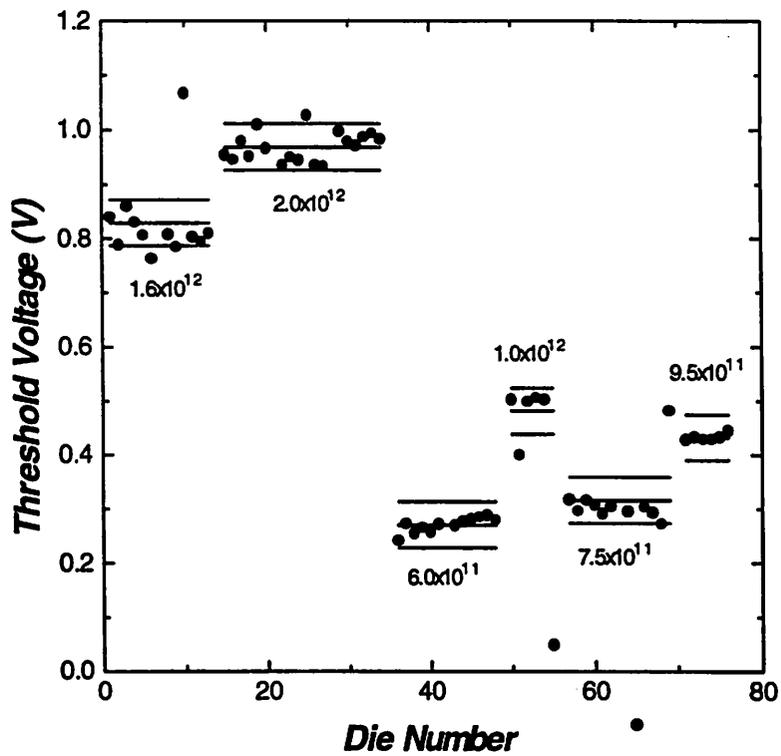
Fig.2 The evolutionary control chart of for Vt control.
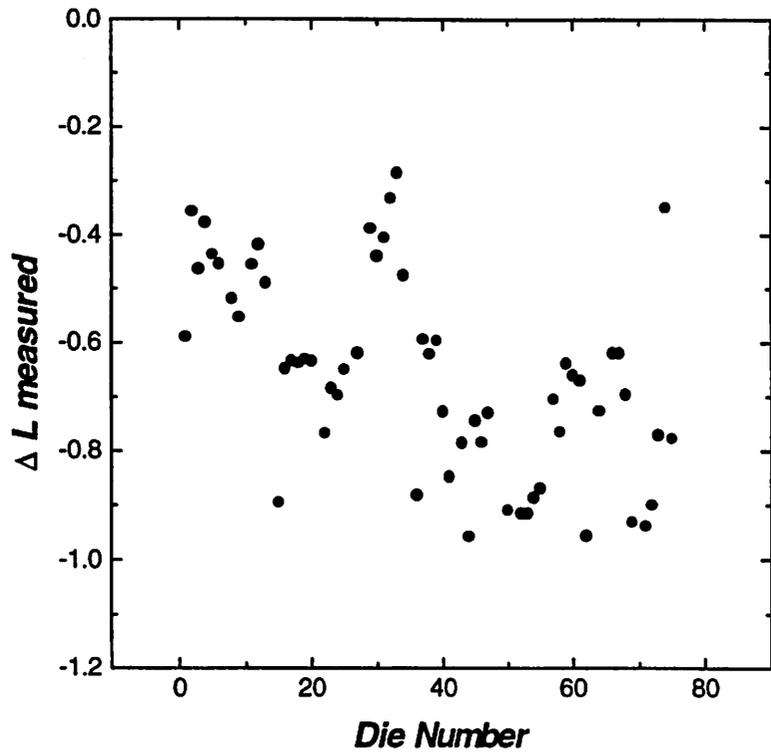
Fig.3 Raw data of delta L measurement from CMOS 19, 20, and 22.
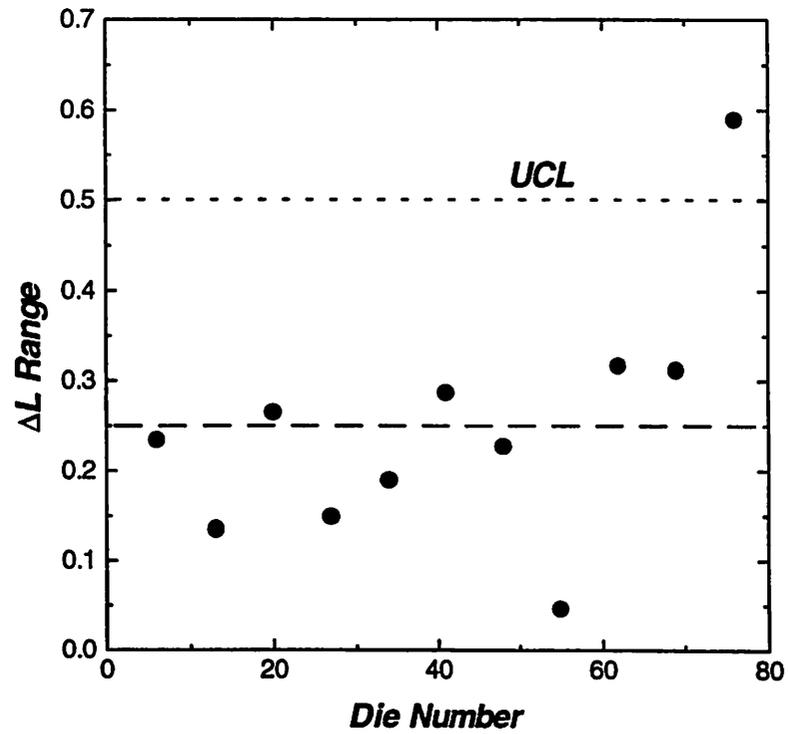
Fig.4 The R chart of delta L measurement with devices on a wafer as a group.
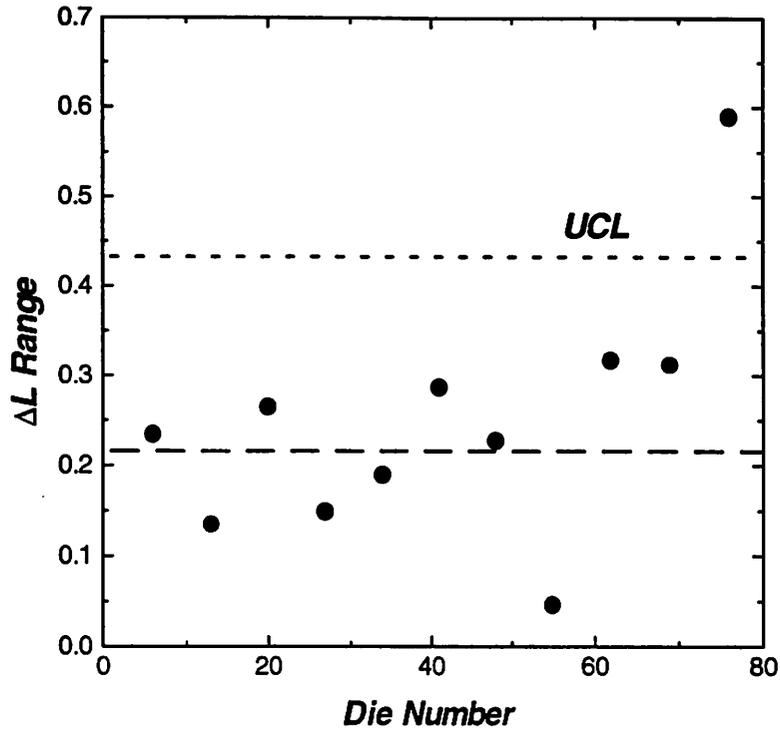
Fig.5 The estimated experimental error is 0.216um.
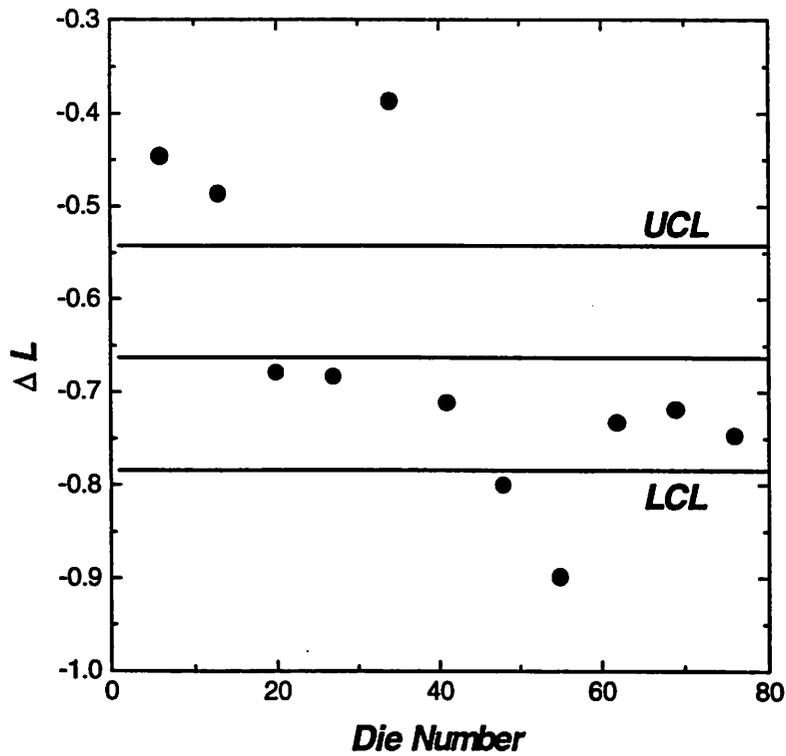
Fig.6 The X chart of delta L measurement. The process is not in control

```c
#include "sunbase.h"

#define totaldev 2
#define points   4
#define vgstart  1.6
#define vgstep   0.6
#define vd       0.05
#define delvt    0.125

static double vgleff[points],lmask[totaldev];
static double rm[points][totaldev],rmmeas[points][totaldev];

void leff_init()
{
    int i;

    lmask[0]=-25.0;
    lmask[1]=-1.5;
    for (i=0;i<=points-1;i++)
        vgleff[i]=vgstart+i*vgstep;
}

double strtoarr(char *spastring, int dev) {
    int i;
    D(printf ("< strtoarr\n");)
    D(printf ("%s\n",spastring);)
    for (i=0;i<=4;i++) {
        spastring[13 * i]='-';
        spastring[13 * i + 12]='\0';
        rm[i][dev]=vd/atof(spastring + 13 * i);   /*Array data processing*/
    }                                             /*according to the output*/
}                                                 /*data format of HP4145*/

void L_EFFECTIVE (char **paramlist, FILE *output) {
    char *result;
    int i,j;
    double sumx,sumy,sumxy,sumxsq,sumysq,denom,sumdeltal,sumrext,deltal,rext;
    double intcpts[points],rho[points],slopes[points];
    D(printf ("< L_EFFECTIVE\n");)
    leff_init();
    while (*++paramlist) {
        if (**paramlist == '+') {
            printf ("option %s\n",*paramlist);
        }
        else {
/* probe first device */
            moveto (*paramlist++);
            hold (BULK,'V',0,.01);
            hold (SOURCE,'V',0,.01);
            hold (DRAIN,'V',0.05,0.01);
            sweep (GATE,'V',vgstart,vgstart+(points-1)*vgstep,vgstep,.01);
            graph (GATE,'V','X',0,5);
            graph (DRAIN,'I','Y',0,0.00005);
            trigger(SINGLE);
            result = dump (DRAIN, 'I');
            strtoarr(result,0);

/* delvt compensates for change in Vt */
/* probe second device */
            moveto (*paramlist);
            hold (BULK,'V',0,.01);
            hold (SOURCE,'V',0,.01);
            hold (DRAIN,'V',0.05,0.01);
            sweep (GATE,'V',vgstart+delvt,vgstart+delvt+(points-1)*vgstep,vgstep,.01);
            graph (DRAIN,'V','X',0,5);
            graph (DRAIN,'I','Y',0,0.0001);
            trigger(SINGLE);
            result = dump (DRAIN, 'I');
            strtoarr(result,1);

/* calc Leff */
            for (i=0;i<=points-1;i++) {
                sumx=sumy=sumxy=sumxsq=sumysq=0.0;
                for (j=0;j<=totaldev-1;j++) {
                    sumx+=lmask[j];
                    sumy+=rm[i][j];
                    sumxy+=lmask[j]*rm[i][j];
                    sumxsq+=lmask[j]*lmask[j];
                    sumysq+=rm[i][j]*rm[i][j];
                }
                denom=totaldev*sumxsq-sumx*sumx;
                intcpts[i]=(sumy*sumxsq-sumx*sumxy)/denom;
                slopes[i]=(totaldev*sumxy-sumx*sumy)/denom;  /* totaldev is int! */
                rho[i]=(sumxy-sumx*sumy/totaldev)/(sqrt((sumxsq-sumx*sumx/totaldev)
                       *(sumysq-sumy*sumy/totaldev)));
            }

            sumdeltal=0.0;
            sumrext=0.0;
            for (i=0;i<=points-2;i++) {
                for (j=i+1;j<=points-1;j++) {
                    sumdeltal=sumdeltal+(intcpts[j]-intcpts[i])/(slopes[i]-slopes[j]);
                    sumrext=sumrext+(intcpts[i]*slopes[j]-intcpts[j]*slopes[i])/
                            (slopes[j]-slopes[i]);
                }
            }

            deltal=sumdeltal/(points*(points-1)/2);
            rext=sumrext/(points*(points-1)/2);
            printf ("DeltaL = %f and Rext = %f\n",deltal,rext);
            sprintf (result,"DeltaL = %f and Rext = %f\n",deltal,rext);

            fprintf (output,"\n*%s\n",*paramlist);
            fprintf (output, result, strlen (result));
        }
    }
    D(printf ("> L_EFFECTIVE\n");)
}
```

```c
#include "sunbase.h"

#define points   101
#define vgstart  0.0
#define vgstep   0.05
#define vd       0.05

static double vg[points];
static double id[points];
static double gm[points];

double strtoarr2(char *spastring, int dev) {
    int i;
    D(printf ("< strtoarr\n");)
    D(printf ("%s\n",spastring);)
    for (i=0;i<points;i++) {                 /*Processing output data*/
        spastring[13 * i]='-';               /*from HP4145. For more*/
        spastring[13 * i + 12]='\0';         /*details see manual*/
        id[i]=atof(spastring + 13 * i);
    }
}

void VT_init()
{
    int i;

    for (i=0;i<points-1;i++)
        vg[i]=vgstart+i*vgstep;
}

void VT (char **paramlist, FILE *output)
{
    char *result;
    int i,j;
    double Vt,gmmax;
    D(printf ("< Vt_MEASUREMENT\n");)
    VT_init();
    while (*++paramlist)
    {
        if (**paramlist == '+')
        {
            printf ("option %s\n",*paramlist);
        }
        else
        {
            /* probe first device */
            moveto (*paramlist);
            hold (BULK,'V',0,.01);           /*Substrate and*/
            hold (SOURCE,'V',0,.01);         /*and Source grounded, compilance 0.01*/
            hold (DRAIN,'V',0.05,0.01);      /*Small drain voltage 0.05*/
            sweep (GATE,'V',0,5,.05,.001);   /*Sweep the gate voltage*/
            graph (GATE,'V','X',0,5);        /*Plotting*/
            graph (DRAIN,'I','Y',0,0.00005);
            trigger(SINGLE);                 /*Start measurement*/
            result = dump (DRAIN, 'I');      /*Get data*/
            strtoarr2(result,0);

            /* calc Vt */
            for (i=0; i<=points-1 ; i++)
/*              printf("vg= %f,id= %e\n",vg[i],id[i]); */
                gm[i]=(id[i+1]-id[i])/(vg[i+1]-vg[i]);

            gmmax=0.0;
            for (i=0; i <= points-1 ; i++)
                if ( gm[i] > gmmax )           /*Find max gm*/
                {
                    gmmax=gm[i];
                    j=i;
                }

            Vt=vg[j]-id[j]/gm[j];              /*Find Vt by Vt = Vg - Id/gmmax*/
            sprintf (result,"Vt = %f\n",Vt);
            fprintf (output,"\n%s\n",*paramlist);
            fprintf (output, result,strlen (result));
        }
    }
    D(printf ("> Vt\n");)
    D(printf ("%f\n",Vt);)
}
```

Development of a Statistical Front End for the SAMPLE Process Simulator

Marco Zuniga

SAMPLE is a program for simulation of profiles in lithography and etching in integrated circuit devices. This projects explores its use in experimental design through the development of a statistical front end, which is then employed to compare the results of designed experiments with and without the inclusion of noise in the input parameters.

1.0 Introduction

SAMPLE is a user oriented FORTRAN program for Simulation and Modeling of Profiles in Lithography and Etching [1]. It is capable of simulating the time evolution of topographical features of IC devices as they undergo multiple process steps. SAMPLE can simulate the following basic processing steps: optical, ebeam, x-ray and ion lithography, wet and dry etching, and the deposition of metals and insulators. The simulator also includes contrast enhancement layers and inorganic resists, etching of non-planar layers, and links to the two-dimensional optical simulator SPLAT [2].

As a simulation tool, one of the uses of SAMPLE is to minimize time spent with experimental work in characterizing various processing conditions. As such, its use in quasi-experimental design is of interest. Although computer simulation cannot explicitly include the effects of the natural variability present in experiments, it can still serve as a valuable exploratory tool.

The purpose of this project is to develop a statistical front end for SAMPLE, which will allow for the explicit introduction of variation in the input variables, and the capability of running complete blocks of designed experiments. Our objective is to demonstrate the capability of SAMPLE as a tool for exploratory experimental designs.

2.0 Methodology

A c-shell script file is used to iteratively run SAMPLE according to a specified experimental design. The user specifies a SAMPLE input file, the input variables, the noise level for each variable, and the desired response. The runs to be simulated can be specified according to a full factorial design or an orthogonal array which best fits the input conditions. A model for the response is then determined accordingly. Table 1 specifies possible input variables and responses which can be simulated with SAMPLE.

Optical Lithography commands Electron Beam lithography commands Ion Beam Lithography commands X-ray Lithography commands	Characteristics of final resist profiles as specified through lithography commands
Deposition Commands	Characteristics of final profile after deposition is completed
Etching Commands	Characteristics of final profile after etch is completed

TABLE 1. General inputs and expected responses as simulated with SAMPLE

As an example to demonstrate SAMPLE's capabilities, a three factor, two level factorial is run for a lithography process. Here we wish to establish the dependence of linewidth with respect to exposure dose, mask linewidth, and development time. Noise levels consisting of a sigma of ten percent are added to each input variable, and a three percent noise level is added to the output itself. The model established is then compared with a model obtained with a simulation without noise in order to elucidate the inherent curvature introduced with SAMPLE.

Table 2 includes the input variables, nominal, high and low values, and desired response for the factorial example.

Input variables: dose: 25, 30, 20mj/cm2	Output response
mask linewidth: 0.4,0.42,0.38 µm development time: 60, 70,50 sec	Resist linewidth of 0.4 µm as measured from bottom of resist

TABLE 2. Input variables and output response for lw experiment

3.0 Implementation

The user specifies the input file and places it in the same directory where the script is located. The program then prompts the user to enter the number of variables, followed by the actual names of the variables in the input file. The user then specifies whether orthogonal arrays or full/fractional factorials should be generated with the variables specified. The number of levels of each variable are then entered; if the number of levels of any single variable is greater than two, the program defaults to running orthogonal arrays for the experiment. The user subsequently enters the values corresponding to the levels of each variable.

The user has the option of performing sensitivity analysis of the output versus each of the input variables by specifying noise levels for each. Noise levels are entered as a percentage deviation from nominal values of each variable, which is subsequently mapped into a random number ranging from 0 to 1, and then added to the value specified at each level.

The script subsequently generates the required input files that cover the design space specified, runs SAMPLE accordingly, and extracts the desired parameter from the output files. A model is then calculated relating the response to the specified input variables. Three dimensional response surfaces can then be displayed by employing an interface to the program Mathematica [3].

The input and output files generated are placed in the same directory of the original input file, and are labeled as specified by the user.

As an example, a chemically-amplified negative tone resist process is chosen to run the proposed experimental design. Figure 1 is an example input file with nominal values for the input variables of this experiment.

#Optical Lithography Example

#Input file samop10

lambda 0.248; #exposure wavelength

proj 0.42; # NA of system

linespace 0.4 0.4; #mask definition

parcohdef 0 0.5 0; #sigma and defocus

vertrespts 200; #number of vertical grid divisions

imagerun;

resmodel (0.248)(-0.712 1.157 0.0029)(1.79 -0.02) (1.0);

#resist exposure parameters

dose 25; #exposure dose in mj/cm^2

shipleyahr; resist bake parameters

layers (1.7 -3.38); # matched substrate index

exposerun;

devrate 4 (6.5 0.0350 6.3); #resist development parameters

devtime 60; #

developrun;

FIGURE 1. Input file describing the computer-based experiment

---

The input file is made up of keywords, followed by a parameter listing. Comments are preceded by the '#' sign. *Lambda* specifies the wavelength of the exposure tool, in this case 0.248 μm. *Proj* defines the NA (numerical aperture of the system) to be 0.42. The *linespace* keyword defines a mask of equal lines and spaces 0.4 μm wide. *Parcohdef* specifies the partial coherence and defocus of the exposure system. When the first parameter is set to one, pure incoherence is indicated; otherwise a partial coherence factor is indicated by the second parameter. The third parameter indicates the amount of defocus present in the system. The keyword *vertrespts* defines the number of vertical sublayers into which the photoresist is divided for numerical computations, in this case 200. *Imagerun* simply runs the image specified in the input file. The *resmodel* command specifies the exposure wavelength for the resist, the ABC parameters as defined by Dill's exposure model [4], the resist's index of refraction at the specified wavelength, and the resist's thickness. *Dose* sets the exposure dose specified in mj/cm^2. The command *shipleyahr* specifies default post exposure bake (PEB) conditions for the Shipley's SNR248 resist. The default parameters are as follows: bake time-60 sec, bake temperature 130C, and diffusion length 0.08μm. *Layers* indicates the number of layers present, excluding photoresist, and the refractive index of each. *Exposerun* runs the exposure simulation. *Devrate* specifies the development model with its corresponding parameters as defined in the SAMPLE user guide. *Devtime* sets the development time, in this case 60 seconds. Finally, *developrun* runs the development simulation. A plot of a simulated resist profile is included in the Appendix.

4.0 Results and Discussion

The following table summarizes the results obtained for the proposed factorial experiment. The goal is to achieve 0.4μm resist lines.

Mask Line width (μm)	Exposure dose (mj/cm^2)	Development Time (sec)	Resist Linewidth (μm)
0.38-	20-	50-	0.40
0.42+	20+	50-	0.45
0.38-	30+	50-	0.34
0.42+	30+	50-	0.40
0.38-	20-	70+	0.43
0.42+	20+	70+	0.48
0.38-	30+	70+	0.37
0.42+	30+	70+	0.42

TABLE 3. Results from factorial experiment

The experimental design corresponds to a three factor, two level, full factorial experiment. A full factorial design was chosen since only eight simulations are needed for a complete calcula-

tions of the effects in the experiment. For larger experiments, however, full factorials can lead to an excessive number of runs for complete determination of the effects. For this reason, there is an option included with which the user can run orthogonal arrays. Several arrays are available ranging from 3 parameters at 3 levels to 11 parameters at 2 levels and 12 parameters at 3 levels. If the proposed experiment does not match any of the arrays available, the user is given a choice of running the experiment with a full factorial or ending the session.

For the example described, employing Yates algorithm implemented in the script, we obtain the following effects: (Here X, Y, and Z represent mask linewidth, exposure dose and development time respectively)

$$L = .411 + (0.05/2)X + (-0.058/2)Y + (2.5e\text{-}3/2)XY + (0.0275/2)Z + (-2.5e\text{-}3/2)XZ + (-2.5e\text{-}3/2)YZ + (-2.5e\text{-}3/2)XYZ \tag{1}$$

The significance of each effect and curvature are determined through comparison with a Minimum Significant Factor and Minimum Significant Curvature respectively [5]. These comparison factors are specified by the user. Using the values in the table indicate that A, B, C are significant. Curvature is found by taking the absolute value of the average of the factorial experiment subtracted from the center point value which in this case is 0.026. Thus, we conclude that curvature is significant.

Moreover, the simulator can be used in order to determine the resist linewidth sensitivity with respect to the three input variables through direct sensitivity analysis by measuring the change in the response divided by the change in the input variable. For example, development time is a hard to control input variable. After conducting a direct sensitivity analysis, we obtain the smallest deviation in linewidth when the resist is processed at $0.42\mu m$ mask lines, 20 mj/cm^2 exposure dose, and 70 sec development time. This is at odds with the objective of the experiment: the printing of 0.4 micron lines in the resist, since this corner yields a high value for the resist line width. This exercise elucidates one of the many trade-offs in experimental work: the region of operation which is least sensitive to noise in a specific input variable may not overlap that which is best for the performance goal. Such analysis can be carried out for each or all of the input variables. According to Taguchi, first you should the noise sensitivity and then target the process.

Furthermore, one can explore the possibility of improving the characterization by including the trend of exposure dose as predicted without noise. We expect exposure values to introduce significant curvature in our response. SAMPLE generates a matrix of values corresponding to the photoactive compound present upon exposure based upon the iterative solution of Dill's exposure model based on the ABC parameters of the resist. Thus, we can run the simulator with exposure dose set at its three noise levels while holding the other two variables at their center values.

The curvature should be reduced by subtracting the corresponding output result from each run in the factorial design. This procedure should elucidate to what degree does exposure introduce curvature in SAMPLE.

Running the simulator at the high, nominal and low values of exposure dose we obtain linewidths of 0.442, 0.409, and 0.387 respectively. Recalculation of curvature gives a value of 0.008, which is comparable to the new MSC of 0.0075.

Mask Linewidth (μm)	Exposure dose(mj/cm^2)	Development Time (sec)	Resist Linewidth (μm)
0.428 (+√2)	25 (0)	60 (0)	0.4
0.372 (-√2)	25 (0)	60 (0)	0.35
0.4 (0)	32.1 (+ 2)	60 (0)	0.4
0.4 (0)	17.9 (- 2)	60 (0)	0.32
0.4 (0)	25 (0)	74.1 (+ 2)	0.37
0.4 (0)	25 (0)	45.9 (- 2)	0.4

TABLE 4. Summary of Experiment

The user can then plot three dimensional contours of the response by using an interface to Mathematica. For the present model we obtained the following equation:

$$LW = 0.411 + ((0.05)/2)X + ((-0.058)/2)Y + ((0.0275)/2)Z \qquad (2)$$

By setting the model to different constant values of linewidth, we obtain equations of planes which can be solved and plotted by Mathematica. The user simply specifies the value of the output for which the equation should be solved. Three dimensional contour plots are included in appendix II.

Finally, the user also has the option of fitting a quadratic model when significant curvature is present in the linear model. If this is the case, the user has the option of running more simulations as needed to fit a quadratic model for the data. For the previous example, additional input files are generated in order to complete a "star" design consisting of six axial points.[6] The additional simulated runs are depicted in table 4. These data is fitted to a second degree polynomial approximation of the following form:

$$Y = b0 + b1(A) + b2(B) + b3(C) + b4(A^2) + b5(B^2) + b6(C^2) + b7(AB) + b8(AC). \qquad (3)$$

The second order equation fitted by least squares from the results of the composite design is as follows:

$$y = 0.353 + 0.023(A) - 0.010(B) + 0.006(C) + 0.018(A^2) + 0.0106(B^2) + 0.023(C^2) + 0.00125(AB) - 0.00125(AC) - 0.00125(BC). \qquad (4)$$

Run	Mean	X	Y	Z	XY	XZ	XYZ	L
1	+	-	-	-	+	+	-	0.40
2	+	+	-	-	-	+	+	0.45
3	+	-	+	-	+	-	+	0.34
4	+	+	+	-	-	-	-	0.40
5	+	-	-	+	-	-	+	0.43
6	+	+	-	+	+	-	-	0.48
7	+	-	+	+	-	+	-	0.37
8	+	+	+	+	+	+	+	0.42
Sum +		1.75	1.53	1.65	1.64	1.64	1.64	
Sum -		1.54	1.76	1.64	1.65	1.65	1.65	
Overall Sum		3.29	3.29	3.29	3.29	3.29	3.29	
Difference		0.21	-0.23	0.01	-.01	-.01	-.01	
Effect		0.05	-.058	.0025	-.0025	-.0025	-.0025	

Center Point=0.39

MSF=9.9e-3
MSC=0.016

TABLE 5. An interface to Mathematica was used in order to solve the necessary matrix equations. A three dimensional contour plot is included in the Appendix.

5.0 Conclusions

The statistical front end for SAMPLE allows the user to run matrixes of designed experiments in batch mode by simply specifying the input variables, their values and noise values at each level requested. The c-shell script is general enough to allow the user to run designed experiments with SUPREM by simply specifying the format for the call to the program.

In addition, if replications from laboratory experiments are available, the noise levels of the input variables could be tuned as to match the spread of variability evidenced by the experimental replications. This could be done by conducting a sensitivity analysis of the input variables in the design space specified, determining which are significant and then iteratively running the simulator until the spread of the output matches the spread of the replications as determined by an F-test with the appropriate degrees of freedom.This assumption is true only if the prerequisites for the F-test are met by the design in question. Design experiments could then be carried out with minimal verification with laboratory work.

Simulated_Resist_Profile
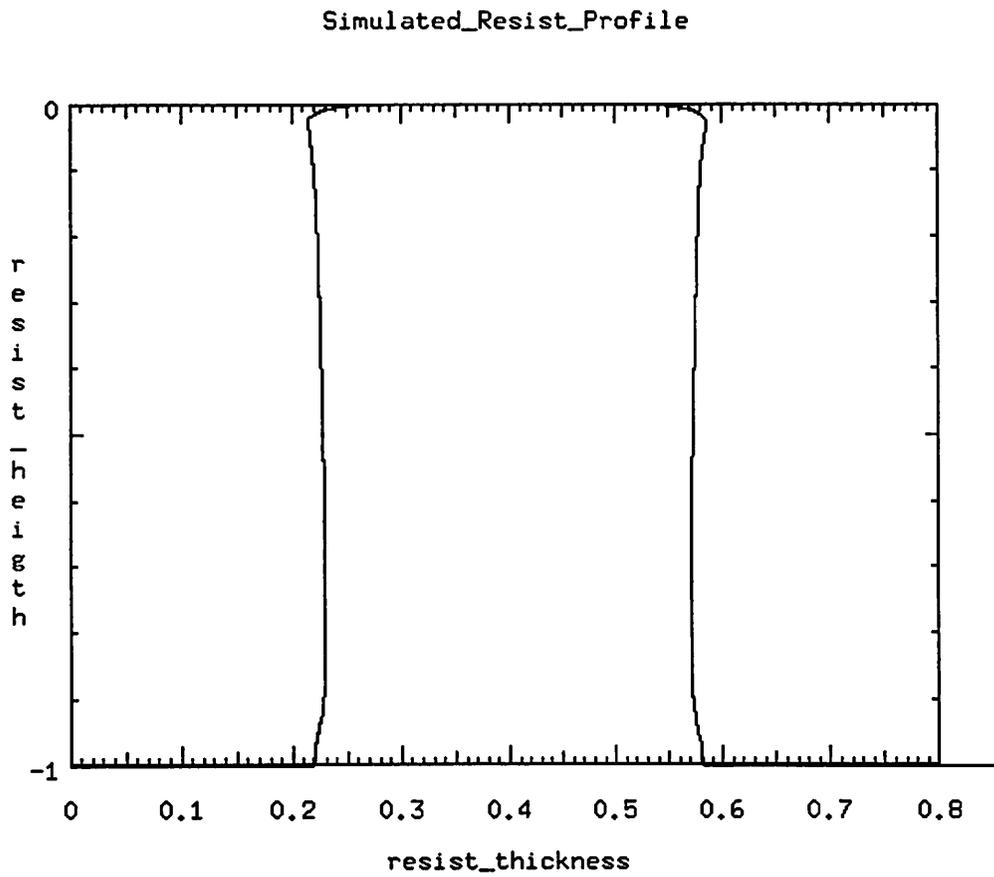
FIGURE 1. Simulated resist profile obtained from SAMPLE input file described for the example factorial design.

Contour Plot for Linewidth=0.4um (linear model)
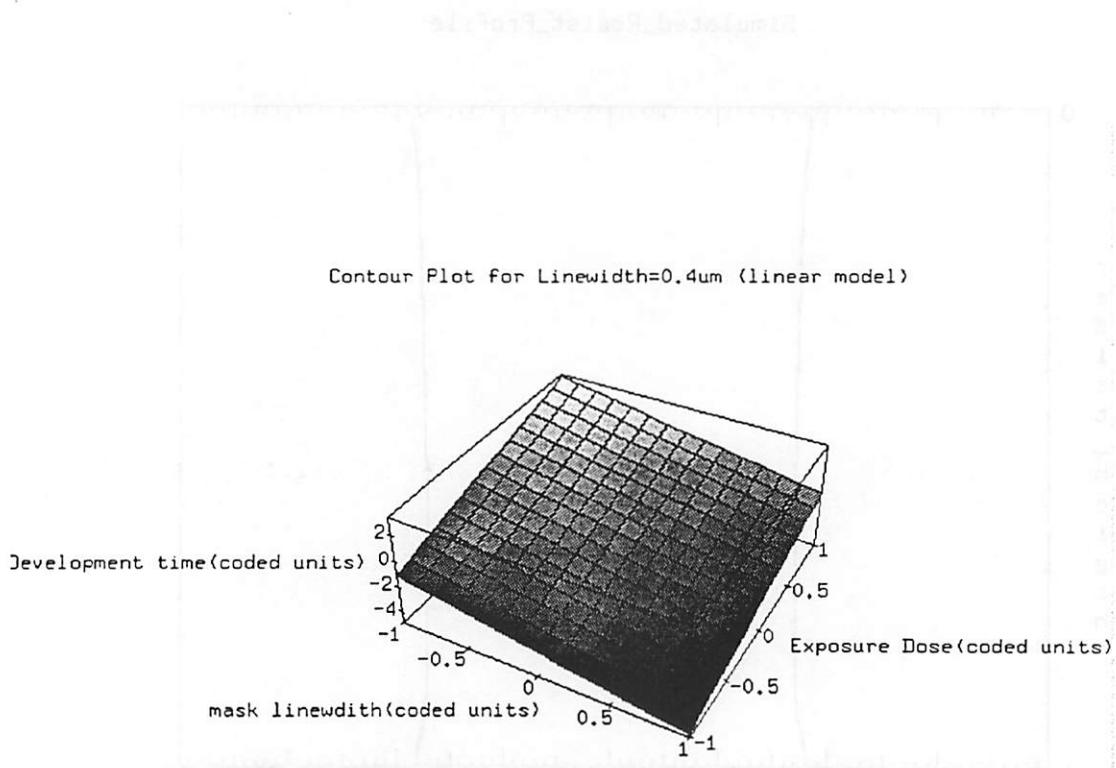
FIGURE 2. Three dimensional contour for constant linewidth of 0.4μm using linear model for which curvature was significant.
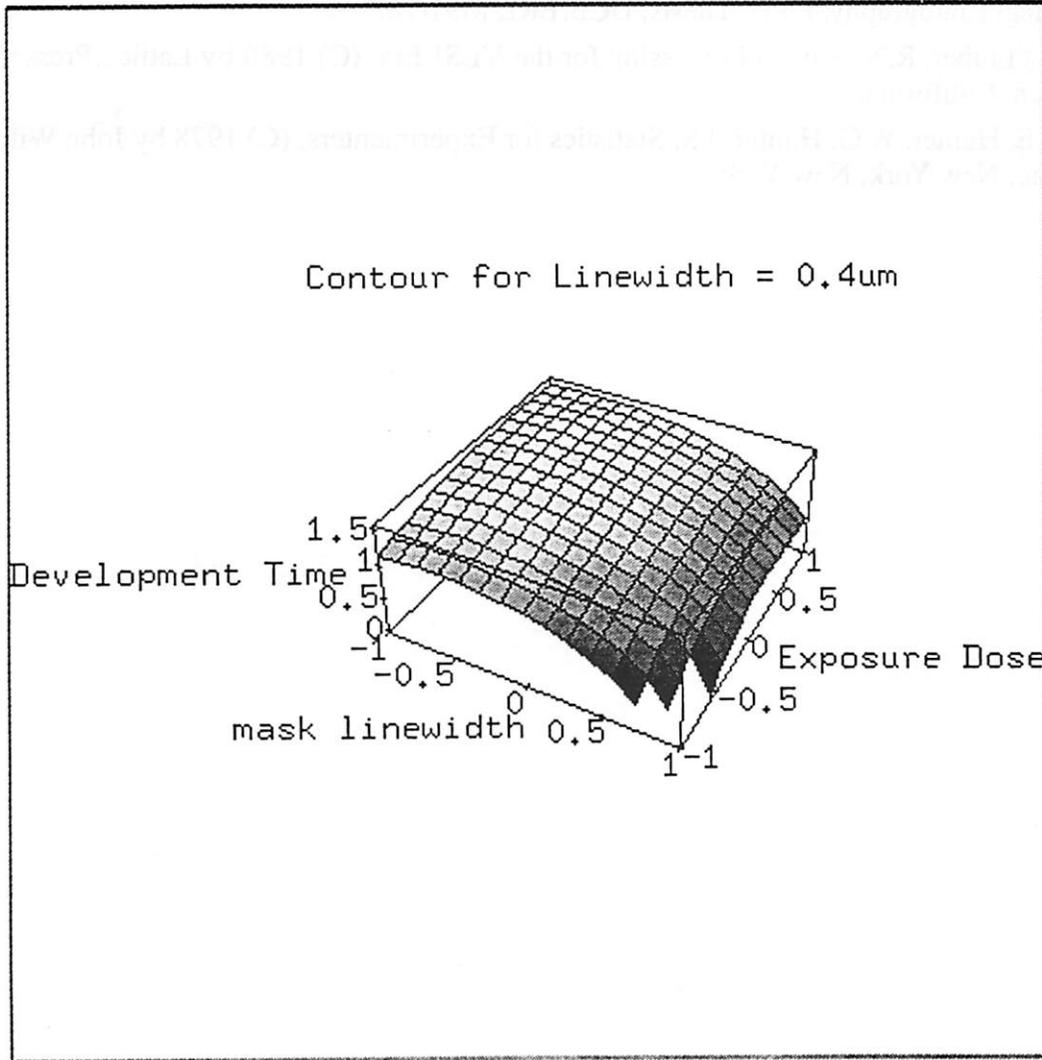
FIGURE 3. Three dimensional contour of constant linewidth of 0.4um using quadratic model for which curvature was not significant.

[1] SAMPLE User Guide, (Version 1.8a, June 1, 1991) (C) 1991, Electronics and Research Laboratory, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley.

[2] SPLAT User Guide (Version 4.2 April 12,1993) (C) 1991, Electronics and Research Laboratory, Department of Electrical and Computer Sciences, University of California at Berkeley.

[3] Mathematica (C) 1988 by Wolfram Research Inc.

[4] Ferguson, R.Modeling and Simulation of Reaction Kinetics in Advanced Resist Processes for Optical Lithography, Ph.D. Thesis, UCB/ERL M91/78.

[5] Wolf, S. Tauber, R.N. Silicon Processing for the VLSI Era, (C) 1986 by Lattice Press, Sunset Beach, California.

[6] Box, G.E. Hunter, W.G. Hunter, J.S. Statistics for Experimenters, (C) 1978 by John Wiley & Sons, Inc. New York, New York.

Employing SPC on a Low Pressure Chemical Vapor Deposition Process

Pamela Tsai

Control charts for individuals were designed for the monthly process monitoring data from tylan9 LPCVD Si Nitride furnace tube in the Microlab. These control charts monitor essential aspects of the tube's performance, such as the deposition rate, within wafer non-uniformity and tube non-uniformity. The definitions of within wafer non-uniformity and tube non-uniformity have been defined in a new, more statistically sound form. These control charts will replace the current standards of control for this furnace.

1.0 Introduction

In order to maintain the quality and efficiency of the Microlab, process monitoring is performed regularly to verify that essential processes are within control standards. The LPCVD Si Nitride furnace tube ("tylan9") is an important and frequently used piece of equipment that is monitored once a month.

Part of the current tylan9 monitoring procedures consists of loading three 4" standard clean, bare Si test wafers into the tube, running the standard recipe (with a deposition time of 30 minutes and a temperature = 800 C) and then measuring the film thickness on each wafer with the ellipsometer at five points (top, center, flat, left, and right, 2 cm from the edge). The average thickness for each wafer and for all three wafers are then automatically calculated. Also calculated are the deposition rate (= average for all three wafers / 30 minutes), thickness uniformity across each wafer (= [Total Range/Total Mean] * 100) and thickness uniformity between wafers (= [Wafer Range/Wafer Mean]) * 100). The control standards currently used are as follows:

- Deposition Rate: 40 A/min +/- 10%

- Within wafer% thickness non-uniformity: 5%

- Wafer-to-wafer% thickness non-uniformity: 20%

While this approach can give a general idea of whether or not the tube's behavior is in control, using statistically defined control charts to monitor the appropriate measurements would provide a more systematic and rigorous indication of the tube's performance. Using control charts would also provide immediate visual feedback by indicating trends and other anomalous behavior that otherwise might not be obvious. In contrast to the current standards which were set somewhat arbitrarily, the control limits would be statistically derived. The goal of this project is to propose and implement the appropriate control charts for this operation.

2.0 Methodology and Results

2.1 Choosing the Type of Control Chart

Because the tylan9 is monitored only once a month, it is impractical to require sample sizes of n>1 for each control point on the chart. Therefore, control charts for individual measurements were selected to monitor this data. Control charts for individual measurements use the moving range of two successive observations to estimate the process variability, defined as $MR_i = |X_i - X_{i-1}|$. For the control chart for individuals, the parameters are:

- Upper Control Limit $= \overline{X} + 3\overline{MR}/d_2$

- Center line $= \overline{X}$

- Lower Control Limit $= \overline{X} - 3\overline{MR}/d_2$

Where $\overline{X}$ is the total average, $\overline{MR}$ is the average of all the moving ranges, and d_2 is a statistically defined constant. The value for this constant for n=2 is 1.128. Although it is also possible to establish a control chart on the moving ranges, it was inappropriate to do so in this case. This is because the values that are ultimately being plotted on the control charts are actually averages of measurements, and it turns out that these measurements are not independent. This issue will be discussed later in some detail.

2.2 Choosing the Appropriate Parameters

The next step was to choose the appropriate parameters to plot on the control charts. Ultimately, three values were of interest: the total average thickness, the within tube non-uniformity, and the within wafer non-uniformity.

2.2.1 The Total Average Thickness

The variation in the total average thickness appeared to be random, so no further statistical analysis was considered in calculating its values. The total average thickness was calculated in the same manner as before, which was as the average of the three wafer averages (which is equivalent to the average of all the thickness measurements taken for the run).

2.2.2 The Tube Non-uniformity

When performing the monitoring runs, the 3 control wafers are loaded in the following manner:

- the first wafer, called the "load wafer", is loaded in the front of the front boat

- the second wafer, called the "center wafer", is loaded in the front of the rear boat

- the third wafer, called the "source wafer", is loaded in the back of the rear boat

The remaining slots are loaded with clean dummy wafers in every other slot.

In almost all of the runs, the load wafer had the thickest film, and the source wafer had the thinnest (due to gas depletion effects.) Therefore, the film was deposited throughout the tube in a deterministic manner. In order incorporate this deterministic observed behavior, the tube non-uni-

---

formity was calculated as the average of the "load wafer" minus the average of the "source wafer". This new value will replace the one previously called wafer-to-wafer variation, which did not take into account the deterministic behavior. The previous definition also involved dividing by the global average, which would further convolute things as the global average is already being monitored by the total average thickness control chart.

2.2.3 The Within Wafer Non-uniformity

The issue of defining the within wafer non-uniformity had to be dealt with carefully. The definition used before for the within wafer non-uniformity for each wafer was to take the difference of thickest measurement on the wafer and the thinnest measurement on the wafer and divide it by the average thickness of that wafer.

However, this definition may not be appropriate, because the film is not necessarily deposited on the wafer in a uniform, random way; the deposition may have a deterministic pattern. For example, the profile of the deposited film on the wafer may always be concave in the middle, bowl-shaped, heavier always on the left side, etc. In this case, it appeared as though the flat, center, and left points were consistently thinner than the right and top points. This was confirmed with a paired t-test (see Table 1), where, for each run, the average of the flat, center, and left points for each wafer was compared to the corresponding average of the right and top points. Note the extremely small p-values for the null hypothesis which states that the average of the flat, center, and left points is not smaller than the average of the right and top points. Therefore, this null hypothesis can be rejected with a high level of confidence.

Therefore, a more appropriate measurement for the within wafer non-uniformity would be to take the difference of the average of the right and top points and the average of the flat, center, and left points and then dividing by the smaller of the two, which is almost always going to be the average of the flat, center, and left points.

The next step for selecting the appropriate value to use in the within wafer non-uniformity control chart was to find out whether or not the within wafer non-uniformities for each of the three wafers could be combined into one figure of merit, thereby necessitating one control chart for within wafer non-uniformity instead of three (a separate one for each wafer.)

In order to combine the three values into one chart, there could not be any reason to believe that they were any different from each other, i.e. that there was no deterministic behavior in the wafer ranges. For example, it would not be appropriate to combine the three wafer ranges if the source wafer range was consistently much higher than those of the center and load wafers. In this case, it appeared that the difference in the wafer ranges for the three wafers were random. This was confirmed with a paired t-test that was performed for each pair of wafer ranges. (See Table 2.) The null hypothesis stated that the wafer ranges for each pair of wafers being compared were no different. For all three tests, the p-values were high, ranging from 0.27 to 0.60, which means that the null hypothesis could not be rejected with a high level of confidence. Therefore, the within wafer non-uniformities for each of the three wafers were combined into one figure of merit and their average value was monitored by one control chart.

2.3 Plotting the Control Charts

Once the appropriate values have been chosen, the next step was to plot these values on their corresponding control charts. If out-of-control points or trends were noticed, then assignable causes would have to be found for them. Then new control limits can be calculated, excluding the out-of-control points.

2.3.1 Plotting the Control Charts: the First Pass

Table 3 shows the run numbers considered in building the control charts and their corresponding dates. Figure 1 shows the three control charts with all runs included in the calculation of the control limits. There are clearly out-of-control points and trends evident in these first round charts. In the total average thickness control chart, there is as shift around run 9-11 where the average seems to have shifted to a higher value. Also in the total average thickness chart, runs 22 and 23 seem to be out of control. In the within wafer% non-uniformity chart, it is evident that run 10 had some problems. In the tube non-uniformity chart, run 22 plotted out of control, and runs 14 and 21 look like they are pushing the upper control limit. The next step would be to try to explain these occurrences with assignable causes.

2.3.2 Finding Assignable Causes for Out of control Points and Trends

All problems and comments concerning the tylan9 furnace tube were recorded on the Wand. Therefore, it was possible to find potential causes for the seemingly out-of control points on the control charts.

In the total average thickness chart, the shift in the mean occurring in runs 10 to 11 may be explained by the fact that a leaky bellows assembly was replaced before the shift. This may have influenced the pressure in the tube, which in turn may have affected the deposition rate.

Between runs 21 and 22, there was a tube lining change and a gas filter change which may be why run 22 is so much higher than the other runs. Changing the gas filter may have contributed to an unusual build-up in the area where the process pressure is monitored; the gauge that monitors the process pressure, if damaged or clogged, will have erroneous pressure readings, resulting in out-of-control deposition rates. Run 23 is also very high; there was very little activity between 22 and 23 as it was Christmas break. The condition seemed to have corrected itself from run 23 to 24, where the average thickness returned down to its lower value.

In the tube non-uniformity control chart, runs 14, 21 and 22 seem to have some problems. In the period right before run 14 was performed, some repair work was done on tylan9, which may have caused the higher than normal tube non-uniformity. Before run 21, new quartzware was installed, which may have affected the tube uniformity. As noted before, there was a tube lining change and a gas filter change before run 22 which may have caused some disturbances in the monitor run.

In the within wafer% non-uniformity chart, run 10 seemed to have exhibited some out of control behavior. During that month, a third boat was added in the tube, which may have disrupted the usual gas flow in the tube during the monitoring. This could have been the cause of the greater wafer non-uniformities.

Figure 2 shows the control charts with the explained out-of-control points excluded from the control limit calculations.

2.3.3 Re-evaluating Control Chart Limits

The control limits for these charts should be re-evaluated whenever it appears that the process has reached a new, acceptable state of statistical control. The total average chart seems more prone to having to have its limits changed than the non-uniformity charts. The shift in the total average chart from 10 to 11 is an example of a situation where new control limits should be calculated.

3.0 Implementation

This method of using control charts to monitor the behavior of the tylan9 may be implemented by using the xspread program. A proposed program may involve the process staff entering the raw thickness values (5 per wafer). The program would then calculate the following values:

- total average thickness = average of all the readings

- tube non-uniformity = average of the load wafer - average of the source wafer

- wafer% non-uniformity (for each wafer) = (avg of right and top measurements-avg of the flat, center, and left measurements) /(avg of the flat, center, and left measurements)

- within wafer% non-uniformity = average of the 3 wafer% non-uniformities

and put the average thickness, tube non-uniformity, and the within wafer% non-uniformity values into the control charts.

4.0 Conclusions

Control charts have been designed for the tylan9 furnace tube in the Microlab. They consist of 3 charts, one for the total average wafer thickness, within wafer% non-uniformity, and tube non-uniformity. The definitions for within wafer% non-uniformity and tube non-uniformity have been altered to be more statistically sound. These control charts would provide a more rigorous indication of the tube's behavior than the current method of using qualitatively set ranges. They provide visual information that can facilitate detecting trends and shifts that may otherwise gone unnoticed. Their control limits must be re-evaluated whenever it appears that the process has reached a new state of statistical control. A program to automatically plot the appropriate measurements on the control charts is currently being worked on. This procedure can also be applied to other Microlab systems.

[1] D. C. Montgomery, "introduction to Statistical Quality Control," John Wiley & Sons, New York, 1985.
[2] D. hebert et al., "The Microlab Manual," Ch. 1.14. pp. 1, 17.

TABLE 1. Within wafer uniformity paired analyses

<u>Within Wafer Paired Analysis: Load</u>

Null Hyp: Median of population of differences (1 - 2) >= 0
Alt Hyp: Median of population of differences (1 - 2) < 0

Median of sample of differences (1 - 2) -15.42000
Interquartile Range 11.75000
# of points 20.00000
SE of Median diff 2.44410
95 percent Confidence Interval Median diff < -11.1939
t statistic -6.30910
p-value 0.00005
Null hypothesis can be rejected at the 0.01 percent level.

<u>Within Wafer Paired Analysis: Center</u>

Null Hyp: Median of population of differences (1 - 2) >= 0
Alt Hyp: Median of population of differences (1 - 2) < 0

Median of sample of differences (1 - 2) -8.170000
Interquartile Range 13.085000
# of points 20.000000
SE of Median diff 2.721800
95 percent Confidence Interval Median diff < -3.4637
t statistic -3.001700
p-value 0.003667
Null hypothesis can be rejected at the 0.37 percent level.

<u>Within Wafer Paired Analysis: Source</u>

Null Hyp: Median of population of differences (1 - 2) >= 0
Alt Hyp: Median of population of differences (1 - 2) < 0

Median of sample of differences (1 - 2) -9.505000
Interquartile Range 16.000000
# of points 20.000000
SE of Median diff 3.328100
95 percent Confidence Interval Median diff < -3.7503
t statistic -2.856000
p-value 0.005055
Null hypothesis can be rejected at the 0.51 percent level.

TABLE 2. Pairwise analyses of the wafer ranges

Wafer Range Paired Analysis: load and center

Null Hyp: Median of population of differences (1 - 2) = 0
Alt Hyp: Median of population of differences (1 - 2) <> 0

Median of sample of differences (1 - 2) 0.310500
Interquartile Range 1.330000
# of points 20.000000
SE of Median diff 0.276600
95 percent Confidence Interval -0.2685 < Median diff < 0.8895
t statistic 1.122400
p-value 0.275692
Null hypothesis cannot be rejected at the 20 percent level.

Wafer Range Paired Analysis : load and source

Null Hyp: Median of population of differences (1 - 2) = 0
Alt Hyp: Median of population of differences (1 - 2) <> 0

Median of sample of differences (1 - 2) 0.280000
Interquartile Range 1.190000
# of points 20.000000
SE of Median diff 0.247500
95 percent Confidence Interval -0.2381 < Median diff < 0.7981
t statistic 1.131200
p-value 0.272053
Null hypothesis cannot be rejected at the 20 percent level.

Wafer Range Paired Analysis: center and source

Null Hyp: Median of population of differences (1 - 2) = 0
Alt Hyp: Median of population of differences (1 - 2) <> 0

Median of sample of differences (1 - 2) -0.135000
Interquartile Range 1.310000
# of points 20.000000
SE of Median diff 0.272500
95 percent Confidence Interval -0.7053 < Median diff < 0.4353
t statistic -0.495400
p-value 0.625975
Null hypothesis cannot be rejected at the 60 percent level.

TABLE 3. [1]dates when the runs were performed

run #	date
1	1/7/91
2	1/29/91
3	3/8/91
4	4/5/91
5	6/5/91
6	7/3/91
7	7/29/91
8	9/9/91
9	10/4/91
10	11/8/91
11	12/30/91
12	2/7/92
13	3/24/92
14	5/1/92
15	5/18/92
16	6/2/92
17	6/22/92
18	7/29/92
19	8/26/92
20	10/2/92
21	11/10/92
22	12/18/92
23	1/5/93
24	1/12/93
25	2/1/93
26	3/3/93
27	3/31/93
28	4/28/93

1. the data needed to calculate within wafer % non-uniformity were unavailable in runs 1-7; therefore, its control chart begins at run8.
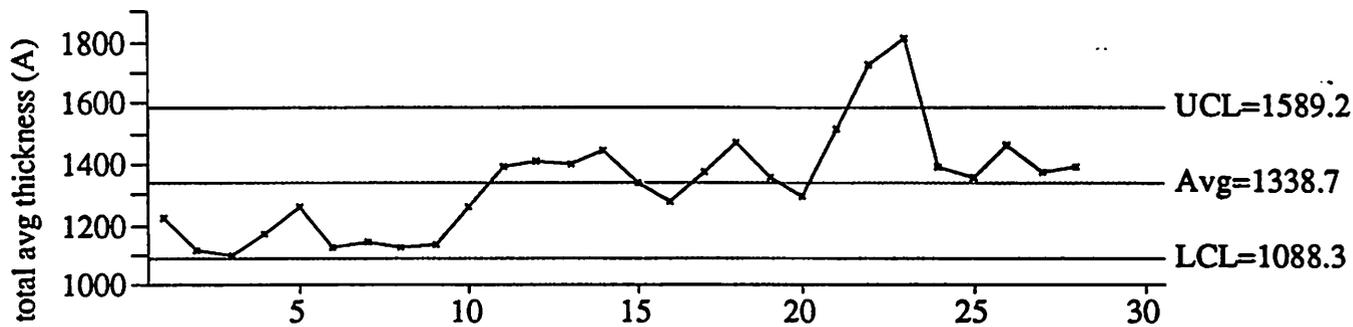
total avg thickness (A)
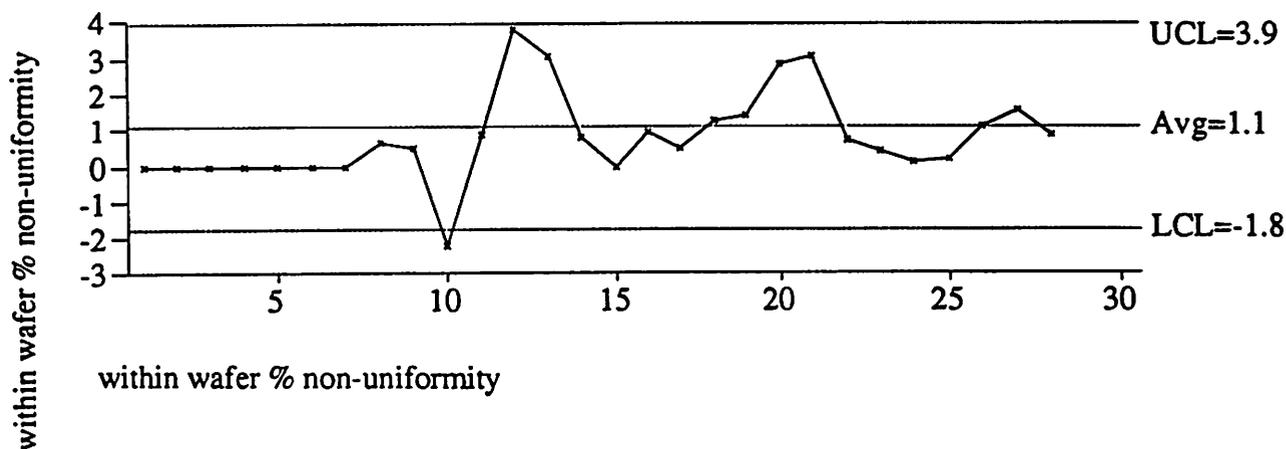
within wafer % non-uniformity
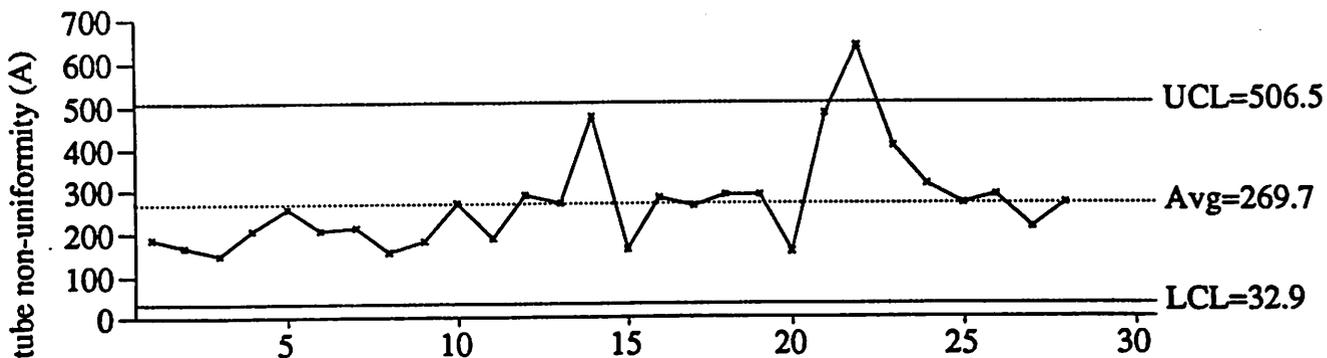
tube non-uniformity (A)

FIGURE 1. Control Charts using all runs to calculate the control limits.
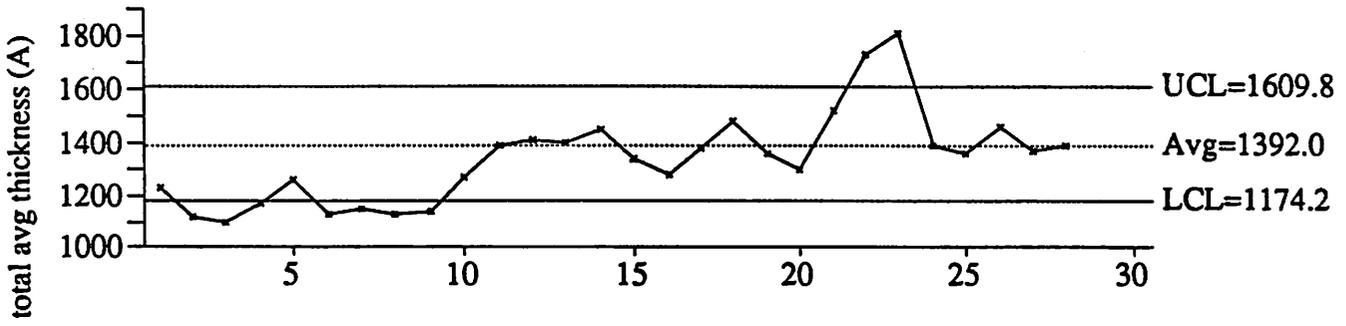
total avg thickness (A)
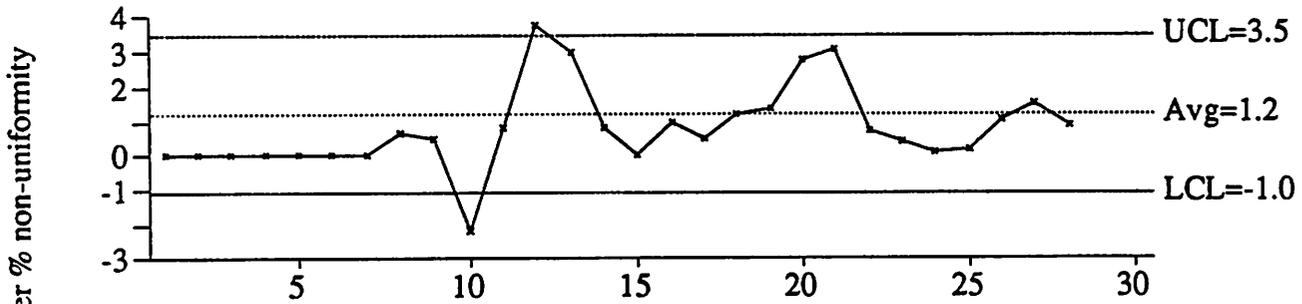
within wafer % non-uniformity
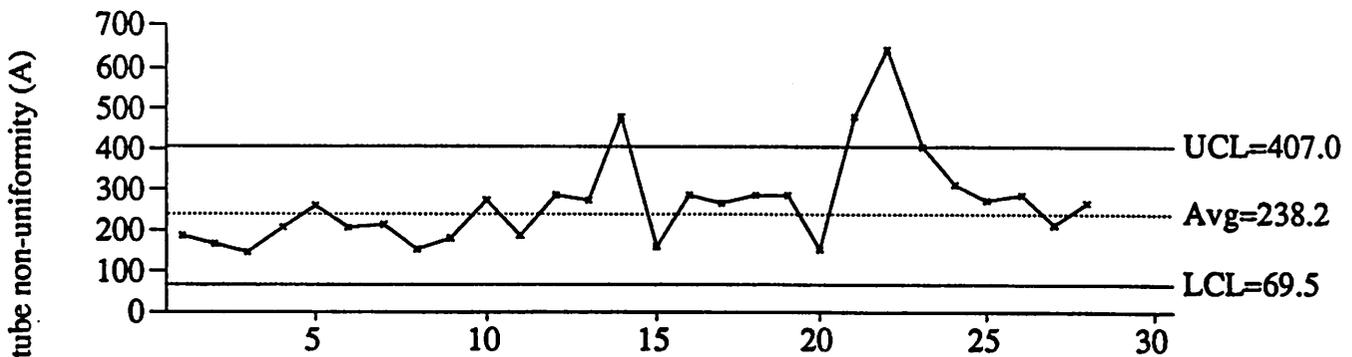
tube non-uniformity (A)

FIGURE 2. Control Charts excluding the explained out-of-control runs when calculating the control limits.

Modeling Photoresist Etching in a High Frequency RIE Discharge

John Hutchinson

The use of high frequency reactive ion etching (RIE) of photoresist is particularly suited for silylated resist processes, exhibiting lower self (DC) bias, higher selectivity, etch rate and lower post etch residues. In this work, orthogonal arrays are used to model the effect of process parameters on the self bias, and the etch rate of unsilylated resist. Understanding of the self bias and resist etch rate is key to modeling of more complex systems such as the silylated resist etching.

1. Introduction

As feature sizes on integrated circuits shrink below 0.3 microns, the process window available at 365 nm (i-line) and 248 nm (KrF excimer laser) decreases significantly. Near and top-surface imaging is seen as a route to improved resolution for optical lithography. Surface imaging can be accomplished by selective incorporation of silicon (silylation) into the resist film, which acts as an etch barrier during subsequent oxygen plasma etching. Gas-phase silylation of resist is the focus of much current research, and is typified by the DESIRE process[1].

A significant challenge in developing a production worthy silylated resist process is the elimination of post dry development residues in unsilylated areas. Post-etch residues are believed to result from sputtering of the oxide mask into previously etched areas. MERIE, helicon and ECR etchers have been investigated as equipment based solutions to residue formation. One inexpensive alternative to ECR and magnetron etchers is to retain the parallel plate diode reactor design and utilize a higher driving frequency. As the frequency driving an RF discharge is raised, the DC bias is lowered, resulting in a lower ion energy, which gives an etch process a lower sputtering rate and a higher selectivity of unsilylated resist over silylated resist etching rates. In addition, a reduced sputtering rate should result in decreased resputtered oxide residue formation.

The purpose of this study is to experimentally verify and model the behavior of DC bias and unsilylated resist etch rate in a high frequency RIE reactor. Once the models of unsilylated etch rate and self bias are built, this will form a framework for future work in the modeling of silylated etch rate and selectivity in high frequency RIE discharges.

2. Methodology

An L9 orthogonal array across the four factors of etch frequency, generator power, oxygen pressure and oxygen flow is used to examine the process parameter effects on the self bias and the unsilylated resist etching rate. The four factors are varied at three levels, and the responses of DC bias and etch rate are measured(See table 1). Although an orthogonal array is not typical method used for traditional response surface modeling, the data collected is still relevant to the problem. The data collected were actually used in an earlier experimental design in order to find a first pass optimal process based on etch rate and uniformity as the reactor was being built.

The self (or DC) bias is expected to follow the etch frequency to the -1.5 power for constant generator power (see appendix A) and should be linear with generator power for constant frequency. The effects of flow and pressure are expected to be negligible compared to the frequency and power effects. A model based on the effects of power and frequency will be fit.

Mechanistic models for the etch rate of resist have been proposed[2]. Joubert proposes that the etching of organic resist is an ion enhanced etching process which is maximized in the presence of atomic oxygen and bombardment flux, and can be described by:

$$\frac{1}{ER} = \frac{A}{J_i} + \frac{B}{[O]}$$

where ER is the etch rate of resist, J_i is the ion flux striking the wafer, [O] is the atomic oxygen concentration, and A and B are fitting parameters depending on the reactor and the resist material. However, the ion flux and neutral oxygen concentration are difficult to measure and predict from parameters that can be varied in a typical etch process recipe. Therefore a semi-empirical approach to modeling the etch rate is taken.

3. Implementation

The etcher used in this study is a prototype RIE reactor constructed at the University of California, modified for RF etch frequencies between 10 and 80 MHz. Wafers were coated with a standard DESIRE resist, Plasmask 200G. The wafers were not silylated before etching. Film thicknesses were measured on a Nanospec 210 spectrophotometer. Etch rates were determined by measuring resist etch rates at 5 points across the wafer before and after a fixed two minute etching time. Self bias was measured for each run with a low pass filter at the powered electrode connected to a multimeter.

The L9 orthogonal array effect values and measured responses are summarized in Table 1. Run #1 was replicated twice to determine experimental error in the etch rate measurement. However, the self bias was not re-measured for the replicates of run #1.

4. Results and discussion

4.1 DC Bias

ANOVA analysis of four factors of frequency, power, flow and pressure showed that frequency (ω) and power were significant effects (Power: F = 7.7260 Prob > F 0.0498, Frequency: F = 58.5173 Prob > F 0.0044) while pressure (F=0.8772 Prob > F 0.4900) and flow (F=0.5897 Prob > F 0.5626) were not significant as expected. Removing the pressure and flow effects from the model and transforming the frequency effect to a -1.5 power increased the R^2 value to 0.94812 from the four-factor linear model (R^2 = 0.913482). Both factors were very significant (Power: F = 19.3265 Prob > F 0.0046, $\omega^{-1.5}$: F = 90.3246 Prob > F 0.0001). ANOVA of the whole model yields an F of 54.8255 (Prob > F 0.0001). Resulting plots of measured DC bias versus predicted DC bias, with model fit and 95% confidence interval are shown in figure 1. The result of the model fit yields:

$$V_{rf} = -101.5707 + 2.4 * P_a + 32574.515 * \omega^{-1.5}$$

where P_a is the generator power in Watts, and ω is the frequency in MHz. Because the self bias was not measured during the replicates, lack of fit was not determined for this process.

4.2 Etch Rate

Because a mechanistic model similar to the DC bias model does not exist for the resist etch rate, a semi-empirical model based on a linear combination of the four effects of frequency, power, flow and pressure was attempted to fit first (Figure 2). Results of an effect F-test (See table 2) revealed that oxygen flow rate (F=1.4131) did not have a significant effect. This indicates that in this regime from 10 to 30 sccm of oxygen flow that the etch process is not starved for atomic oxygen, which should increase with oxygen flow rate. However, this simple linear model showed significant lack of fit (F=45.5377).

In order to improve the model, the oxygen flow term was first eliminated. However, significant lack of fit in the model still existed (F=45.1041). Analysis of residuals versus effect (Figure 3) showed that the model fit poorly at high levels of both power and frequency. According to the Joubert model, the etch rate should be proportional to the ion flux, for fixed atomic oxygen concentration. However, Jurgensen[3] suggests that the sputter yield of resist is proportional to the bombardment energy or dc bias, and the etch rate should be proportional to the product of the ion flux and the ion energy. Since the ion energy should be linear with power, and the ion flux increases with the frequency, a new term in the model of the product of power and frequency is suggested. Adding the cross term of frequency and power to the existing three linear terms of power, pressure and frequency yields a model with better lack of fit (F=17.9353 Prob > F 0.0535). However, when including the frequency and power interaction term, the estimate of frequency effect is 0.07744 with a standard error of 4.1900. The estimate of the frequency and power interaction term is -0.289144 with a standard error of 0.08174. This indicates that even with the correlation between

the frequency effect and the interaction term, there is significant evidence that the interaction term is more applicable and the frequency main effect cannot be distinguished from zero. In addition, elimination of the frequency main effect from the model including the interaction term improves the lack of fit (from Prob > F of 0.0535 to 0.0663) still further, providing further evidence that the interaction effect is significant even in the presence of the correlated variables. With this model, the resulting ANOVA shows that all effects of power, pressure and power crossed with frequency are significant. The lack of fit for this model is better (F=14.3490, Prob >F = 0.0664). This means that the null hypothesis cannot be rejected at the 5% level of significance. The resulting model equation for this fit is:

$$\text{Etch Rate} = -220.7118 + 30.065 * P_a + 8.5017 * P_r + -0.2877 * P_a * \omega$$

where ω is the frequency in MHz, P_a is the generator power in Watts, and P_r is the etch pressure in millitorr. The whole model test of measured etch rate versus predicted etch rate is shown in Figure 4. The results of the ANOVA for effects and lack of fit in the revised model are in Table 3.

5. Conclusions

A mechanistic model of the self bias in a high frequency RIE reactor was fit based on a linear dependence on generator power and frequency to the -1.5 power in close accordance with theoretical analysis of self bias effects with frequency ($R^2 = 0.94812$). A semi-empirical model of photoresist etch rate based on factors of generator power, etch pressure and a cross term of frequency and generator power compared well with experimentally measured values ($R^2 = 0.987312$). The resulting models of etch rate and self bias can be used to generate a framework for understanding the effects of high frequency RIE in resist etching.

6. Acknowledgments

The author would like to acknowledge Kent Kalpakjian for his assistance with the etcher and the DC bias scaling equations, and Sovarong Leang for his assistance with the JMP analysis software.

7. References

[1] F. Coopman, B. Roland, Proc. SPIE 631, 34 (1986).
[2] O. Joubert, J. Pelletier, and Y. Arnal, "The etching of polymers in oxygen-based plasmas: A parametric study," J. Appl. Phys. 65 (12), 15 June 1989, p5096-5100.
[3] C. Jurgensen and A. Rammelsberg, "Oxygen reactive ion etching mechanisms of organic and organosilicon polymers," J. Vac. Sci. Technol. A 7(6), Nov/Dec 1989, 3317-3324.
[4] K. Kalpakjian, "High Frequency Reactive Ion Etching of Silylated Resist," M.S. Report, University of California, May 1993.

Appendix 1: Derivation of self bias dependence on frequency in an RIE discharge[4]

Power balance in an discharge dictates that the RF power input to the reactor must equal the power lost within the system. This is expressed by:

$$P_{in} = \left(1 + \frac{\varepsilon_i}{\varepsilon_e + \varepsilon_L}\right)(P_\Omega + P_s) \approx \left(1 + \frac{\varepsilon_i}{\varepsilon_e + \varepsilon_L}\right)P_s$$

where ε_i and ε_e are the energies with which ions and electrons are lost to the walls, respectively, and ε_L is the total loss per electron-ion pair created. The ion loss term is dominated by energy that ions gain crossing the sheath (i.e. the DC bias) and is approximately 0.41 V_{rf}. P_Ω and P_s represent the two primary electron heating mechanisms in RIE discharges, ohmic and stochastic heating. Ohmic heating refers to electrons losing energy through bulk plasma collisions and then gaining energy via the applied electric field. Stochastic heating comes about by electrons reflecting off of the oscillating sheaths and gaining kinetic energy. Typically, the stochastic heating term dominates the power. ($P_\Omega \gg P_s$). The electron loss term is given by $2kT_e$, approximately 2 to 10 volts, and is typically negligible. The total loss per electron-ion pair created is the amount of energy an electron loses before ionizing a neutral oxygen atom. This parameter is highly dependent on the electron energy and the type of neutral species.

Substituting for P_s, and typical values for the energy loss terms:

$$P_{in} = \left(1 + \frac{(0.41)V_{rf}}{306}\right)K'_s \omega^2 V_{rf}$$

which shows a V_{rf} proportionality to ω^{-2} dependency for low V_{rf} and a V_{rf} proportionality to ω^{-1} for large V_{rf} for a constant input power. In the frequency range of 20 - 100 MHz, since ε_i is proportional to V_{rf}, the DC bias, or ion energy is expected to drop as the frequency to the -1.5 power, as the input power is held constant. For increasing power at constant frequency, the DC bias should scale linearly.

Table 1. Orthogonal array effect values and responses

Run #	Power (W)	Pressure (mT)	Flow (sccm)	ω (MHz)	DC Bias (V)	ER (A/min)
1	20	10	10	20	257	329
1'	20	10	10	20		329
1"	20	10	10	20		363
2	20	20	20	40	105	341
3	20	30	30	60	29	201
4	50	10	20	60	105	609
5	50	20	30	20	380	1136
6	50	30	10	40	156	1073
7	80	10	30	40	210	1301
8	80	20	10	60	108	917
9	80	30	20	20	505	1990

Table 2. Effect test for linear model of etch rate

Effect	F	DF	Prob > F
Power	132.6	1	0.0000
Pressure	12.9929	1	0.0113
Flow	1.4131	1	0.2795
Frequency	33.8239	1	0.0011
Whole Model	45.8427	4	0.0001
Lack of Fit	45.5377	4	0.0216

Table 3. Effect test for revised linear model of etch rate

Effect	F	DF	Prob > F
Power	387.784	1	0.0000
Pressure	9.0943	1	0.0195
Freq*Power	109.3175	1	0.0000
Whole Model	181.5650	3	0.0000
Lack of Fit	14.3490	5	0.0664

DC Bias(V) Measured
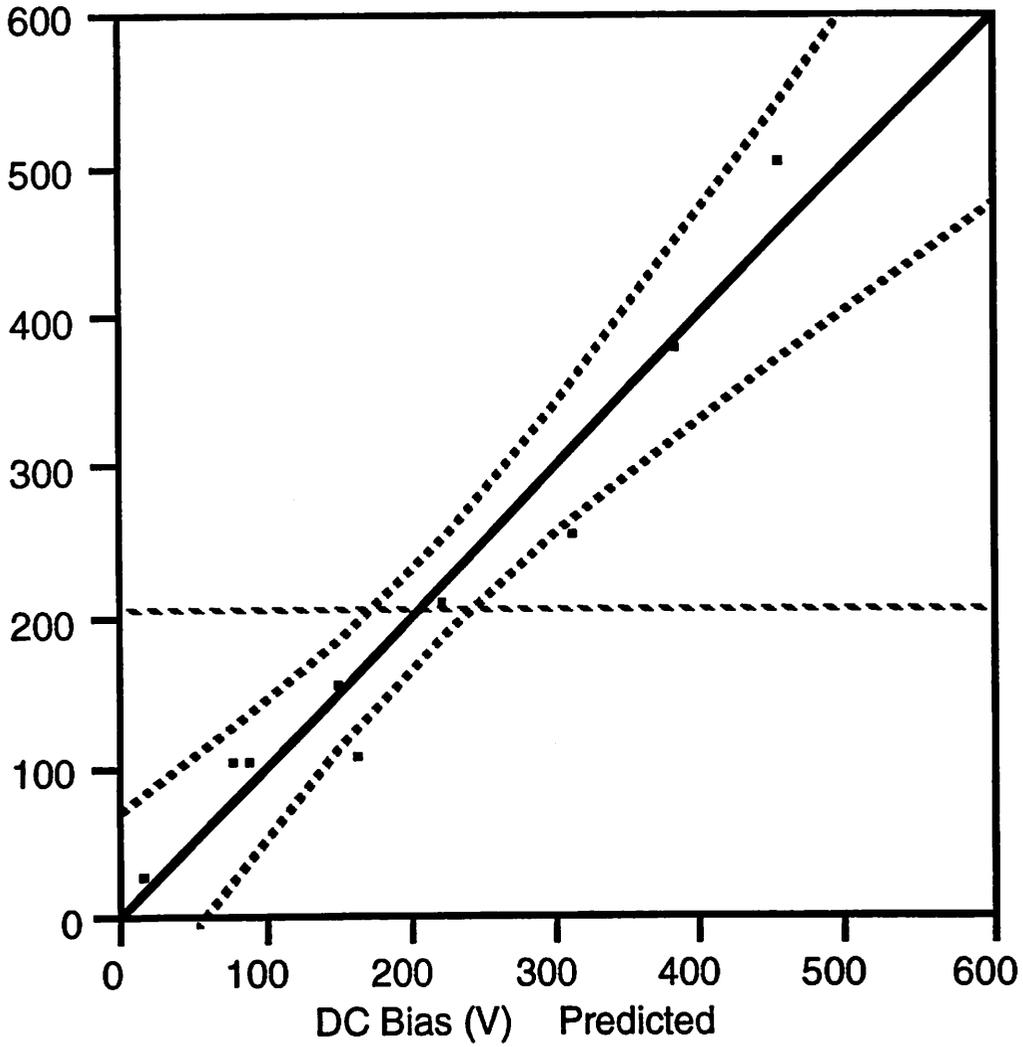
Figure 1. DC Bias measured versus predicted for model based on power and ω⁻
1.5

Etch Rate (A/min) Measured
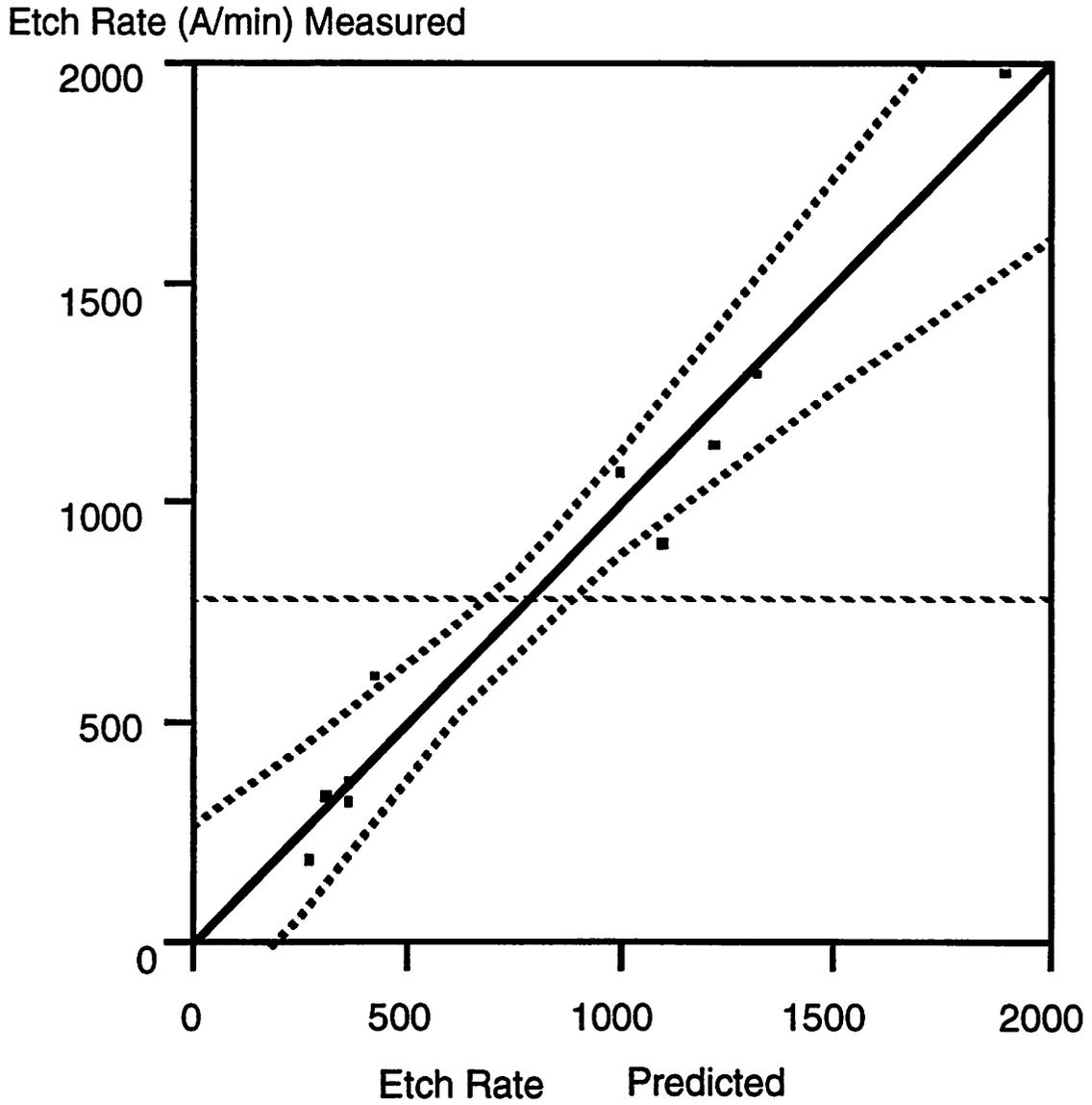
Figure 2. Resist etch rate measured versus predicted for linear model based on all four factors of frequency, power, pressure and flow.
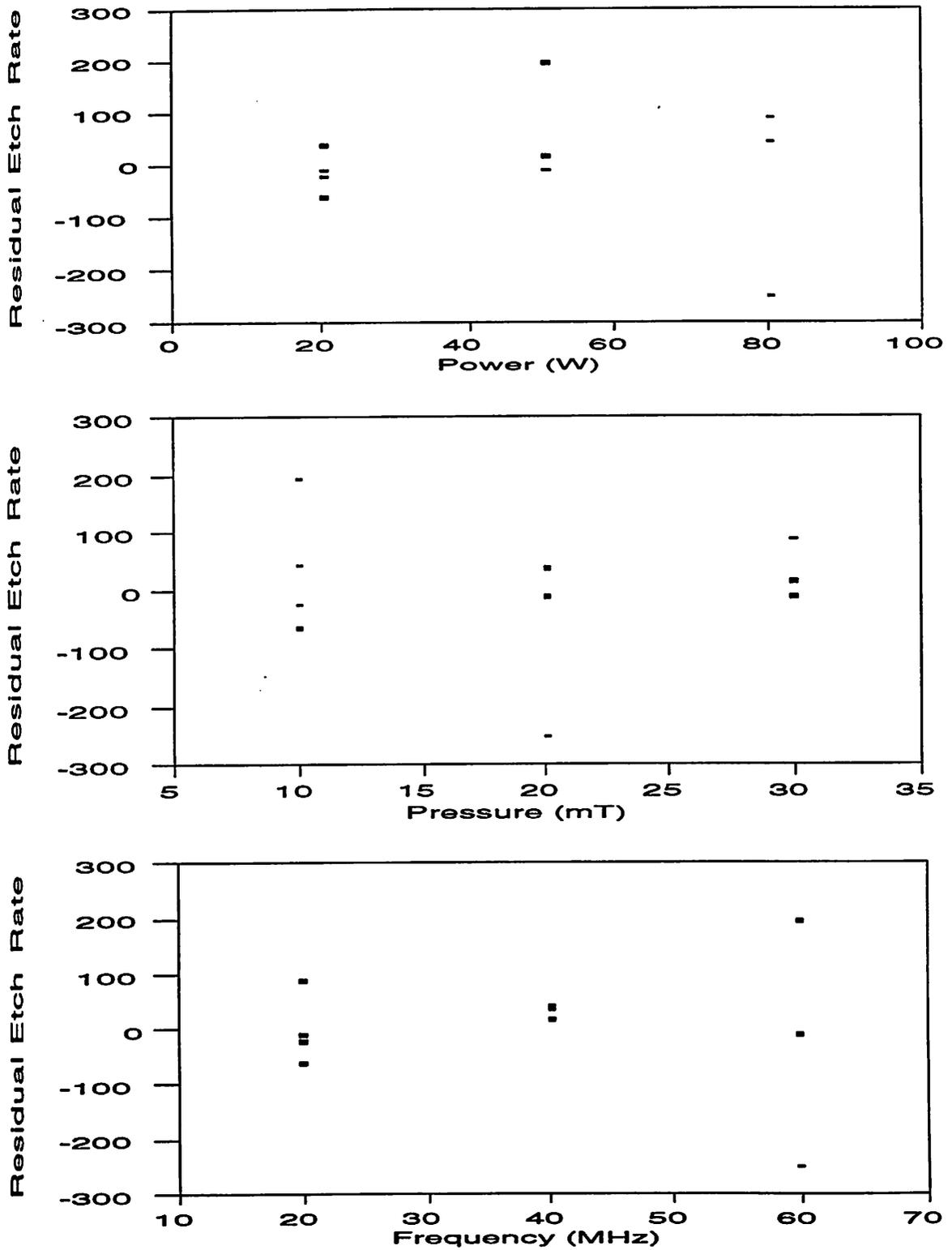
Figure 3. Residuals of revised linear model without flow term.
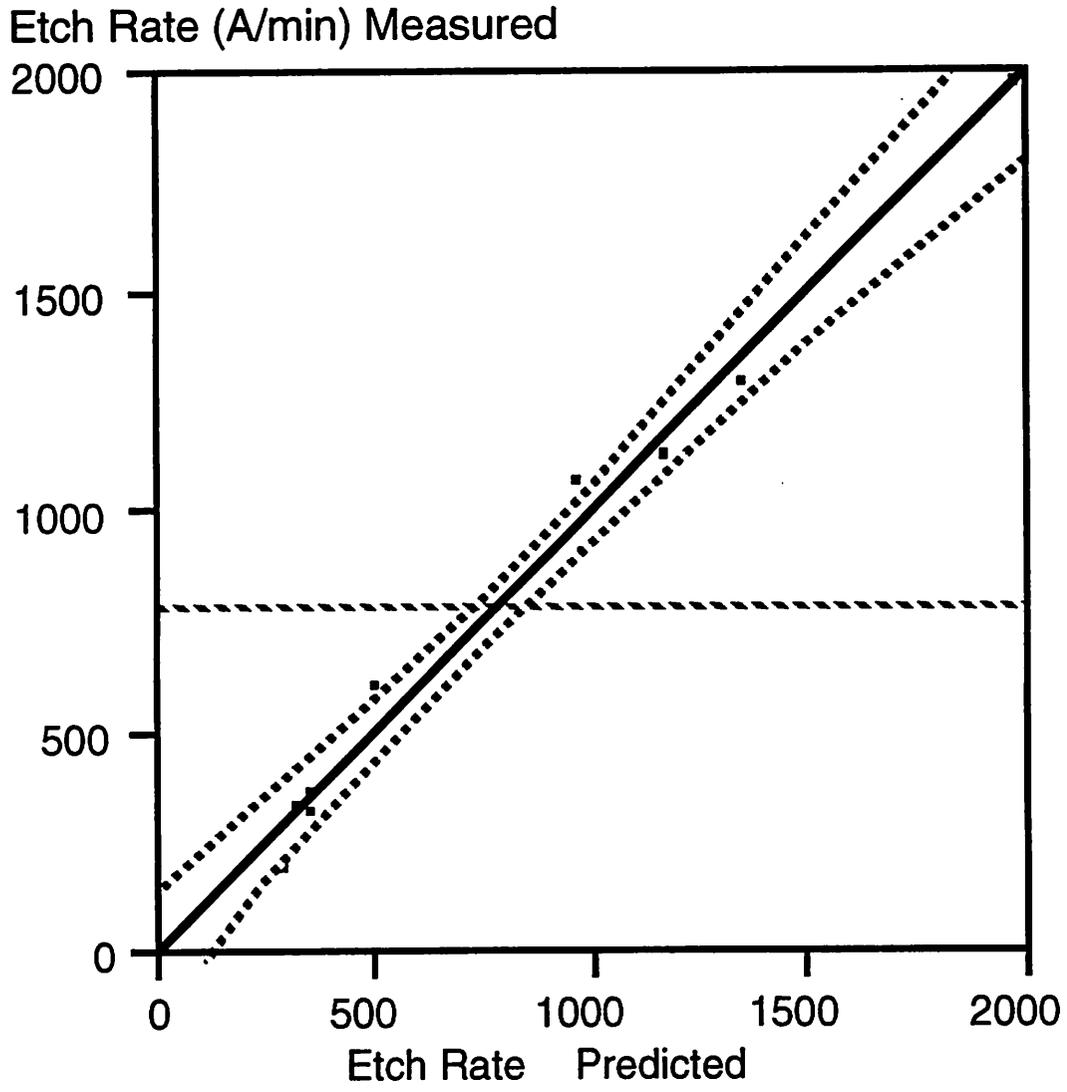
Figure 4. Resist etch rate measured versus predicted for linear model based on power, pressure and cross term of frequency and power.