MODELING AND CHARACTERIZATION OF SEMICONDUCTOR MANUFACTURING EQUIPMENT: AN APPLICATION TO LPCVD REACTORS

by

Kuang-Kuo Lin

Memorandum No. UCB/ERL M90/44

25 May 1990

MODELING AND CHARACTERIZATION OF SEMICONDUCTOR MANUFACTURING EQUIPMENT: AN APPLICATION TO LPCVD REACTORS

by

Kuang-Kuo Lin

Memorandum No. UCB/ERL M90/44

25 May 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

MODELING AND CHARACTERIZATION OF SEMICONDUCTOR MANUFACTURING EQUIPMENT: AN APPLICATION TO LPCVD REACTORS

by

Kuang-Kuo Lin

Memorandum No. UCB/ERL M90/44

25 May 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

MODELING AND CHARACTERIZATION OF SEMICONDUCTOR MANUFACTURING EQUIPMENT: AN APPLICATION TO LPCVD REACTORS

Ph.D. Kuang-Kuo Lin EECS Department

ABSTRACT

A systematic equipment modeling and characterization methodology has been developed for automated VLSI manufacturing. This methodology is based on the development of generic first principle process models. These generic models are subsequently refined and fitted to specific manufacturing equipment, using a multi-stage D-optimal experimental design. The methodology has been successfully applied to a low pressure chemical vapor deposition (LPCVD) furnace for undoped polysilicon deposition. A 2-stage D-optimal experiment with 24 runs has yielded statistical fitted models for the film growth rate and film residual stress. The calibrated models agree well with the experimental data, and account for the observed process variations.

Achieving consistent and high quality operation for each VLSI manufacturing process step can be challenging. The task of optimizing the process yield can be simplified through the application of statistical semi-empirical equipment models for automated process design and control. In order to accommodate the multiple objectives of semiconductor manufacturing, a numerical optimizer is integrated with a highly interactive interface. The combination can assist the process engineer in choosing the best compromise of equipment settings in terms of product and equipment performance. Currently, the statistical LPCVD models we developed are being used in a computer-aided design system that synthesizes optimal manufacturing process steps for the LPCVD of undoped polysilicon. The system will generate recipes that achieve objectives not just related to the average value of the film properties, but to their uniformity as well.

ACKNOWLEDGEMENT

It is with deep gratitude that I wish to acknowledge the invaluable guidance of Professor Costas Spanos and Professor David Hodges throughout the course of my graduate studies. They have provided tremendous motivation and encouragement in both my academic and non-academic pursuits. This dissertation would not have been accomplished without their knowledge, expertise and patience.

I sincerely thank Prof. Sadashiv Adiga for serving on my dissertation committee, as well as advising me in various aspects of this research. I would also like to thank Dr. Chi-Yung Fu, Prof. Larry Rowe, Prof. Roger Howe and Prof. Nathan Cheung for their many helpful technical insights.

I wish to express my heartfelt appreciation for the hard work and contributions of Jiahua Huang in her involvement with the experimental areas of this research, and also of Chung Liu for his efforts and assistance in the development of the recipe generator.

It has been a great pleasure working with my colleagues in my research group. They have greatly enhanced my learning experience. It is not possible to list all their names here. Nevertheless, I would especially like to thank Norman Chang, Sherry Lee, Gary May, Haifang Guo, Chris Hegarty, Eric Boskin, Chris Williams and Dr. Zhi-Min Ling. My thanks also go to Carl Galewski, Gino Addiego and Kenny Toh of the IC processing and simulation group, Jyuo-Min Shyu of the CAD group and Chuen-Chien Lee of the fuzzy control group for the many useful discussions we had. The help received from the PICASSO user interface group, as well as from the staff of the Berkeley Microfabrication Laboratory, is also much appreciated. Thanks should also go to Boon-Khim Liew for his advice on typesetting the figures of this dissertation.

Thanks to Ywh-Pyng Harn, Weijie Yun and Gani Jusuf for adding much joy and fun during this entire period. Last, but not least, I thank my family for their love and support through the years. The sacrifice made on their part in order to provide me with

excellent education is greatly appreciated.

The work is jointly sponsored by the Semiconductor Research Corporation (SRC), MICRO, Harris Semiconductor, IBM, Intel, National Semiconductor, Philips/Signetics, Rockwell International, Siemens AG, Texas Instruments, and National Science Foundation under Grant No. MIP 8715557.

Table of Contents

CHAPTER 1
INTRODUCTION

1.1. Motivation

In light of today's stiff economical competition and rapid technological advancement in the semiconductor industry (Table 1), manufacturing science is recognized for its leverage on productivity and quality. The design of a Computer-Integrated-Manufacturing (CIM) system architecture for manipulating critical manufacturing data sets [1], as well as the development of efficient methodologies for process design and control, can offer a competitive edge for the industry.

TRENDS IN VLSI WAFER FABRICATION			
	LSI	VLSI	ULSI
	(1975)	(1985)	(1990)
Products (DRAM)	16K	256-1M	4M
Throughput, wafer/mo.	10K	30K	50K
Total process steps/lot	100	230-400	550
No. of equipment types	40	100	120
Total equipment count	70	300	400
No. of process conditions	200	800	1,500
Database records/lot for stable production	100	5,000	10,000

Source: K. Sato, Toshiba
(JTECH/NSF visit to Japan, 1988)

Table 1.1 Trends in VLSI wafer fabrication

Accurate and efficient process/equipment models [2,3], if coupled with the capabilities of a well-designed CIM architecture, can be instrumental in supporting many crucial manufacturing functions. The subject of this dissertation is a systematic and economical methodology towards building and calibrating *equipment-specific process models* for manufacturing applications. We choose to call these models "*process models*" because they actually describe the physical and chemical phenomena in the processing chamber of the equipment. The fact that these models must be characterized through measurements for specific pieces of equipment makes them "*equipment-specific*". In comparison to purely physically-based models and to strictly empirical models, this approach is probably more suitable for process design and control in a semiconductor manufacturing environment, in terms of its simplicity of computation and capability of extrapolation.

The methodology has been successfully applied for the modeling of a low pressure chemical vapor deposition (LPCVD) furnace for undoped polysilicon. A 2-stage D-Optimal experimental design with 24 runs was used to calibrate and refine physically-derived LPCVD models for the specific furnace. These models describe the equipment not only for the nominal process responses, such as the *film deposition rate* and *residual stress*, but also for the associated process uniformities. The models agree well with the experimental data, and account correctly for the observed process variations.

Currently, the LPCVD models have been integrated in the Berkeley Computer-Aided-Manufacturing (BCAM) system for various applications, such as process design [4], control, and diagnosis [5]. One application utilizing the developed models and described in length herein, is a computer-aided process design system. In order to accommodate the multiple objectives of semiconductor manufacturing, a numerical optimizer is integrated with a highly interactive interface. The combination can assist

the process engineer in choosing the best compromise of equipment settings in terms of product and equipment performance. The computer-aided design system is currently being used to synthesize optimal manufacturing process steps for the LPCVD of undoped polysilicon. The system will generate recipes that achieve objectives not just related to the average values of the film properties, but to their product uniformity and process stability as well.

1.2. Dissertation Outline

Chapter 2 provides a conceptual overview of the proposed architectures of the Berkeley Computer-Integrated/Aided-Manufacturing (BCIM/BCAM) systems. The development of semi-empirical models, for both the nominal process behavior and the associated process variations, is described in Chapter 3. The design of experiments for model characterization is introduced in Chapter 4. Chapter 5 depicts the statistical characterization process using data derived from a designed experiment. The evaluation of statistical semi-empirical models is given in Chapter 6. Chapter 7 is devoted to a discussion on automating optimal process design using the developed models. Finally, the dissertation concludes with ideas for future research in Chapter 8.

REFERENCES

[1] D. Hodges, L. Rowe, and C. Spanos, "Quality and Productivity in Semiconductor Manufacturing", *Research Summary 1989*, University of California, Berkeley, College of Engineering, EECS/ERL, 1989.

[2] R. Cavin, W. Miller, "Design for Manufacturing", *Semiconductor Research Corporation (SRC) Newsletter*, March 1989.

[3] E. Sachs, R-S. Guo, S. Ha, and A. K. Hu, "Process Control System for VLSI Fabrication", *Proceedings of the Fifth Symposium on Automated Integrated Circuits Manufacturing*, Proceeding Volume 90-3, pp. 78-93, The Electrochemical Society, Inc., 1990.

[4] K-K Lin, C. Y. Fu, Norman H. Chang and D. A. Hodges, "Recipe Generator for LPCVD Deposition," Symposium on Automated Semiconductor Manufacturing, 172nd Electrochemical Society Meeting, Honolulu, Hawaii, October 1987, *Extended Abstracts, Volume 87-2*, abstract no. 653, pp. 919-920.

[5] N. Chang, C. Spanos, "Chronological Equipment Diagnosis with Evidence Integration", Presented at the SPIE Conference on Applications of Artificial Intelligence VIII, 1990.

CHAPTER 2
THE BERKELEY COMPUTER INTEGRATED MANUFACTURING FRAMEWORK

2.1. Introduction

The objective of the Berkeley Computer Integrated Manufacturing (BCIM) project is to improve the productivity and quality of IC manufacturing. This objective is being achieved by the development of both algorithmic methodologies and software modules for controlling VLSI processing steps. Crucial to this approach is the design of a flexible architectural model for linking different software modules into a computer integrated system to support design, manufacturing and testing. This chapter gives a conceptual introduction to BCIM and its associated modules, where a set of distinct new approaches to IC manufacturing are proposed.

2.2. The Berkeley CIM Architecture

The success of a CIM system depends on the degree of integration of knowledge and software across disciplines such as business, engineering, operations research, computer science, management information systems (MIS), etc. Unfortunately, many of the important pieces of knowledge and software in today's manufacturing environment are "stand-alone" entities. This lack of integration often delays corporate decisions that depend on the information gathered from various levels and disciplines in the corporation.

A cost-effective model of the system architecture, shown in Fig. 2.1, is proposed to support IC-CIM applications [1].

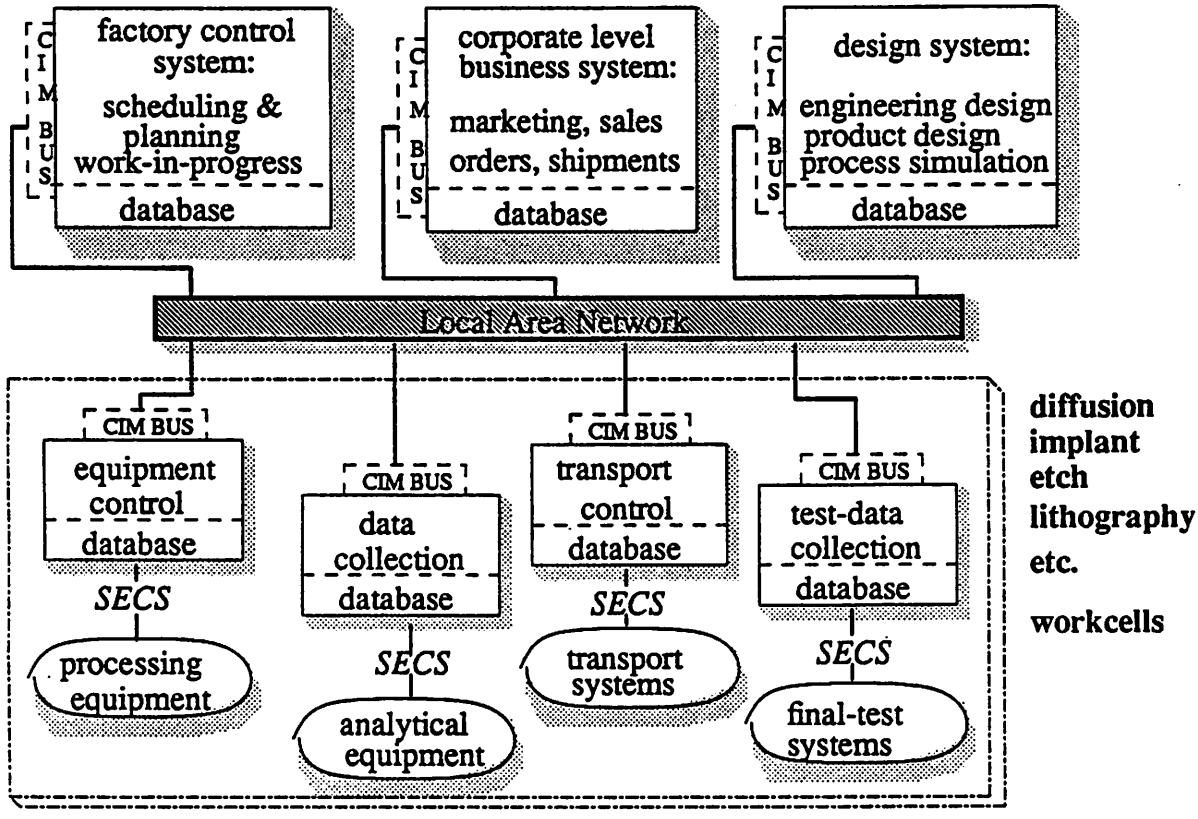


Fig. 2.1 The BCIM architecture.

Special features and requirements of employing this architectural model include:

(1) A multi-tasking operating system, such as UNIX.

(2) A high speed local area network (LAN) for linking different computing systems. The LAN makes it possible to link process control applications in different systems directly to each other and to the fabrication equipment.

(3) A physically distributed but logically integrated relational database.

(4) A two-level partition of the system. The lowest level in this architecture, which is physically closest to the manufacturing floor, includes the embedded controllers that

provide real-time control of the semiconductor analysis and fabrication equipment. Personal computers used as equipment controllers or for other tasks are included in this first level.

The second level comprises a distributed network of multi-tasking computer workstations, file servers, and computing processors linked by a common distributed relational Data Base Management System (DBMS). Different computing processors are linked by a high-speed LAN at this level. Some computer workstations, generally known as work-cell controllers, will be located in specific fabrication areas for supervising a given equipment cluster. The functions of the work-cell controllers include collection of real-time and in-line process data, statistical process analysis (SPC), and graphical display interfaces to equipment operators. The work-cell controllers communicate with the embedded equipment controllers via the SECS (Semiconductor Equipment Communication Standard) or compatible protocols to facilitate vital data collection. Due to their multi-tasking capability, the work-cell controllers can support multiple concurrent operations within the work-cell.

Other multi-tasking workstations at the second logical level (but physically further from the manufacturing floor) may be devoted to general logistic operations, such as engineering design systems, corporate-level business systems, factory control systems, etc.

(5) A distributed relational database, coupled with the definition of a common data interchange format (such as CIMBUS [2]), enables a user to issue the same operating commands to enter or query data in the database at any location or level in the architecture. Similarly, different pieces of software can communicate with each other at any location or level in the architecture.

This simple architectural model has great flexibility, which enables future extension and adaptation to meet changing requirements, such as the addition of more computing resources. The physical location of database files and the specific computer on which application software is executed can be easily changed as needed.

2.3. The Berkeley CAM Architecture for a Work-cell Controller

A Berkeley CAM architecture [3], shown in Fig. 2.2, is proposed for a work-cell controller suitable for supervising the operation of semiconductor manufacturing equipment.

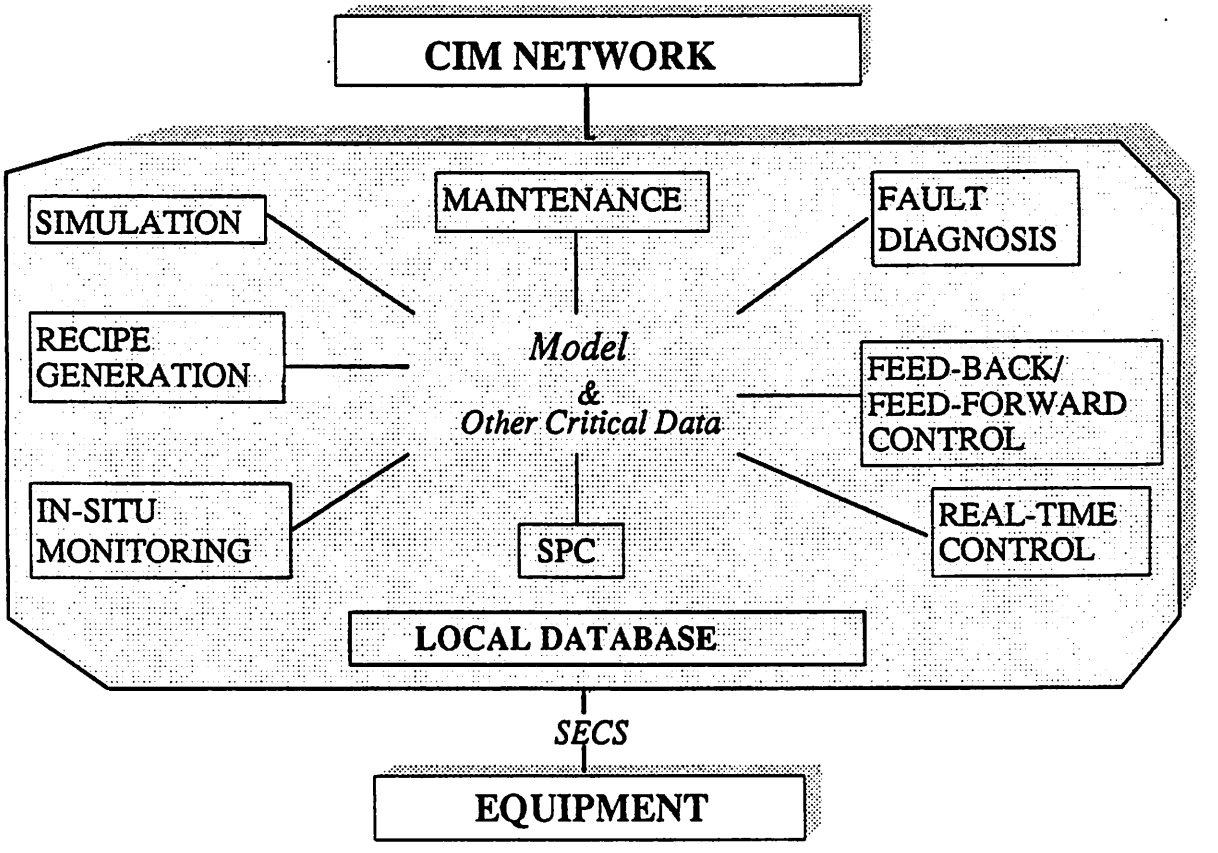


Fig. 2.2 The BCAM architecture.

The six data sets, listed below, are identified as being critical to the operation of a work-cell controller [1]:

(i) production lot history data

(ii) real-time process monitoring data

(iii) in-line physical and electrical measurements

(iv) final electrical test data

(v) equipment maintenance data

(vi) equipment-specific process models

There are several useful work-cell controller capabilities. They include: *equipment simulation, recipe generation and management, real-time monitoring, fault diagnosis, statistical process control (SPC), preventive maintenance, real-time control, and feedback/feed-forward control.* These capabilities are implemented using techniques from statistical experimental design, statistical process control (SPC), numerical optimization, artificial intelligence [4], etc. These capabilities are tightly coupled; for example, SPC can be used in the presence of noise to trigger an on-line diagnostic procedure, which in turn will use an equipment model to render its diagnosis.

The software structure of BCAM can be described with an object-oriented paradigm [5], shown in Fig. 2.3, to facilitate implementation efficiency, such as inheritance of generic capabilities (such as the ones described above) for different pieces of equipment, and modularity for software maintenance.

A model library [6], containing the necessary equipment models for the work-cell controller applications, also follows the object-oriented paradigm (Fig. 2.4).

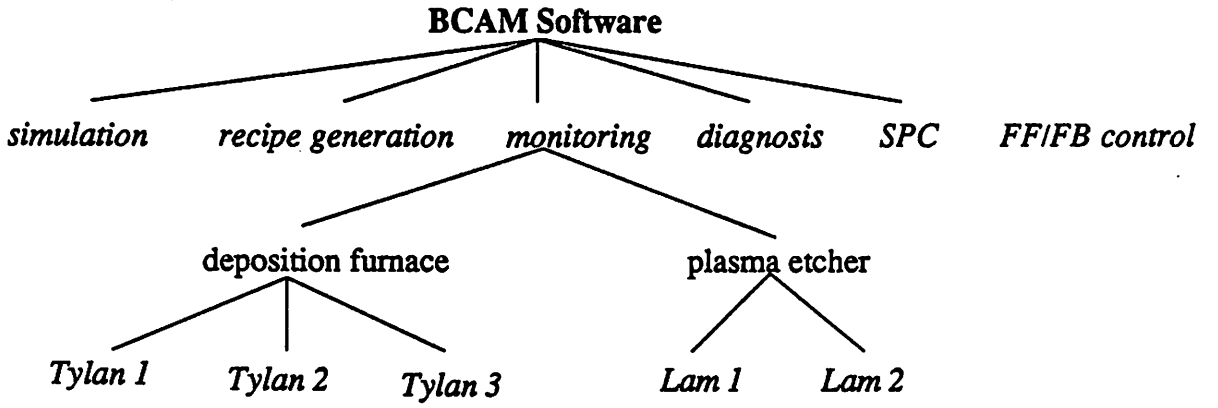


Fig. 2.3 The BCAM software structure.

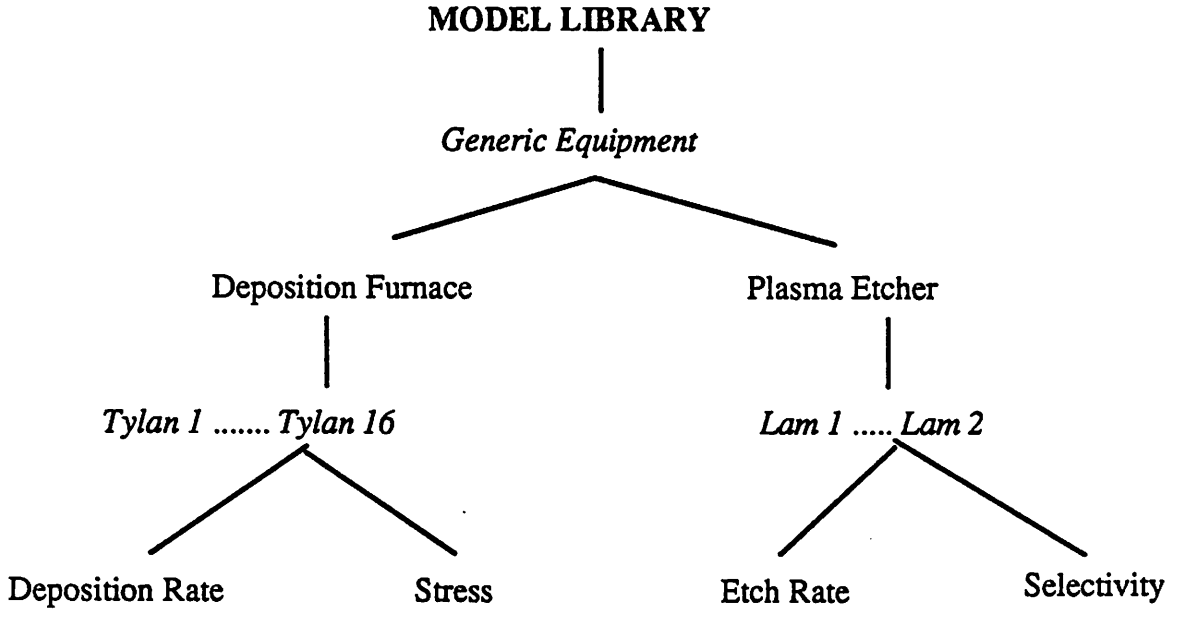


Fig. 2.4 The BCAM model library.

The implementation platform for the BCAM work-cell controller includes high resolution bit-mapped display workstations based on X window primitives; high-level programming languages such as C, C++, Common Lisp, CLOS (a Common Lisp Object System), and the INGRES relational database. Other commercial packages, such as RS/1 (a statistical analysis environment) and NAG (a numerical analysis library), are also used within the BCAM work-cell controller system.

2.4. Equipment-Specific Process Models for Manufacturing Applications

Accurate process models are indispensable for supporting the various functions of a work-cell controller. In order to support a wide range of manufacturing applications, including real-time computation, the manufacturing-based process models must meet certain criteria:

(i) The process models should be characterized for each individual piece of equipment for accurate representation. Henceforth, we describe them as *equipment-specific process models*.

(ii) The models should be simple and analytical for two reasons. First, they must be easy to use and modify. Second, they must be computationally efficient.

(iii) The models should capture both the nominal process behavior (such as the process average of a wafer lot) and the process variation (such as the process uniformity of a wafer lot).

Given these criteria, a semi-empirical approach is chosen to construct the *equipment-specific process model.* A semi-empirical model is physically derived from simplified physical principles, but is statistically characterized for a piece of specific equipment by a designed experiment. The semi-empirical approach combines the benefits of both the first-principle models, which allow for extrapolation and prediction, and those of the

empirical models [7], which are simple and statistically significant.

The remainder of the dissertation is devoted to the construction of semi-empirical *equipment-specific process models*, with LPCVD of undoped polysilicon as the research vehicle of the modeling and characterization methodology. A sample application of the developed models for recipe generation will be demonstrated at the end of the dissertation.

REFERENCES

[1] D. Hodges, L. Rowe, and C. Spanos, "Quality and Productivity in Semiconductor Manufacturing", *Research Summary 1989*, University of California, Berkeley, College of Engineering, EECS/ERL, 1989.

[2] D. Hodges, *Private communication*, University of California, Berkeley, 1990.

[3] C. J. Spanos, "The Berkeley Computer-Aided Manufacturing System", *Research Summary*, pp. 93-94, University of California, Berkeley, College of Engineering, EECS/ERL, 1990.

[4] C. Y. Fu, N. Chang, K.K. Lin, "Smart Integrated Circuit Processing", *IEEE Transaction on Semiconductor Manufacturing*, No. 4, pp. 151-158, Nov. 1989.

[5] Sonya E. Keene, "Object-Oriented Programming in Common Lisp: A Programmer's Guide to CLOS", Addison-Wesley Inc., 1988.

[6] Gary May, *private communication*, January 1990.

[7] Fariborz Nadi, "Modeling Complex Manufacturing Process via Integration of Influence Diagrams and Neural Networks", *Memorandum No. UCB/ERL M89/123*, University of California, Berkeley, November 1989.

CHAPTER 3
THE LPCVD MODEL DEVELOPMENT

3.1. Introduction

The objective of this investigation is to develop simple, yet precise equipment models that can be used in a semiconductor manufacturing environment. The LPCVD of undoped polysilicon has been selected as a test vehicle for this investigation due to its importance in many IC applications. Although the deposition of undoped polysilicon is better understood than other LPCVD processes, it is still difficult to achieve good wafer uniformity by this batch-mode deposition process. It is expected that a model can help improve this situation.

3.2. Process Description

Fig. 3.1 depicts the schematic of the *Tylan* furnace system for the LPCVD of undoped polysilicon. This process deposits thin polycrystalline silicon film on the wafers. Wafers equally spaced are placed inside quartz boats with porous covers to ensure uniform gas flow into the boats. The boats are supported by an alumina cantilever rod extending into the furnace. Pure silane SiH_4 (with no other carrier gas) is injected from the load end (front end) of the furnace, and pumped out at the exhaust end (rear end). There are three heating elements, each controlled by a thermocouple[1]. Together, they provide the required process temperature for the load (front), center, and exhaust (rear) zones. The desired pressure is attained by a mechanical vacuum pump controlled by a pressure sensor located at the load zone. A disposable quartz liner is inserted within the furnace to protect it from undesired accumulation of polysilicon.

[1] A thermocouple is placed close to the external heating elements for optimal response of the furnace temperature control. The drawback of this configuration is the existence of a temperature offset between the wafers and the thermocouple.

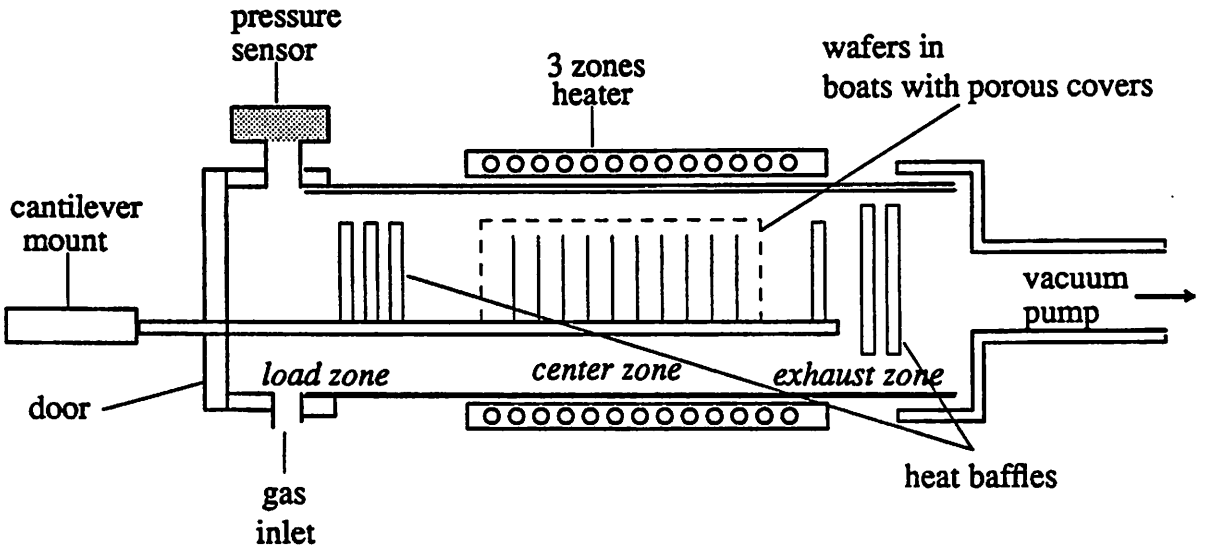


Fig. 3.1 The Tylan LPCVD furnace

There are many process design parameters affecting the performance of this processing step. They are listed in Tables 3.1 and 3.2 in order of importance [1-14].

Primary LPCVD Process Design Parameters
(1) The 3-zone furnace temperatures.
(2) The furnace pressure.
(3) The front-end injected silane (SiH_4) flow rate.
(4) The deposition time.
(5) The furnace geometry.
(6) The wafer spacing (or the furnace load factor).
(7) The history of the furnace (amount of accumulated polysilicon).
(8) The history of the wafer (existing films and underlying structures)
(9) The wafer cleaning procedure.

Table 3.1 Process design parameters for LPCVD of undoped polysilicon

LPCVD Process Responses
(1) The deposition rate (or thickness for a fixed deposition time).
(2) The built-in residual polysilicon film stress.
(3) The etch properties of the deposited film.
(4) The film smoothness.
(5) The film grain size.
(6) The film texture (crystalline orientation).
(7) The step coverage properties of the film deposited on complex features.
(8) The refractive index of the film.

Table 3.2 Process responses of LPCVD undoped polysilicon

3.3. Practical Limitations on Process Parameter Settings

Not all the parameters listed in Table 3.1 are controllable or easily adjustable. The only easily controllable parameters are: *temperature, pressure, injected silane flow rate,* and *deposition time.* The settings of these primary parameters are traditionally dictated by the practical considerations listed below [1-14]:

Parameter settings				
process objectives	*temperature*	*pressure*	*silane*	*time*
high throughput	high	high	high	
surface–reaction–limited[2]	low	low		
wafer-to-wafer uniformity	low		high	
within-wafer uniformity		low	high	
smooth film	low	low		
small grain-size	low	high	low	low

Table 3.3 Practical settings of the primary process parameters.

[2] As the name implies, a surface-reaction-limited process is controlled by the rate at which the reactants can react at the wafer surface. A *surface-reaction-limited* process generally produces more uniform results.

The ranges for some of these parameters are dictated by mechanical limitations.

Other parameters, such as furnace geometry, are not easily adjustable. Engineering judgement and experience must be exercised in finding the optimal settings of these secondary parameters. A reasonable design space, based on the above requirements for the primary parameters is shown in Fig. 3.2 (the boundaries depicted in Fig. 3.2 are set for the particular reactor that was used for the experimental part of this dissertation).

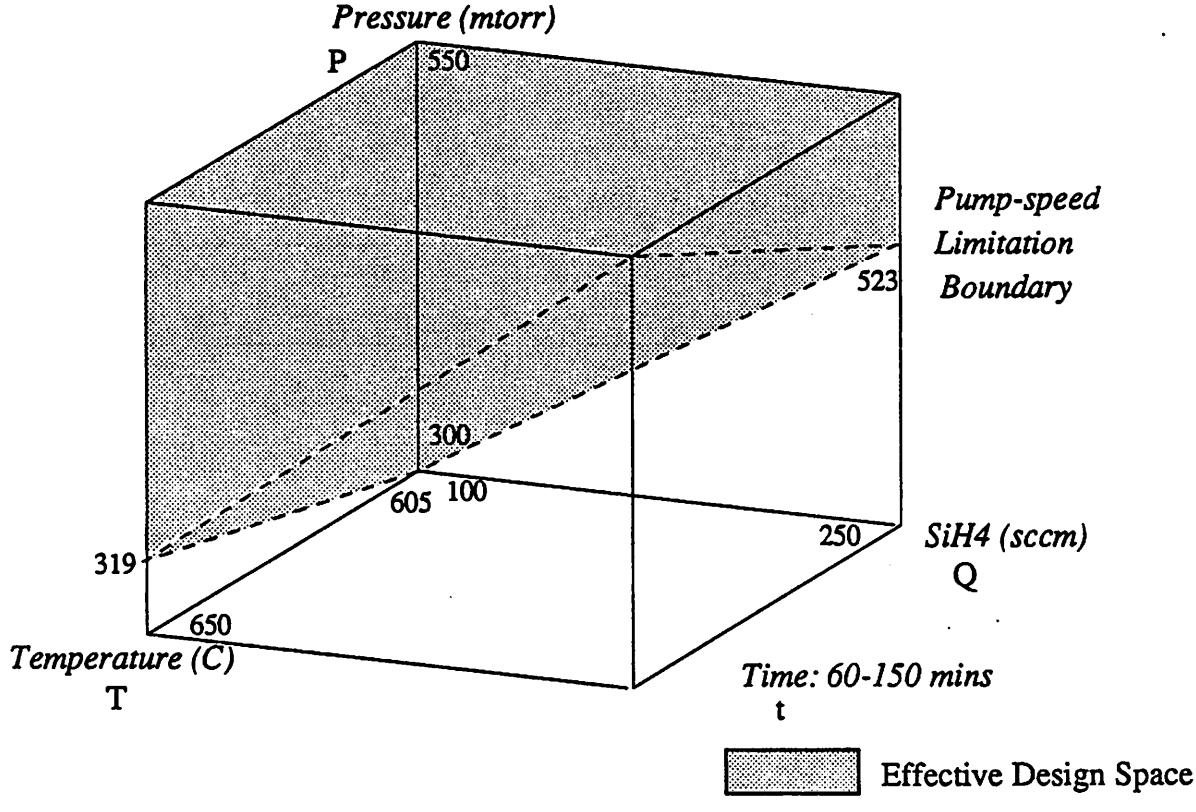


Fig. 3.2 The LPCVD design space

For the output responses, only the deposition rate and stress are easily measurable. Other responses can only be qualitatively evaluated and related to the primary process parameters.

3.4. The Development of a Quantitative Model

We intended to model both deposition rate and stress from first principles. However, the only well-understood response of the LPCVD undoped polysilicon process is the deposition rate [2]. Most of the deposition rate models in the literature are numerically complex and cannot be easily adapted to represent a specific item of manufacturing equipment [12,14]. Here, our goal is to develop a closed-form *equipment-specific process model*, in order to facilitate process characterization, design, and control. We will start with a model that relates the primary parameters to the deposition rate.

3.4.1. Model Assumptions

Given the chosen design space, the following assumptions are used to simplify the derivation of the deposition rate model [1-14]:

(i) The deposition is a surface-reaction-limited process such that any limitation due to diffusion tranport of the reactant can be neglected.

(ii) The temperature is constant and uniform[3] in the volume of the furnace.

(iii) Silane is transported by laminar flow along the furnace.

---

[3] The assumption of uniform temperature is very important since some polysilicon properties, such as the stress and film quality, are known to be greatly affected by non-uniform temperature profiles [1]. Although this assumption is needed for model simplicity, it can be relaxed for non-uniform temperature profiles [21].

3.4.2. The Initial Deposition Rate Model

Based on the above assumptions, a first order physically-based analytical model for the deposition rate kinetics has been developed [1,5,12]. This model yields the deposition rate in Å/min at a given furnace position z in *cm* from the pure silane SiH_4 inlet.

$$R(z) = A \exp(\frac{-\Delta E}{kT}) P_{SiH_4}(z) \tag{3.1}$$

where A is the Arrhenius frequency factor[4] in Å/*min-mtorr*, ΔE is the activation energy in *kcal/mole*, k is the universal gas constant given as 1.98719 *cal/mole*-K, T is the temperature in K, and $P_{SiH_4}(z)$ is the partial silane pressure in *mtorr*. Eq. (3.1) can be expressed in terms of the total system pressure P:

$$R(z) = A \exp(\frac{-\Delta E}{kT}) C(z) P \tag{3.2}$$

where C(z) is the mole fraction of silane at position z, and P is the total furnace pressure in *mtorr*. We now define a new term $\eta(z)$ to represent the mole fraction of silane already consumed at position z. The following integral represents the ratio of reacted versus injected silane SiH_4 between the first position zero and z.

$$\eta(z) = \int_0^z \frac{s(z)\, R(z)}{Q} C_{gs}\, dz;\ \eta(0) = 0 \text{ at } z=0 \tag{3.3}$$

where $s(z)$ is the effective deposition area per unit length in *cm* (which includes the wafer surfaces, boat surfaces and the furnace wall), Q is the total front-end injected silane flow in *sccm*, and C_{gs} is the "gas to solid" conversion factor for silicon (1.85×10^{-5} *cm*/Å) . We can simplify Eq. (3.3) given an effective deposition surface

[4] For a discussion on the *Arrhenius frequency factor* refer to Appendix 3 of [1].

area[5] of S cm^2 and a furnace length of L *cm*, between position zero and z.

$$\eta(z) = \int_0^z \frac{S\,R(z)}{LQ} C_{gs}\, dz \tag{3.4}$$

The surface reaction of polysilicon deposition can be summarized by the pyrolysis of silane [1,2].

$$SiH_4 \rightarrow Si + 2H_2 \tag{3.5}$$

From this we see that every $\eta(z)$ mole of consumed SiH_4 will release $2\eta(z)$ moles of H_2. Hence, C(z) is related to $\eta(z)$ by:

$$C(z) = \frac{\text{Mole fraction of remaining } SiH_4 \text{ at z}}{(\text{Mole fraction of remaining } SiH_4 \text{ at z}) + (\text{mole fraction of released } H_2)}$$

or, equivalently

$$C(z) = \frac{1-\eta(z)}{[1-\eta(z)] + 2\,\eta(z)} = \frac{1-\eta(z)}{1+\eta(z)} \tag{3.6}$$

Substituting this relationship back into the deposition rate model of Eq. (3.2), we obtain:

$$R(z) = A \exp(\frac{-\Delta E}{kT})\, P\, \frac{1-\eta(z)}{1+\eta(z)} \tag{3.7}$$

3.4.3. Translation of the Coordinate System

The equations derived so far are based on the assumption that the furnace is operating at full capacity, with the first wafer deposition starting at 'z=0' (the very front end of the furnace). However, our furnace is operating at half capacity, with the first

[5] In general, the effective deposition surface area is a function of temperature (T). S is approximated with a constant value within the experimental temperature range of 605 to 650 °C.

wafer positioned somewhere near the center of the furnace (Fig. 3.3). To accommodate this arrangement, we introduce a new coordinate system " z' ". The new coordinate system is related to the old by a simple translation by z_0:

$$z' = z - z_0 \tag{3.8}$$

where z_0 is the distance of the first wafer from the front end of the furnace.

To account for the small amount of silane η_0 depleted up to z_0, we also have:

$$\eta' = \eta - \eta_0, \quad \eta_0 = \eta(z_0) \tag{3.9}$$

where η' is the additional silane depletion downstream from z_0.

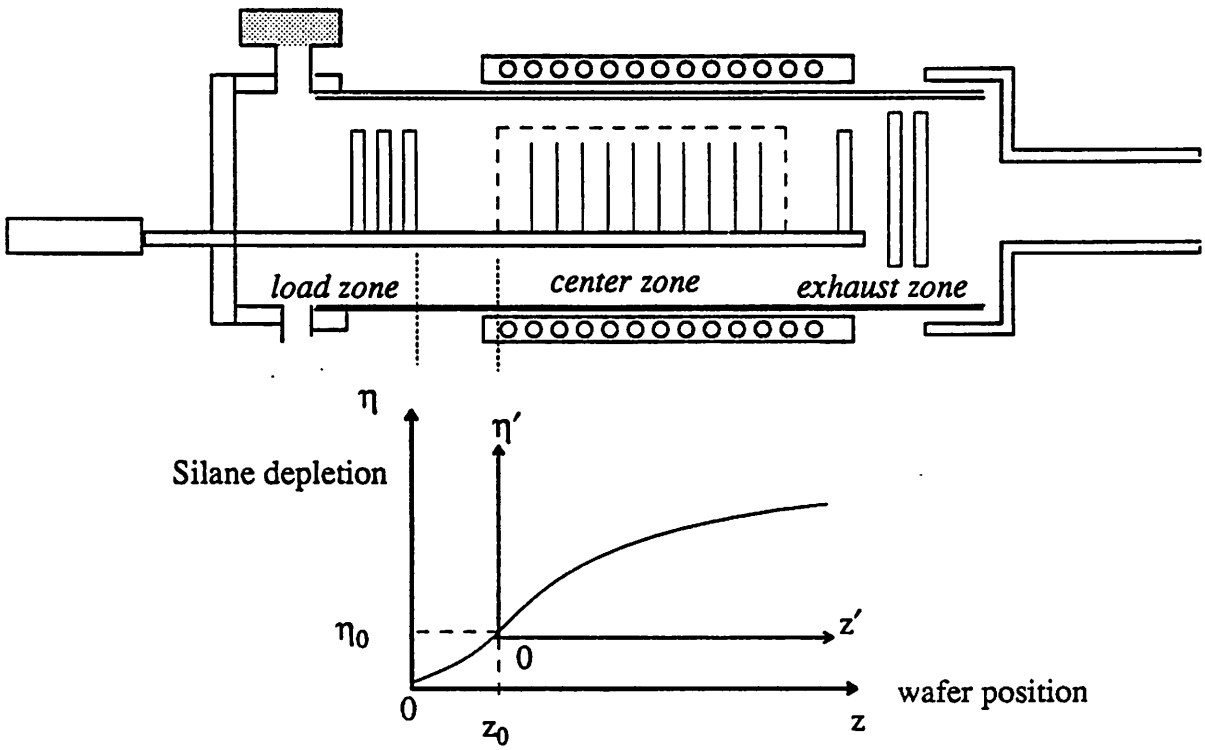


Fig. 3.3 Translation of the furnace coordinate system.

Substituting the above relationships back Eq. (3.7), we have:

$$R(z) = A \exp(\frac{-\Delta E}{kT})\, P\, \frac{1-\eta_0}{1+\eta_0} \left[\frac{1 - \frac{\eta'(z)}{1-\eta_0}}{1 + \frac{\eta'(z)}{1+\eta_0}} \right] \tag{3.10}$$

Defining R_0 as the deposition rate at position z_0, we have:

$$R_0 = A \exp(\frac{-\Delta E}{kT})\, P\, \frac{1-\eta_0}{1+\eta_0}$$

$$R(z) = R_0 \left[\frac{1 - \frac{\eta'(z)}{1-\eta_0}}{1 + \frac{\eta'(z)}{1+\eta_0}} \right] \tag{3.11}$$

The portion in Eq. (3.11) involving η' can now be expressed in terms of z' and the process parameters P, Q, and T. Recall that:

$$\eta(z) = \int_0^z \frac{S\, C_{gs}}{QL} R(z)\, dz \tag{3.12}$$

Substituting the formal definition of R(z) (Eq. (3.7)) into the integral gives:

$$\eta(z) = \int_0^z \frac{S\, C_{gs}}{QL}\, A\, P \exp(\frac{-\Delta E}{kT}) \left[\frac{1-\eta(z)}{1+\eta(z)} \right] dz$$

Denoting all terms outside the large bracket in the expression above as "ζ", and separating the variables:

$$\int_{\eta_0}^{\eta(z)} \left[\frac{1+\eta(z)}{1-\eta(z)} \right] d\eta = \int_{z_0}^{z} \zeta\, dz, \quad \eta_0 = \eta(z_0) \tag{3.13}$$

Using Eqs. (3.8) and (3.9) to change variables, we get:

$$\int_0^{\eta'} \left[\frac{(1+\eta_0)+\eta'}{(1-\eta_0)-\eta'} \right] d\eta' = \int_0^{z'} \zeta \, dz', \quad \eta'(0) = 0 \tag{3.14}$$

After algebraic manipulation of the integral in Eq. (3.14), we have:

$$-\eta' - 2 \ln \left[1 - \frac{\eta'}{1-\eta_0} \right] = \zeta z' \tag{3.15}$$

We keep only the first term from the expansion of the natural-logarithm, by assuming that η' is small.

$$-\eta' + 2 \frac{\eta'}{1-\eta_0} \approx \zeta z' \tag{3.16}$$

The approximation error introduced from Eq. (3.15) to (3.16) is about 0.5% when $\eta' = 0.1$, and up to 2% when $\eta' = 0.2$. For most cases (including our furnace), η' can be assumed to be less than 0.2. Under these conditions, Eq. (3.16) can now be manipulated to yield η' as a *linear* function of z'.

$$\eta'(z') \approx \zeta z' \frac{(1-\eta_0)}{(1+\eta_0)}$$

The above equation can be further expressed in term of R_0, which in turn can be expressed in terms of the process parameters:

$$\eta'(z') \approx \frac{SC_{gs}}{QL} R_0 z' \tag{3.17}$$

3.4.4. The Final Deposition Rate Model

The final closed-form, first order deposition rate model is shown below:

$$R(z') = R_0 \left[\frac{1 - \frac{\eta'(z)}{1-\eta_0}}{1 + \frac{\eta'(z)}{1+\eta_0}} \right] \tag{3.18}$$

$$\eta'(z) \approx \frac{S'C_{gs}}{QL'} R_0 z' \tag{3.19}$$

(assuming $s(z) = S'/L'$ for $z_0 \le z \le L'$, and $\eta'(L') < 0.2$)

$$R_0 = A\, P \exp\left(\frac{-\Delta E}{kT}\right) \frac{1-\eta_0}{1+\eta_0} \tag{3.20}$$

$$\eta_0 = \eta(z_0) = \int_0^{z_0} \frac{S_0 C_{gs}}{L_0} Q\, R(z)\, dz \approx \frac{\lambda}{Q} \tag{3.21}$$

(assuming $s(z) = S_0/L_0$, $R(z) \approx R(0)$ for $0 \le z \le z_0$ and λ is some constant)

The models in Eqs. (3.18) to (3.21) are to be used as the physical basis for the creation of the semi-empirical *equipment-specific process model* (see Ch. 3 and 4).

At this point, unlike the deposition rate, the physical mechanism responsible for the built-in residual stress of the polysilicon film is not as well understood. Hence, we built an empirical polynomial *stress* model from the experiments (see Ch. 4).

3.5. The Development of Qualitative Models for LPCVD Operations

Extensive study on some critical process responses, based on the development of either first-principle or empirical models, is frequently limited by the lack of physical understanding and measurement data. To optimize these critical responses, process heuristics based on past experience are often used in the absence of precise quantitative models.

The heuristics, acquired from local experts or from published reports [2,6-11], are summarized in Table 3.4. Here we describe the qualitative dependency of *film roughness, grain size, refractive index* and *texture* on some of the primary process parameters. It is important to be able to reason about the process even though these process heuristics are in general approximate, ambiguous, and incomplete. To this end, we are now investigating a strategy that uses *fuzzy* principles to represent the acquired heuristics [15-20].

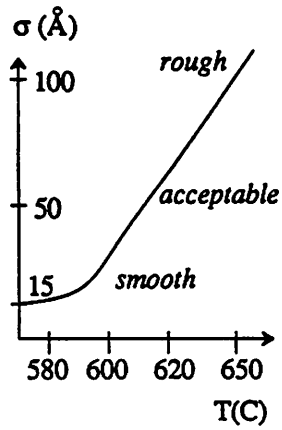
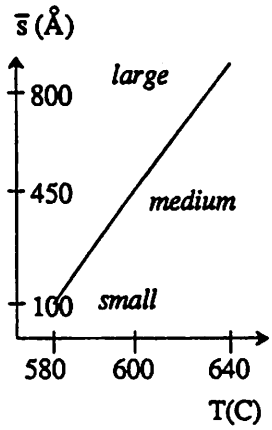
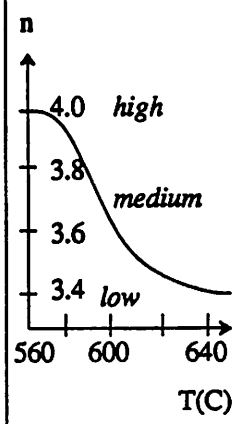
response / process settings	Surface Roughness (σ)	Grain Size ($\bar{s}$)	Refractive Index (n)	Film Texture
T (C)	 at P=? Q=? t=?	 at P=? Q=? t=?	 at P=? Q=? t=?	T(C) | Texture < 600 | <311> ≈ 620 | <110> > 650 | <100> at P=? Q=? t=?
P (mtorr)	P is high => σ is rough P increases => σ increases *and vice versa*	P(mtorr) / T(C) | 100 | 1000 580 | 600 | 390 Å 625 | 1160 | 920 Å 700 | 1360 | 1050 Å at Q=? t=?	? no information	T(C) / P(mtorr) | 600<T<625 P > 500 | <110> 10<P<100 | <100> P<0.2 | random T(C) / P(mtorr) | T≈580 P≈1000 | <311> 50<P<100 | <100> or <110> at Q=? t=?
Q (sccm)	? no information	Q is high => $\bar{S}$ is large Q increases => S increases *and vice versa*	? no information	? no information
t (min)	? no information	t is long => $\bar{S}$ is large t increases => S increases *and vice versa*	? no information	? no information

Table 3.4 Process heuristics for some process responses.

3.5.1. Knowledge Representation

The heuristics in Table 3.4 are first generalized as fuzzy rules in a computer system. For example, "surface roughness" can be qualitatively modeled as:

```
(T = low) => (surface = smooth)                                    (3.22)
(T = medium) => (surface = acceptable)
(T = high ) => (surface = rough)
(P = low) => (surface = smooth)
(P = medium) => (surface = acceptable)
(P = high) => (surface = rough)
```

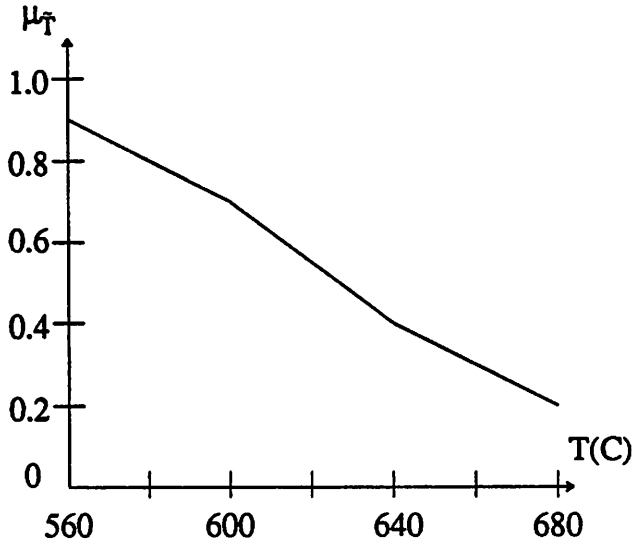
The premises and consequences of the rules are described by fuzzy concepts such as "T = low" and "Surface = smooth". A fuzzy concept can be described by a fuzzy set $\tilde{F}$ with its ordered-pairs defined as follows:

$$\tilde{F} = \left\{ (u, \mu_{\tilde{F}}(u)) \mid u \in U \right\} \tag{3.23}$$

Here, U is a set of objects in the universe of discourse, denoted generically by u. $\mu_{\tilde{F}}(u)$ is the *membership function* or *degree of truth* of u in $\tilde{F}$. The function maps U to the membership space M, whose normalized range is usually from 0 to 1. As an example, U can be the pressure (P), u can be a value of P such as 300 *mtorr*, and $\tilde{F}$ can be the fuzzy set to represent "P = low". The membership function will be used to calculate the degree of truth of the statement "a pressure of 300 *mtorr* is *low*". Other membership functions, such as $\tilde{T}$ for "T = low" and $\tilde{S}$ for "surface = rough", are shown in Table. 3.5 and 3.6. For the convenience of illustration, we use the tabulated format of the membership function.

T(C)	$\mu_{\tilde{T}}$
560	0.9
600	0.7
640	0.4
680	0.2

(a)



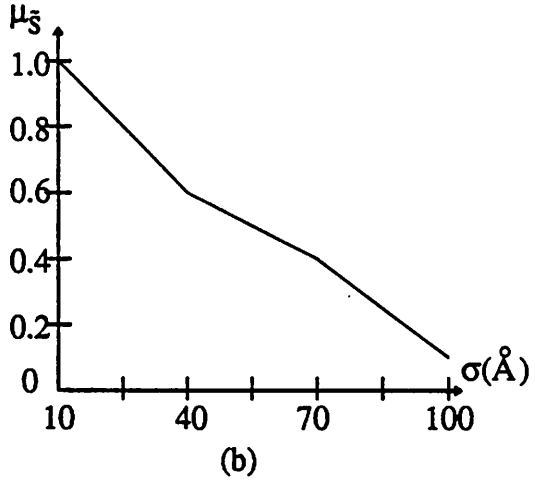
(b)

$\tilde{T}$ for "T = low"

Table 3.5 Membership function for the fuzzy set "T = low".

σ(Å)	$\mu_{\tilde{S}}$
10	1.0
40	0.6
70	0.4
100	0.1

(a)



(b)

$\tilde{S}$ for "Surface = smooth"

Table 3.6 Membership function for the fuzzy set "Surface = smooth".

3.5.2. Approximate Reasoning

The rules outlined in Eq. (3.22) can also be modeled by fuzzy relations to facilitate useful inference about the process responses, given the values of the input process parameters. For example, let $\tilde{T}$ and $\tilde{S}$ be the fuzzy sets representing the concepts "T = low" and "Surface = smooth" as appear in the first rule of Eq. (3.22). The fuzzy relation $\tilde{R}$ that relates $\tilde{T}$ and $\tilde{S}$ is represented as a matrix in Table 3.7. The elements of the matrix for $\tilde{R}$ can be computed by the fuzzy operator *Cartesian product*[5] ($\times$) on $\tilde{T}$ and $\tilde{S}$:

$$\mu_{\tilde{R}} = \mu_{(\tilde{T}\times\tilde{S})}(\tau,\sigma) = \min\left\{\mu_{\tilde{T}}(\tau)\,,\,\mu_{\tilde{S}}(\sigma)\right\} \tag{3.24}$$

Here τ and σ are specific values of temperature and surface roughness. With the definition of the fuzzy relation $\tilde{R}$ for the first rule, we are able to compute the fuzzy set $\tilde{C}$ for its conclusion that "surface = smooth" at a particular temperature τ. The fuzzy operator, *composition* (o), is used for the computation:

$$\text{composition: } \tilde{C} = \tilde{T} \text{ o } \tilde{R} \tag{3.25}$$

$$\mu_{\tilde{C}}(\sigma) = \max_{\tau}(\min(\mu_{\tilde{T}}(\tau)\,,\,\mu_{\tilde{R}}(\tau,\sigma)))$$

A simplified illustration of *approximate* reasoning is summarized below:

Step 1: Given a particular temperature τ = 600°C, compute $\mu_{\tilde{T}}$ (600°C) = 0.7 from Table 3.5.

Step 2: The grade of membership for $\tilde{T}$ of 600°C is then compared with the column of $\mu_{\tilde{R}}$ (Table 3.7) at 600°C. The minimum of the comparison yields the membership function for the fuzzy conclusion $\tilde{C}$. The membership function $\mu_{\tilde{C}}$ of

[6] In this case, the *Cartesian product* has the same definition as the *and* ($\cap$) operator, i.e. $\mu_{\tilde{T}\cap\tilde{S}} \triangleq \mu_{(\tilde{T}\times\tilde{S})}$.

$\tilde{C}$ is shown in Table 3.8.

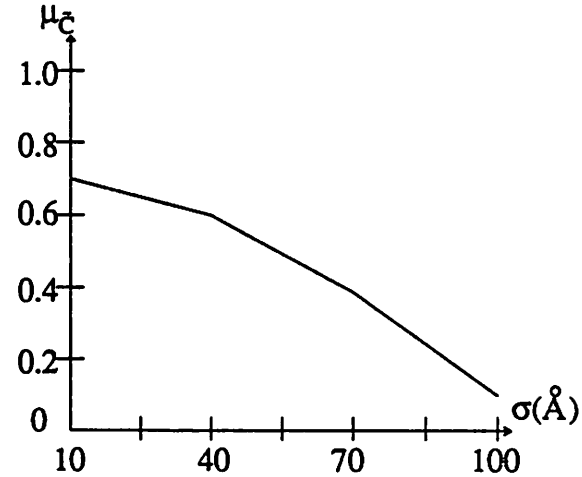
Step 3: The fuzzy set $\tilde{C}$ on surface quality resembles that of $\tilde{S}$, so the proposition "Surface = smooth" can be asserted.

		T°C			
		560	600	640	680
σ(Å)	10	0.9	0.7	0.4	0.2
	40	0.6	0.6	0.4	0.2
	70	0.4	0.4	0.4	0.2
	100	0.1	0.1	0.1	0.1

Table 3.7 Fuzzy relation $\mu_{\tilde{R}}$ for $\tilde{T}$ and $\tilde{S}$.

σ(Å)	$\mu_{\tilde{C}}$
10	0.7
40	0.6
70	0.4
100	0.1

(a)



(b)

$\tilde{C}$ for "Conclusion = Surface is smooth"

Table 3.8 Fuzzy set $\tilde{C}$ for the conclusion at 600°C.

3.5.3. Knowledge Integration

If we follow the approach described in the previous sections, many fuzzy sets representing conclusions of the rules (e.g. Eq. (3.22)) in the knowledge base will be generated. These fuzzy sets must be aggregated to provide a final and consistent solution. Fuzzy operators, such as the weighted or unweighted arithmetic or geometric means, can be used to integrate these fuzzy sets [18,19].

3.5.4. Summary for Qualitative Models

Methodologies employing *fuzzy set* and *fuzzy logic* are suitable for the qualitative modeling of certain process responses. The methodology makes it possible to incorporate approximate, ambiguous and incomplete process heuristics into a computer system.

REFERENCES

[1] S. Wolf, R.N. Tauber, "Silicon Processing; for the VLSI Era; Volume 1 - Process Technology", *Chapter 3 on Chemical Vapor Deposition of Amorphous and Polycrystalline Film*, Lattice Press, pp. 161-197.

[2] T. I. Kamins, "Polycrystalline Silicon for Integrated Circuit Applications", Kluwer Academic Publishers, 1988.

[3] R.J. Gieske, J.J. McMullen, "Low Pressure Chemical Vapor Deposition of Polysilicon". Materials and Molecular Research Division, Lawrence Berkeley Laboratory and Department of Chemical Engineering, University of California, Berkeley, 1977.

[4] W.A. Brown, T.I. Kamins, "An Analysis of LPCVD System Parameters for Polysilicon, Silicon Nitride and Silicon Dioxide Deposition", *Solid State Technology*, July 1979, pp. 51-57, 84.

[5] N. Cheung, "Thin Film Technology for IC Fabrication", EECS 290P course notes, University of California, Berkeley, Spring 1988.

[6] G. Harbeke, L. Krausbauer, E. F. Steigmeier, and A. E. Widmer, "LPCVD Polycrystalline Silicon: Growth and Physical Properties of In-Situ Phosphorus Doped and Undoped Films," *RCA Review* 44, 287-312, (June 1983).

[7] G. Harbeke, L. Krausbauer, E. F. Steigmeier, A. E. Widmer, H. F. Kappert, and G. Neugebauer, "Growth and Physical Properties of LPCVD Polycrystalline Silicon Films," *J. Electrochemical Society,* 131, 675-682 (March 1984).

[8] M. T. Duffy, J. T. Mcginn, J. M. Shaw, R. T. Smith, R. A. Soltis, and G. Harbeke, "LPCVD Polycrystalline Silicon: Growth and Physical Properties of Diffusion-Doped, Ion-Implanted, and Undoped Films," *RCA Review* 44, 313-325, (June 1983).

[9] P. Joubert, B. Loisel, Y. Chouan, and L. Haji, "The Effect of Low Pressure on the Structure of LPCVD Polycrystalline Silicon Films," *J. Electrochemical Society,* 134, 2541-2545 (October 1987).

[10] T. I. Kamins, "Structure and Properties of LPCVD Silicon Films", *J. Electrochem.. Soc: SOLID STATE SCIENCE AND TECHNOLOGY*, Vol. 127, No.3, pp. 686-690, March 1980.

[11] K.L. Chiang, et. al., "Optical Evaluation of Polycrystalline Silicon Surface Roughness", *J.Electrochem. Soc: SOLID-STATE SCIENCE AND TECHNOLOGY*, Vo. 126, No. 12, pp.2267-2269, December 1979.

[12] A. E. T. Kuiper, C. J. H. van den Brekel, J. de Groot, and G. W. Veltkamp, "Modeling of Low-Pressure CVD Processes," *J. Electrochem. Soc. SOLID STATE SCIENCE AND TECHNOLOGY*, Vol. 129, No. 10, pp. 2288-2291, October 1982.

[13] R. T. Howe and R. S. Muller, "Stress in Polycrystalline and Amorphous Silicon Thin Films," *J. Appl. Phys.* 54 (B) 4674-4675, (August 1983).

[14] S. Middleman and A. Yeckel, "A Model of the Effects of Diffusion and Convection on the Rate and Uniformity of Deposition in a CVD Reactor", *J. Electrochem. Soc.: SOLID STATE SCIENCE AND TECHNOLOGY*, Vol. 133, No. 9, pp. 1951-1956, September 1986.

[15] L. Zadeh, "Knowledge Representation in Fuzzy Logic", *IEEE Trans. on Knowledge and Data Engineering*, Vol. 1, No. 1, March 1989.

[16] C. C. Lee, "Fuzzy Logic in Control System: Fuzzy Logic Controller, Part I and II", *IEEE Trans. on Systems, Man and Cybernetics*, Vol. SMC-20, No. 2, 1990.

[17] H. J. Zimmermann, "Fuzzy Set Theory - and its Applications", *Kluwer-Nijhoff Publisher*, 1985.

[18] H. J. Zimmermann, "Fuzzy Sets, Decision Making, and Expert Systems", *Kluwer-Nijhoff Publishers*, 1987.

[19] C. V. Negoita, "Expert Systems and Fuzzy Systems", *Benjamin/Cummings Publishing*, 1985.

[20] S. Adiga, P. Cogez, "An Integrated Fuzzy Knowledge-Based System to Support Shop-Floor Production Dispatching Decisions", Submitted to *IEEE Trans. on Data and Knowledge Engineerings*, 1990.

[21] K-K Lin, C. Y. Fu, and D. A. Hodges, "A Simple Uniformity Algorithm for LPCVD Polysilicon Deposition," presented at the *3rd SRC/DARPA IC-CIM workshop*, Stanford, California, August 1988.

CHAPTER 4
AN EXPERIMENTAL DESIGN FOR LPCVD CHARACTERIZATION

4.1. Introduction

A generic physical model, such as the deposition rate model discussed in Ch. 3, must be characterized in order to represent a specific piece of manufacturing equipment. A *statistically designed experiment*, with a small number of strategically-chosen experimental runs, can help characterize the model with minimal uncertainty and little computational effort.

Among the many statistical experimental design techniques, the D-optimal design technique was chosen for this application because it is best suited for the confirmation of *known* models. Other experimental design techniques (e.g. factorials, orthogonal arrays, etc.) are geared towards *exploring* the design space in order to discover important input parameters or unknown parameter interactions [1-3]. Next, we will briefly describe the theory behind the D-optimal designs, and its application on designing a minimal number of LPCVD experiments towards establishing accurate LPCVD models.

4.2. Background on D-optimal Design

The family of *optimal experimental designs* is used to select from a predetermined design space a number of runs that optimize a certain statistical criterion. For D-optimal designs, in particular, the criterion is the minimization of the variances of the estimated model coefficients. This way we obtain a model with minimal variances and

correlations for the estimated coefficients, and a low[1] maximum variance of prediction [3-6].

4.2.1. The Theory of D-optimal Design

To illustrate the D-optimal design we will explore a simple linear statistical model[2]:

$$\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{4.1}$$

where[3]:

$\mathbf{y}$ is a $n \times 1$ vector consisting of the observed experimental responses.

$\mathbf{X}$ is a $n \times k$ matrix of the modeling terms at the selected experimental points.

$\boldsymbol{\beta}$ is a $k \times 1$ vector of the modeling coefficients to be estimated.

$\boldsymbol{\varepsilon}$ is a $n \times 1$ vector of residuals distributed around zero with variance σ^2.

In general, we need at least k experimental points in order to obtain an estimate of β. For example, consider a quadratic model, $y = \beta_1 + \beta_2 x + \beta_3 x^2$, and 5 experiments. In this case $\mathbf{y}$, $\mathbf{X}$, $\boldsymbol{\varepsilon}$ and $\boldsymbol{\beta}$ are as follows:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}, \mathbf{X} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \\ 1 & x_5 & x_5^2 \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}, \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{bmatrix} \tag{4.2}$$

---

[1] G-optimal design explicitly minimizes the maximum variance of prediction [2,3].

[2] "Linear" in the model coefficients to be estimated. The theory for nonlinear D-optimal design is more complex and out of the scope of our investigation [7,8].

[3] In this report we use lower-case bold-face letters (i.e. $\mathbf{y}$) to represent column vectors, and upper-case bold-face letters (i.e. $\mathbf{X}$) to represent matrices.

We now return to Eq. (4.1). The least-square estimation [9] of β is given by:

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \cdot \mathbf{X}' \mathbf{y} \tag{4.3}$$

and the variance-covariance matrix [9] of $\hat{\beta}$, defined as $VC(\hat{\beta})$ is:

$$VC(\hat{\beta}) = (\mathbf{X}'\mathbf{X})^{-1}\sigma^2 \mathbf{I} = (\mathbf{X}'\mathbf{X})^{-1}\sigma^2 \tag{4.4}$$

Here $\mathbf{X}^{-1}$ is the inverse of $\mathbf{X}$, $(\mathbf{X}^{-1})'$ is the transpose of $\mathbf{X}^{-1}$, $VC(\mathbf{y}) = \sigma^2 \mathbf{I}$, σ^2 is the variance of ε, and $\mathbf{I}$ is the k×k identity matrix.

The D-optimal criterion minimizes the determinant of $VC(\hat{\beta})$ in Eq. (4.4). This determinant is an overall measure of the size of its individual elements [4-6,8]. Equivalently, the D-optimal criterion can be satisfied through the maximization of the determinant $|\mathbf{X}'\mathbf{X}|$ as indicated in Eq. (4.5).

$$\max_{\mathbf{X}} |\mathbf{X}'\mathbf{X}| \tag{4.5}$$

The DETMAX algorithm [4-6] is an iterative search algorithm that selects the experimental points to maximize $|\mathbf{X}'\mathbf{X}|$. This search algorithm will randomly select an initial subset of experimental points from a superset of candidates, and then iteratively add or discard points until the D-optimal criterion, as defined in Eq. (4.5), is met. The RS/1 statistical software[4] supports the DETMAX D-optimal algorithm and was used for this work.

[4] RS/1 is a product of the BBN Corporation.

4.2.2. Unique Features of D-optimal Design

As a *confirmatory* experimental design technique, the D-optimal algorithm has the following advantages:

(1) It can be used to find either the minimal number of runs to achieve a certain variance in the coefficients of the user-supplied model. Alternatively, it can be used to minimize the variance of the model coefficients if the number of runs is given.

(2) A D-optimal design can handle an arbitrary design space (Fig 3.2).

(3) A D-optimal design allows for a flexible number of experiments, depending on the experimental budget.

(4) A D-optimal design permits experimentation in stages. This means that we can perform experiments in several stages in order to refine the models after each stage, and use the refined models to generate the next set of experiments. This multi-stage sequential feature effectively gives to this approach an *exploratory* as well as a *confirmatory* character.

4.3. The LPCVD D-optimal Experimental Strategy

Given a total experimental budget of 24 runs, we divided the experiments into 2 stages. The first stage included 12 runs and it was designed to suit the simplified nominal deposition rate model derived in Ch. 3.

Since no physical model was available for the residual stress [10,11], the stress model was not included in the design of the first experimental stage. The second stage consisted of 12 additional runs and was designed to suit both the calibrated deposition rate model and the empirical residual stress model. The empirical stress model for designing the second stage was derived after the first experimental stage.

4.3.1. Designing the First Stage of the Experiment

In order to generate a D-optimal experimental design, we used the nominal deposition rate model described in Ch. 3, repeated below:

$$R_0 = A\,P \exp\left(\frac{-\Delta E}{kT}\right) \frac{1-\frac{\lambda}{Q}}{1+\frac{\lambda}{Q}} \tag{4.6}$$

The model expresses the deposition rate of the first wafer in terms of the temperature (T), pressure (P) and silane flow (Q).

To simplify the D-optimal design, the model was linearized by employing a Taylor's expansion of the natural logarithm in Eq. (4.6). The resulting 6-term *linear additive* model has 6 coefficients to be identified. Higher order terms in the linearized model were discarded as numerically insignificant.

$$\ln R_0 \approx c_1 + c_2 \ln P + c_3 T^{-1} + c_4 Q^{-1} + c_5 Q^{-2} + c_6 Q^{-3} \tag{4.7}$$

The theoretical values of these coefficients are:

$$c_1 = \ln A,\ c_2 = 1 \text{ (ideally)},\ c_3 = \frac{-\Delta E}{k},\ c_4 = -2\lambda,\ c_5 = -\lambda^2, \text{ and } c_6 = \frac{-2}{3}\lambda^3 .$$

The D-optimal criterion was used to choose the runs from a set of 890 possible experimental points. These 890 points were set to provide a fairly uniform coverage of the design space (Fig. 3.2), taking into account the degree of control for every one of the primary parameters. For example, since the furnace can only control the temperature to within 1 °C, the temperature was set at 5 °C increments. Since the deposition time (t) was expected to affect the built-in residual stress, the process engineers selected its values after the worksheet for temperature (T), pressure (P) and silane flow (Q) was generated.

The worksheet with the first 12 runs is shown in the upper half of Table 4.1. The sequence of the actual runs was randomized to avoid bias errors [1,2]. Before completing the first 12 runs, RS/1 [3] was used to evaluate the design as having an average *Relative error of prediction*[5] of 0.706, and a *G-efficiency*[6] of 0.825, over the 12 runs. Generally, we would prefer a low Relative error of prediction and a high G-efficiency. Since our experiments included very little replication, the Relative error of prediction and G-efficiency were not considered as very accurate measures of the effectiveness of the models.

Fig. 4.1 shows how the 12 runs were scattered in the design space. Since these points had been selected according to the D-optimality criterion, they do not follow any obvious geometrical pattern.

---

[5] The *Relative error of prediction* is an overall measurement of the relative variance of the fitted model. This value ranges from 0 to 1, and a good model tends to minimize it.

[6] The *G-efficiency* is an overall figure of merit for an experimental design. It is inversely proportional to the number of experiments and the variance of the fitted model.

First Stage Design				
Run#	*Pressure(mtorr)*	*Temperature(°C)*	*SiH4(sccm)*	*Dep-time(min)*
1	340	605	125	120
2	320	650	100	60
3	550	605	125	90
4	550	620	250	70
5	430	605	175	110
6	550	610	250	80
7	540	605	100	100
8	370	650	125	60
9	520	650	175	60
10	300	605	100	150
11	550	650	100	60
12	550	650	125	60

Table 4.1 A 2-stage D-optimal experimental design (First stage shown).

Second Stage Design				
Run#	*Pressure(mtorr)*	*Temperature(°C)*	*SiH4(sccm)*	*Dep-time(min)*
13	550	605	100	60
14	300	605	100	150
15	540	650	225	100
16	550	650	100	110
17	320	650	100	150
18	550	605	250	60
19	300	605	100	110
20	550	605	100	60
21	470	625	200	130
22	550	630	100	60
23	470	625	200	120
24	550	630	100	60

Table 4.1 (cont.) A 2-stage D-optimal experimental design (Second stage shown).

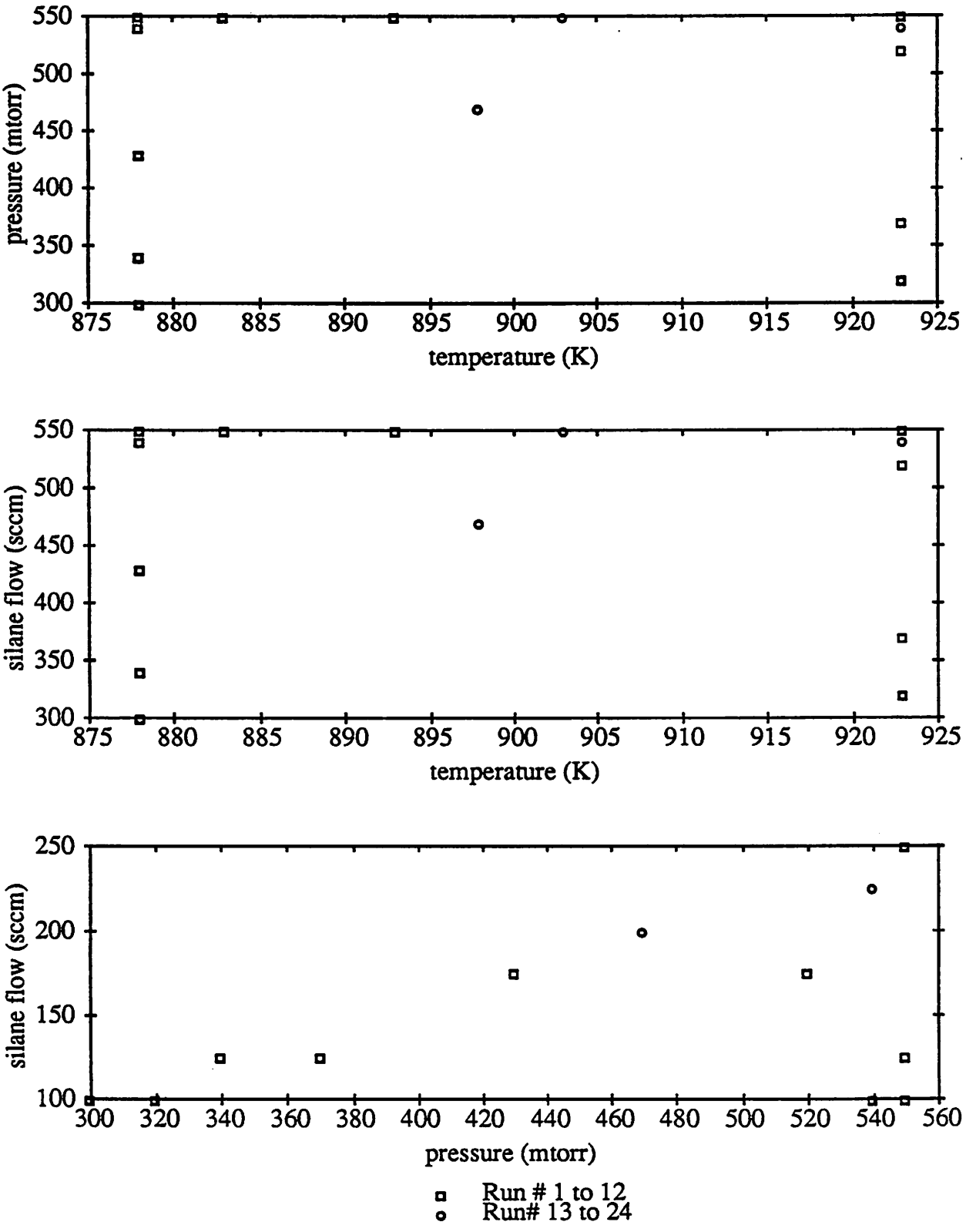


Fig. 4.1 Scatter plots of the design points

4.3.2. Designing the Second Stage of the Experiment

After analyzing the data from the first stage (the analysis will be discussed in detail in Ch. 5) we derived two models for designing the second stage D-optimal experiment. First, we simplified the linearized physically-based deposition rate model to only 4 terms:

$$\ln R_0 \approx c_1 + c_2 \ln P + c_3 T^{-1} + c_4 Q^{-1} \tag{4.8}$$

The coefficients c_1 to c_4 are defined as in Eq. (4.7), and were fitted with specific values. The last two terms in Eq. (4.7) were excluded, since in our data we found no evidence to support the hypothesis that they were different from zero.

Subsequently, we fitted the empirical residual stress[7] model shown below:

$$\text{Stress} = a_1 + a_2 T + a_3 t + a_4 T t + a_5 t^2 + a_6 T^2 \tag{4.9}$$

T is the temperature, and t is the deposition time. Pressure P and silane flow Q were not found to have any significant effect on the residual stress, neither were any higher order effects of time (t) and temperature (T).

Since the D-optimal design can only deal with one model at a time, the challenge was to design the second stage to satisfy the models discussed above. In order to optimize the accuracy of both the deposition rate and stress models, we divided the second stage design into two sub-stages. In the first sub-stage we optimized the linearized deposition rate model of Eq. (4.8), and set the experimental values for temperature (T), pressure (P) and silane flow (Q). To set the values for time (t) we applied the D-optimal criterion on the residual stress model of Eq. (4.9). In the second sub-stage we

[7] The film residual stress has been traditionally associated with the film thickness. The objective of this work is to build a stress model that relates to the independent process parameters.

optimized the residual stress model, from which we set the experimental values for temperature (T) and time (t). The remaining values for pressure (P) and silane flow (Q) were finally set from the linearized deposition rate model.

The design space for the second stage included a total of about 1000 possible runs. Note that in the second stage, the parameter time (t) was represented in the residual stress model, so a step size of 10 minutes was assigned to generate the time (t) settings.

The worksheet for the last 12 runs of the second stage, consisting of 8 runs for the first sub-stage and 4 runs for the second sub-stage, is shown in the second half of Table 4.1. The first 8 runs (# 13-20) showed that the experimental settings of P, Q and T were clustered at the extreme ends of the design space, which according to the D-optimal algorithm, led to a more accurate linearized deposition rate model. The last 4 runs (# 21-24) used T and t values not explored in the first stage.

Before completing the second stage, RS/1 was used to evaluate the second stage design using the G-efficiency and the Relative error of prediction. The G-efficiency and the Relative error of prediction are tabulated against the number of designed runs for both models in Table 4.2. For the linearized deposition rate model, the G-efficiency peaks at around run # 20, and the Relative error of prediction shows only a gradual decline after run # 20. This implies that for the linearized deposition rate model alone, any more runs beyond run # 20 offer only a marginal decrease of the Relative error of prediction with a drop in the G-efficiency. This is an indication that additional runs would not improve the quality of the model.

For the residual stress model however, the G-efficiency increases with the number of designed runs, and the Relative error of prediction decreases gradually with the number of designed runs. This implies that runs beyond our original experimental budget could be used to improve the residual stress model. In the next chapter we will

describe the model characterization procedure in detail.

linearized deposition rate model		
Run #	*G-efficiency*	*Rel. Error of Prediction*
12	0.825	0.706
20	0.835	0.446
24	0.743	0.406

Residual stress model		
Run #	*G-efficiency*	*Rel. Error of Prediction*
20	0.304	0.521
24	0.509	0.488

Table 4.2 "G-efficiency" and "Relative error of prediction" versus number of experimental runs.

REFERENCES

[1] G. E. Box, W. G. Hunter, J. S. Hunter, "Statistics for Experimenters - An Introduction to Design, Data Analysis, and Model Building", John Wiley & Sons, 1978.

[2] G. E. Box, N. R. Draper, "Empirical Model-Building and Response Surfaces", John Wiley & Sons, 1987.

[3] "RS/Discover: User's Guide", *Chapter 3.8.2 and Appendix D on the D-optimal design"*, pp. 3-46 and D-1, BBN Software Products Corporation, 1988.

[4] M.J. Box, N.R. Draper, "Factorial Design, the |X'X| Criterion, and Some Related Matters", *Technometrics, Vol. 13. No. 1*, November 1971, pp. 731-742.

[5] H.P. Wynn, "The Sequential Generation of D-Optimum Experimental Designs", *The annals of Mathematical Statistics, Vol: 41, No. 5*, 1970, pp. 1655-1664.

[6] T.J. Mitchell, "An Algorithm for the Construction of "D-Optimal" Experimental Designs", *Technometrics, Vol. 16, No. 2*, May 1974, pp. 203-210.

[7] I. Ford, D. M. Titterington, and C. P. Kitsos, "Recent Advances in Nonlinear Experiment Design", *Technometrics, Vol. 31. No. 1*, February 1989, pp. 49-60.

[8] R. H. Myers, et. al., "Response Surface Methodology: 1966 - 1988", *Technometrics, Vol. 31, No. 2*, May 1989, pp. 137-159.

[9] N. R. Draper, H. Smith, "Applied Regression Analysis", John Wiley & Sons, Second Edition, 1981.

[10] H. Guckel, et. al., "Fine-Grained Polysilicon Films with Built-in Tensile Stress", *IEEE Transaction on Electron Devices*, Vol. 35, No.6, June 1988, pp.800-801.

[11] H. Guckel, D.W.Burns, H.A.C. Timans, C.C.G. Visser, D.W. DeRoo, T.R. Christenson, P.J. Klomberg, J.J. Sniegowski and D.H. Jones, "Processing Conditions for Polysilicon Films with Tensile Strain for Large Aspect Ratio Microstructures," IEEE

Solid-State Sensors and Actuators Workshop, Hilton Head, SC (6-9 June 1988).

CHAPTER 5
THE CHARACTERIZATION OF THE LPCVD MODELS

5.1. Introduction

The generic physical models developed in the previous chapters were characterized with data collected from the statistically designed experiments (Ch. 4). The models are:

- the *semi-empirical* deposition rate depletion model.
- the *empirical* "within-wafer" deposition rate uniformity model.
- the *empirical* residual stress model.
- the *empirical* residual stress uniformity model.

These models are fitted and refined for our deposition furnace by standard least-square regression analysis [1], with data obtained from the 2-stage D-optimal experimental design. "Homoscedasity[1]" is assumed during the regression analysis of the above models. This assumption is justified by the observation that the equipment noise is random, and is independent of processing conditions for the range covered by our applications.

5.2. Experiment and Data Collection

(1) **Wafer selection:** <111>[2] sample wafers covered with 1000 Å of oxide were used.

---

[1] Constancy of data variance across the experimental space.

[2] <111> wafers have a constant Poisson's ratio ν_s (= lateral-strain/axial-strain) [2] of the substrate, which will later simplify the residual stress computation.

(2) **Deposition setup**: Eight <111> sample wafers were mixed evenly with 18 dummy wafers for the deposition, as prescribed in the worksheet of Table 4.1. The sample wafers were located at positions 4, 7, 8, 9, 12, 17, 21 and 25 of the 26-slot boats. Polysilicon was deposited on both sides of the wafers.

(3) **Thickness (deposition rate) measurement**: Five points (center, upper and lower corners) were measured on each wafer. The thickness of the deposited polysilicon was first checked by the *Nanometrics NanoSpec AFT* thickness measurement system. Since the refractive index of the material was not known exactly, the measurement was calibrated using the more accurate *AlphaStep 200* automatic step profiler. The deposition rate of the polysilicon was obtained by dividing the thickness by the processing time.

(4) **Residual stress measurement**: The following formula was used to compute the polysilicon residual film stress [3,4]:

$$\text{Residual stress} = \frac{E_s}{3(1-\nu_s)}\left[\frac{t_s}{r}\right]^2\left[\frac{f_1-f_2}{t_f}\right] \quad \text{in dyne/cm}^2 \qquad (5.1)$$

t_s : substrate thickness
r : radius of curvature
f_1 : substrate deflection with 2 side poly-si
f_2 : substrate deflection with 1 side poly-si
t_f : polysilicon film thickness
E_s : Young's modulus of the substrate
ν_s: Poisson's ratio of the substrate

Given that:

$$\left[\frac{E_s}{1-\nu_s}\right]_{<111>} = 2.29\times 10^{12}\ \text{dyne/cm}^2 \qquad (5.2)$$

wafer flatness (f_1, f_2) was measured at the same 5 locations by the *Tencor* flatgage, both before and after etching away the front side polysilicon using a *Lam* plasma etcher. The substrate thickness (t_s) was measured by the *Tencor* sonogage, and the film thickness

(t_f) was measured by the *NanoSpec* interferometer.

In all, about 40K bytes worth of data were collected and analyzed by the statistical software package RS/1.

5.3. Characterization of the First Stage Models

5.3.1. Characterization of the First Stage Deposition Rate Model

To characterize the entire deposition rate depletion model described in Eqs. (3.18) to (3.21), it is sufficient to characterize the deposition rate of the first wafer, i.e. R_0 in Eq. (4.6) of Ch. 4. R_0 was obtained by interpolating the deposition rates of the sample wafers, and then fitted to the linearized expression "ln R_0" in Eq. (4.7). The data of "R_0" and the corresponding "ln R_0" are shown in the experimental worksheet of Table 5.1. Eqs. (4.6) and (4.7) are repeated below:

$$R_0 \simeq A\, P \exp\left(\frac{-\Delta E}{kT}\right) \frac{1 - \frac{\lambda}{Q}}{1 + \frac{\lambda}{Q}} \tag{4.6}$$

$$\ln R_0 \simeq c_1 + c_2 \ln P + c_3 T^{-1} + c_4 Q^{-1} + c_5 Q^{-2} + c_6 Q^{-3} \tag{4.7}$$

The ANOVA table for the linearized deposition rate model regression analysis is shown in Table 5.2. The *F-test* revealed that this model and all its coefficients are statistically significant at the 5% level of significance. The two higher-order terms for silane flow rate (Q) were found to be statistically insignificant at the 5% level of significance, and were omitted from this table. This simplifies the linearized deposition rate model "ln R_0" to 4 terms. Without replicates, the associated *random error* cannot be estimated. However, if the model is accurate to the extent that the *lack of fit* of the model can be assumed to be statistically insignificant, then the *residual* is

approximately equal to the *random error*. With this assumption, the model will predict the mean response of the equipment with a one sigma *prediction error*[3] of ± 0.027 ln Å/min (about 3% of the average deposition rate in our range of operation). The actual response of the equipment will vary around the mean value with a one sigma *replication error*[4] of ± 0.055 ln Å/min (around 6% of the average deposition rate).

The fitted first stage linearized deposition rate model is shown below:

$$\ln R_0 = \underset{(0.92)}{21.51} + \underset{(0.08)}{0.22}\,(\ln P) - \underset{(632.83)}{15414.13}\,(T^{-1}) - \underset{(8.34)}{61.92}\,(Q^{-1}) \quad \text{in } \ln \text{Å/min} \qquad (5.3)$$

with P is expressed in *mtorr*, T in K, Q in *sccm*, and R_0 in Å/*min*. The standard errors of the corresponding estimated coefficients are shown underneath the model in parenthesis. Details of model evaluation are presented in Ch. 6.

3 The *prediction error* of the model can be computed approximately by the formula: $p\,\sigma^2/n$, where p is the number of parameters used in the model, σ^2 is the variance, and n is the number of data.

4 The *replication error* measures the variation in equipment response when a run is repeated several times. This error is a characteristic of the equipment quality and it cannot be improved with an increased number of experiments.

Run #	*Pressure* (*mtorr*)	*Temperature* (*K*)	*SiH4* (*sccm*)	*Time* (*min*)	*Rate* (R_0) ($\frac{\text{Å}}{\text{min}}$)	ln R_0 (ln $\frac{\text{Å}}{\text{min}}$)	*Rate-Unif* (%)	*Stress* ($10^9 \frac{\text{dyne}}{\text{cm}^2}$)	*Stress-Unif* ($\log_{10}$ %)
1	339	882	125	120	113.7	4.696	0.73	5.83	1.139
2	318	926	100	60	257.5	5.553	0.80	-3.78	0.957
3	549	881	125	90	127.6	4.796	0.96	6.72	1.071
4	548	897	250	70	220.9	5.321	5.28	-1.46	2.517
5	427	882	175	110	149.0	4.919	1.12	6.44	1.155
6	548	888	250	80	184.2	5.157	3.41	5.81	1.389
7	538	882	100	100	115.0	4.727	0.56	5.27	1.422
8	366	926	125	60	275.0	5.617	0.77	-4.23	1.101
9	517	927	175	60	321.0	5.751	1.43	-4.47	1.271
10	296	882	100	150	91.0	4.560	0.89	-1.64	2.388
11	547	927	100	60	270.0	5.568	0.53	-4.23	1.106
12	547	927	125	60	292.0	5.693	0.66	-3.63	0.958
13	548	879	100	60	122.7	4.810	0.73	5.21	1.370
14	295	883	100	150	97.5	4.576	1.32	-0.68	2.481
15	537	927	225	100	354.2	5.872	0.97	-1.36	1.744
16	548	927	100	110	275.9	5.620	1.96	-2.30	1.445
17	316	927	100	150	262.4	5.573	0.50	-1.61	1.483
18	552	883	250	60	186.8	5.234	2.34	5.08	1.496
19	294	881	100	110	95.6	4.564	1.06	2.17	2.066
20	546	881	100	60	120.3	4.794	0.53	5.38	1.056
21	466	900	200	130	208.5	5.336	2.24	-0.76	2.018
22	546	907	100	60	190.6	5.247	1.84	-5.11	1.485
23	465	897	207	120	190.6	5.247	1.15	0.12	3.403
24	545	904	100	60	183.1	5.209	0.76	-5.35	1.540

(+) stress values imply tensile stress

(-) stress values imply compressive stress

Table 5.1. The experimental worksheet

Source	DF	Sum of Squares	Mean Square	F-Ratio	Significance
Total	11	2.104	0.191		
Regression	3	2.079	0.693	221.600	0.000
Residual	8	0.025	0.003		
Lack of Fit	8	0.025	0.003		
Random Error	0				

Term	Coefficient	Standard Error	T-Value	Significance
1	21.51	0.92	23.27	0.0001
1/T	-15414.13	632.83	24.36	0.0001
log(P)	0.22	0.08	2.87	0.0207
1/Q	-61.92	8.34	7.43	0.0001

Table 5.2 ANOVA table of the first-stage deposition rate model

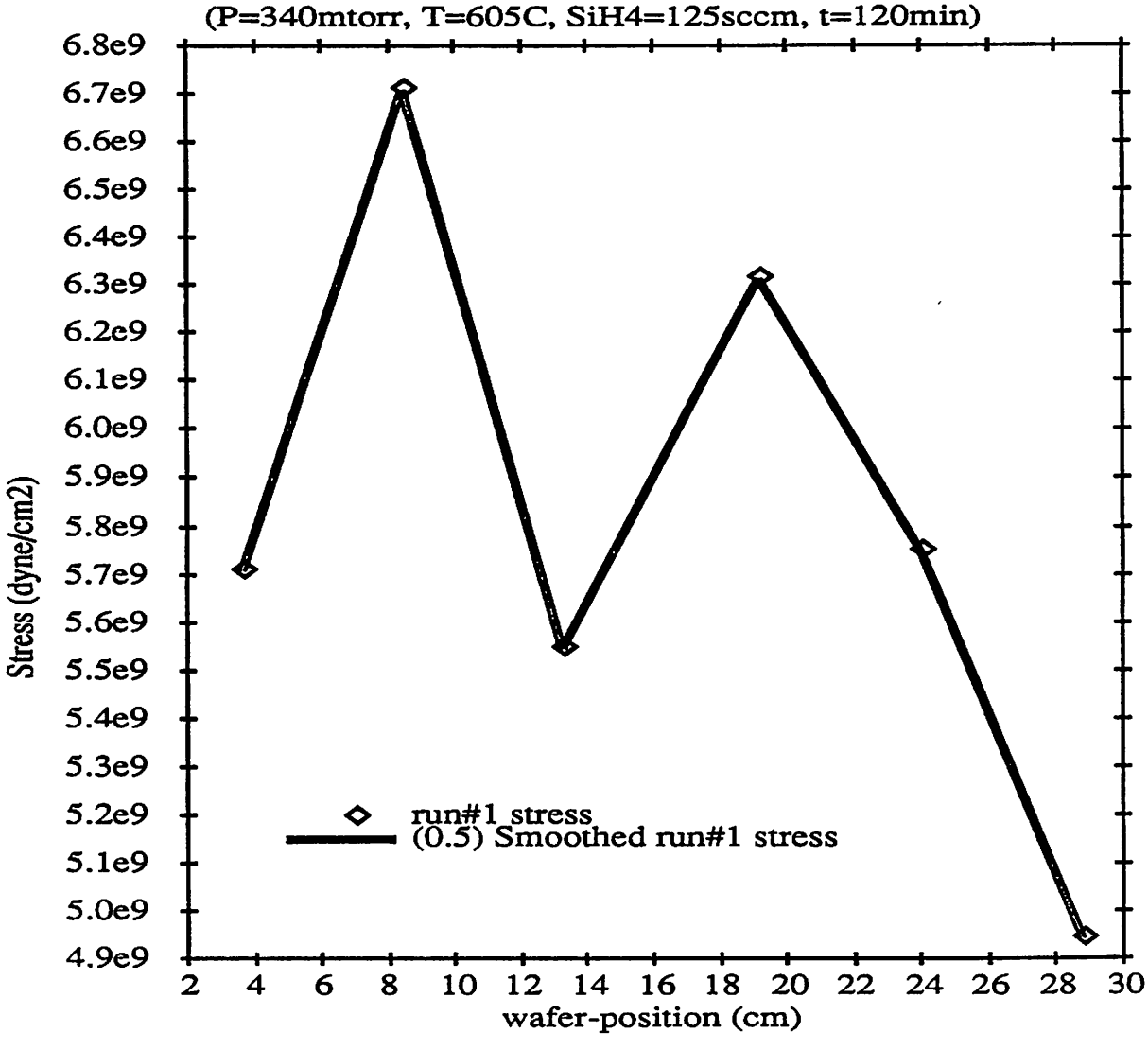


Fig. 5.1 Stress vs. position

5.3.2. Characterization of the First Stage Residual Stress Model

A plot of stress versus wafer position from a typical run is shown in Fig. 5.1. Like all other runs, it revealed no significant dependency of stress on the wafer position. Hence, the measured stress was averaged for each run over all wafers of the two boats. The data of the average stress of each run are also shown in the experimental worksheet in Table 5.1.

The ANOVA table for the stress model is shown in Appendix 1. This analysis suggested that the model is statistically significant at the 5% level of significance, with a 6-term second-order polynomial equation of T and t. The random error of the model is not calculated, but it would be equal to the residual assuming that the model shows no significant lack of fit. This model will predict the mean response for the residual stress with a one sigma prediction error of $\pm 0.33 \times 10^9$ dyne/cm^2. The one sigma replication error of the equipment around the average value of the residual stress is $\pm 0.78 \times 10^9$ dyne/cm^2.

The estimated coefficients and the corresponding standard errors are below:

$$\text{Stress} = \underset{}{164.60} + \underset{(36.67)}{191.19\,T'} - \underset{(41.22)}{182.02\,t'} + \underset{(40.83)}{192.29\,T't'} + \underset{(40.48)}{31.65\,(T')^2} - \underset{(5.01)}{17.17\,(t')^2} \text{ in } 10^9\text{dyne/cm}^2 \quad (5.4)$$

$$\text{(3.35)}$$

$$\text{where: } T' = \frac{T-900.5}{22.5}, \quad t' = \frac{t-120}{60},$$

and T is expressed in K, t in min.

The deposition rate and residual stress models inferred from the first experimental stage were used to design the experiments of the second stage.

5.4. Characterization of the Second Stage Models

We performed three calibration runs prior to the second stage experiment to ascertain the performance of the deposition furnace. The calibration runs confirmed that no systematic differences existed in the furnace between the two experimental stages. We finalized the characterization of the models in the second stage.

5.4.1. Characterization of the Second Stage Deposition Rate Model

The final data points of "ln R_0" shown in Table 5.1 were fitted to the linearized "ln R_0" model in Eq. (4.8), which was derived after the first stage experiments. Eq. (4.8) is repeated below:

$$\ln R_0 = c_1 + c_2 \ln P + c_3 T^{-1} + c_4 Q^{-1} \tag{4.8}$$

The ANOVA table for the final linearized deposition rate model is shown in Appendix 1. The model and all its coefficients are statistically significant at the 5% level. With no replicates and assuming the lack of fit of the model is statistically insignificant, the random error is estimated by the residual. Based on the above assumption, the one sigma prediction error of the linearized deposition rate model on the mean response of the equipment is approximately $\pm$ 0.019 ln Å/min (about 2% of the average deposition rate). The one sigma replication error of the equipment around this mean value is $\pm$ 0.055 ln Å/min (about 6% of the average deposition rate). The replication error is a characteristic of the equipment, and as expected, it did not improve during the second stage. Also as expected, the prediction error of the model decreased. The final linearized deposition rate model is shown below:

$$\ln R_0 = \underset{(0.73)}{20.65} + \underset{(0.06)}{0.29\,(\ln P)} - \underset{(515.92)}{15189.21\,(T^{-1})} - \underset{(5.24)}{47.97\,(Q^{-1})} \quad \text{in } \ln \text{Å/min} \quad (5.5)$$

where P is expressed in *mtorr*, T in K, Q in *sccm*, and R_0 in Å/*min*. The standard errors of the estimated coefficients for the linearized model are shown in parenthesis. The linearized model can either be used as is, or the estimated coefficients in the linearized model can be used to deduce the coefficients in the original non-linear deposition rate model shown below:

$$R(T, P, Q, z) \approx R_0(P, Q, T) \left[\frac{1 - \eta(P, Q, T, z)}{1 + \eta(P, Q, T, z)} \right] \quad \text{in Å/min} \qquad (5.6)$$

$$R_0(P, Q, T) \approx A P^{\alpha} e^{\frac{-\Delta E}{kT}} \left[\frac{1 - \frac{\lambda}{Q}}{1 + \frac{\lambda}{Q}} \right]$$

$$\text{mole fraction of depleted } Si_4 \approx \eta(P, Q, T, z) = \left[\frac{S' C_{gs}}{L' Q} R_0 \right] z$$

z = wafer position in cm. Wafers are 1.2 cm apart.
A = the Arrhenius frequencey factor = 9.29×10^8 Å/min·mtorr$^{\alpha}$
$\alpha = 0.29$ (differs from the ideal value of 1)
ΔE =the activation energy = 30.18 kcal/mol
$\lambda = 23.98$ sccm^{-1}
k = the universal gas constant = 1.98719 cal/mol–K
C_{gs} = the gas solid conversion factor = 1.85×10^{-5} cm/Å
S': Effective deposition area, 4777.8 cm^2 (for T ≈ 605 to 650 °C)
L': Length of the boats, 30.0 cm

5.4.2. Characterization of the Second Stage Residual Stress Model

The final 24 run-average stress data are also shown in Table 5.1 of the experimental worksheet. The stress data was used to fit a 6-term second-order polynomial. This polynomial was slightly different from the first stage stress model. Pressure P was

found to interact with temperature T and time t, which no longer interact with each other. We observed that the residual stress changes drastically at 898 K and 120 minutes. Hence, these threshold parameter values were taken to be the center points in the "normalized" polynomial model.

The ANOVA table for the final empirical stress model is shown in Appendix 1. The final stress model and its coefficients are statistically significant at the 5% level of significance. Given no replicates, the random error is approximated to the residual by assuming the model is accurate with insignificant lack of fit. The model will predict the mean stress of the equipment with a one sigma prediction error of $\pm$ 0.446×10^9 dyne/cm^2. The actual stress of the equipment will vary around the predicted mean value with a one sigma replication error of $\pm$ 1.261×10^9 dyne/cm^2.

The final model is shown below with the standard error of the coefficients in parenthesis:

$$\begin{aligned} \text{Stress} = {} & 3.86\,T' + 3.45\,t' - 2.52\,T'P' - 1.76\,P't' + 3.36\,(T')^2 - 5.14\,(t')^2 \text{ in } 10^9 \text{ dyne/cm}^2 \quad (5.7) \\ & (0.44) \quad (0.68) \quad (0.48) \quad (0.61) \quad (0.44) \quad (0.72) \end{aligned}$$

$$\text{where } T' = \frac{898-T}{22.5},\ t' = \frac{120-t}{45},\ P' = \frac{400-P}{150},$$
and P is expressed in *mtorr*, T in K, t in *min*.

5.4.3. Characterization of the Second Stage Uniformity Models

In addition to characterizing the nominal process responses discussed above, it is also important to model the process uniformity [5] of the deposition step. Here we are focusing on the collective data of the two experimental stages, and we model the following two process uniformities:

The "within-wafer" deposition rate uniformity model: We averaged the "within-wafer" deposition rate uniformity (in %) of all wafers under each experimental run, and empirically[5] modeled the "within-wafer" uniformity in terms of pressure (P), silane flow (Q), and temperature (T). Column 8 of Table 5.1 shows the data of the "within-wafer" deposition rate uniformity data. The ANOVA table for the "within-wafer" deposition rate uniformity model is shown in Appendix 1. The model shows statistical significance at the 5% level, with an empirical second-order polynomial consisting of 4 terms. The model has a one sigma prediction error of ± 0.242%, and a one sigma replication error of ± 0.686%. The final model is:

$$\Delta R = \underset{(0.31)}{1.49} + \underset{(0.25)}{0.99\,P'} + \underset{(0.23)}{1.06\,P'Q'} - \underset{(0.29)}{0.71\,(T')^2} \quad \text{in \%} \qquad (5.8)$$

$$\text{where: } P' = \frac{P-400}{150},\ Q' = \frac{Q-175}{75},\ T' = \frac{T-900.5}{22.5},$$
$$\text{and P is expressed in } \textit{mtorr}, \text{ Q in } \textit{sccm}, \text{ and T in K.}$$

Note that the "between-wafer" non-uniformity of the deposition rate due to reactant depletion is a systematic phenomenon, and it is implicitly modeled in the original deposition rate model of Eq. (5.6).

The residual stress uniformity model: The uniformity of the residual stress (defined as: $\frac{\text{standard error}}{\text{mean}} \times 100\%$) averaged over all the wafers in each run was also modeled. A logarithmic transformation was first applied to the wide range of the uniformity data, before entering them into the experimental worksheet of Table 5.1. The ANOVA table for the log-uniformity model is shown in Appendix 1. The model is statistically

[5] A physical model for "within-wafer" deposition rate uniformity, based on diffusion limitations in the interwafer region, was presented in [6].

significant with a 5-term second order polynomial in terms of the process parameters P, Q, T, and t. The one sigma prediction and replication errors of the model are ± 0.148 and ± 0.418 expressed in log %.

$$\log_{10}(\Delta\text{stress}) = 2.31 + 0.44\,T' + 0.38\,T'Q' - 0.44\,Q't' - 0.88\,(T')^2 \text{ in log \%} \quad (5.9)$$

(0.19) (0.17) (0.16) (0.13) (0.22)

$$\text{where } T' = \frac{T-900.5}{22.5},\ Q' = \frac{Q-175}{75},\ t' = \frac{t-105}{45},$$

and T is expressed in K, Q in *sccm*, t in *min*.

REFERENCES

[1] N. R. Draper, H. Smith, "Applied Regression Analysis", John Wiley & Sons, Second Edition, 1981.

[2] William D. Callister, Jr., "Material Science and Engineering: An Introduction", John Wiley & Sons, Inc., 1985.

[3] S. Wolf, R.N. Tauber, "Silicon Processing; for the VLSI Era; Volume 1 - Process Technology", *Chapter 3 on Chemical Vapor Deposition of Amorphous and Polycrystalline Film*, Lattice Press, pp. 161-197.

[4] R. J. Jaccodine, and W. A. Schlegel, "Measurement of Strains at Si-SiO_2 Interface", *Journal of Applied Physics*, Volume 37, Number 6, pp. 2429-2434, May 1966.

[5] G. Taguchi, M.S. Phadke, "Quality Engineering Through Design Optimization," *Conference Record, GLOBECOM84 Meeting, IEEE Communications Society*, Atlanta, Georgia, November 1984, pp. 1106-1113.

[6] S. Middleman and A. Yeckel, "A Model of the Effects of Diffusion and Convection on the Rate and Uniformity of Deposition in a CVD Reactor", *J. Electrochem. Soc.: SOLID STATE SCIENCE AND TECHNOLOGY*, Vol. 133, No. 9, pp. 1951-1956, September 1986.

CHAPTER 6
THE EVALUATION OF THE LPCVD MODELS

6.1. Introduction

Both the first and second stage models agree well with the experimental data. Moreover, the estimated parameters of the deposition rate model are comparable to those reported in the literature [1-5]. However, we note that the uniformity models of Eqs. (5.8) and (5.9) exhibit high prediction and replication errors, although their terms are statistically significant at the 5% level. The high replication error is possibly due to the noise inherent in the uniformity measurement. Nevertheless, we believe that the models can provide us with first-order process trend information.

6.2. The Deposition Rate Model

The scatter plot in Fig. 6.1 compares the actual deposition rates of the first wafer R_0 with the deposition rate model for the 24 calibration runs. Fig. 6.2 and 6.3 show two experimental deposition rate profiles, under typical processing conditions, which are in excellent agreement with the deposition rate model. The deposition rate model accounts for the "wafer-to-wafer" uniformity in terms of the SiH_4 depletion along the tube. Fig. 6.4 shows one of the three deposition rate profiles which are not in agreement with the experimental profiles. The discrepancy, however, is still within the three sigma limits (18%) of the replication error of our experiment.

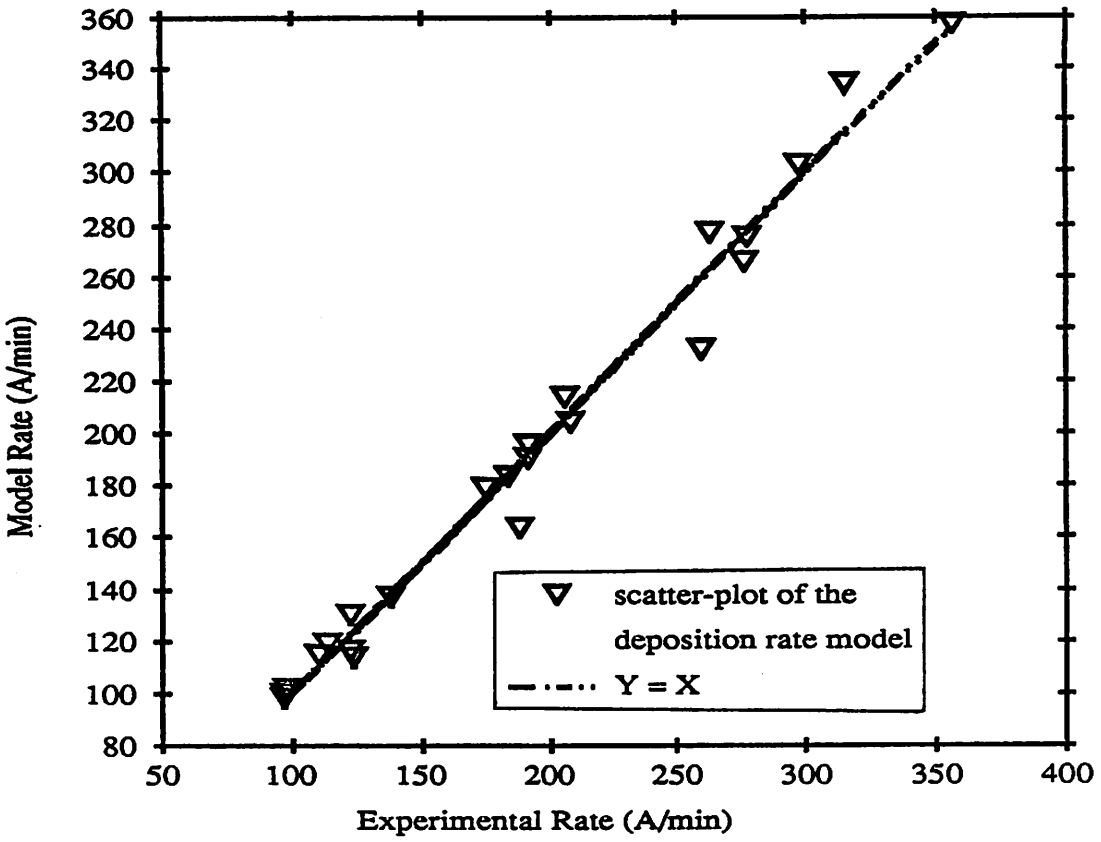


Fig. 6.1 Experimental rate vs. model Rate

(T= 878K, P= 340mtorr, SiH4= 120sccm, time= 120min)

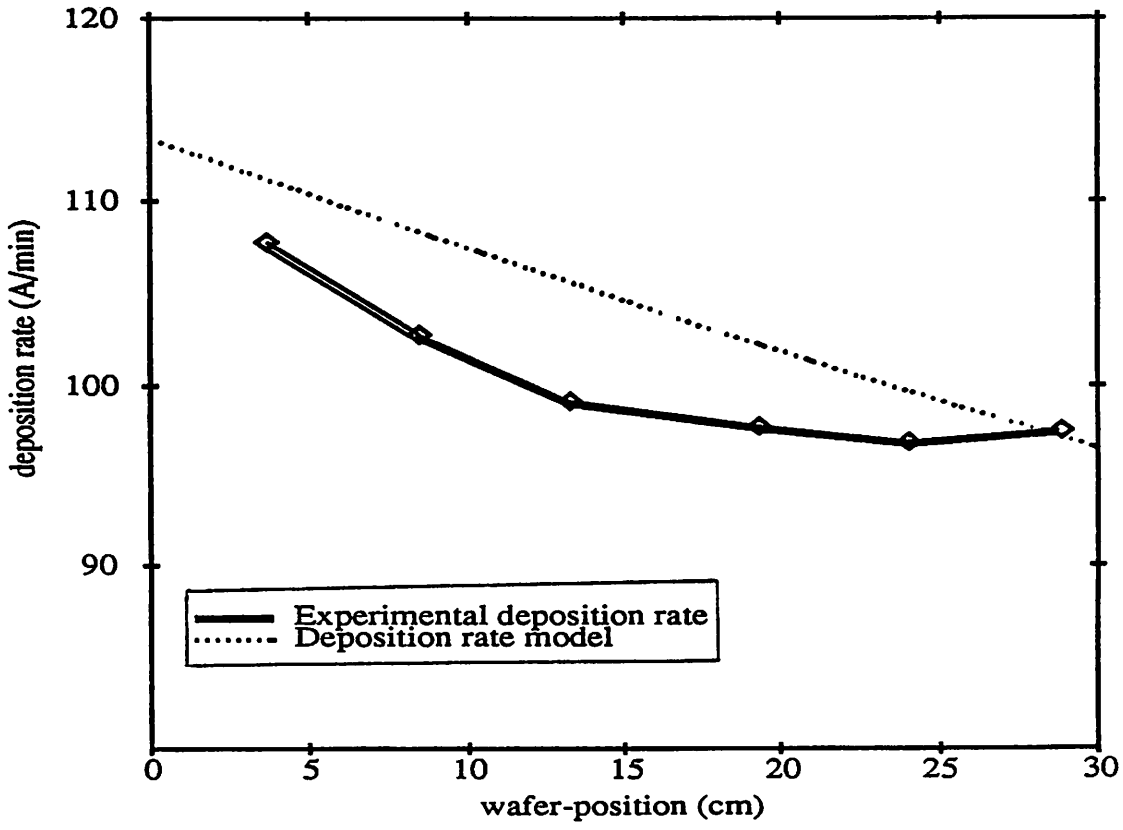


Fig 6.2 Deposition rate vs. wafer position for run # 1

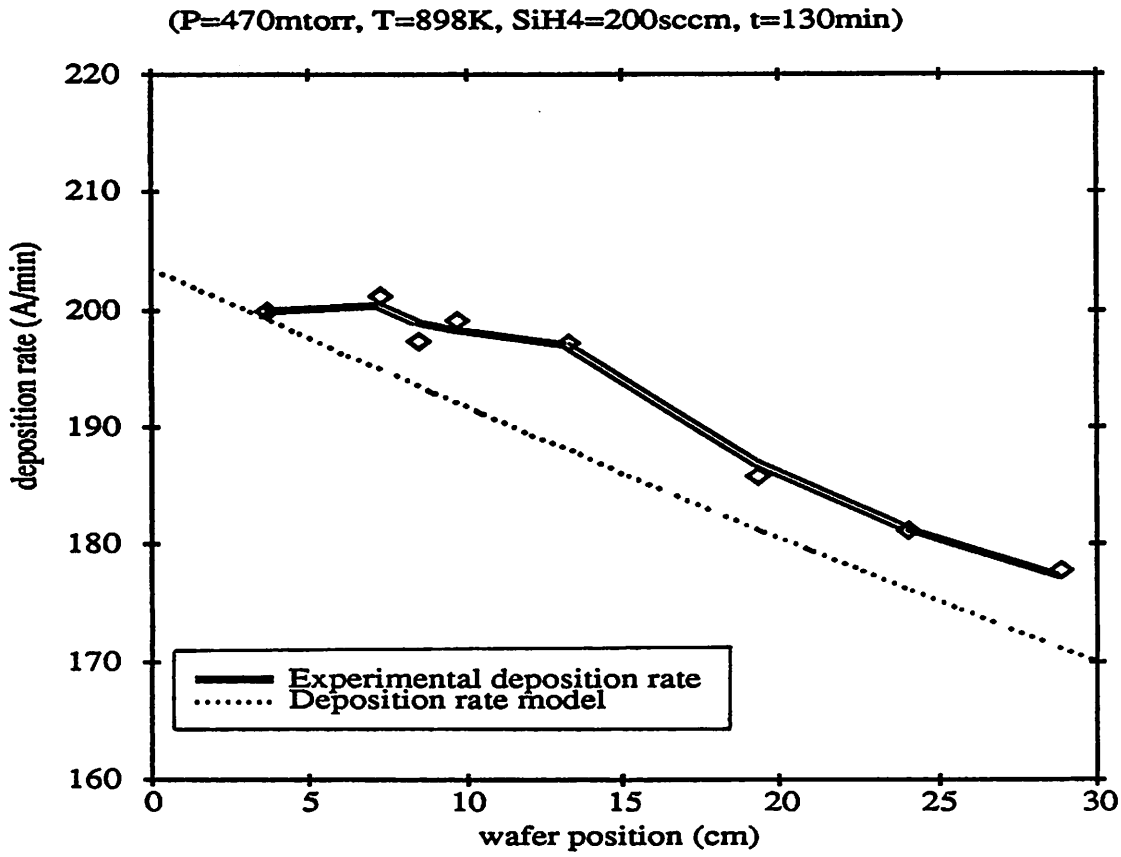


Fig. 6.3 Deposition rate vs. wafer position for run # 21

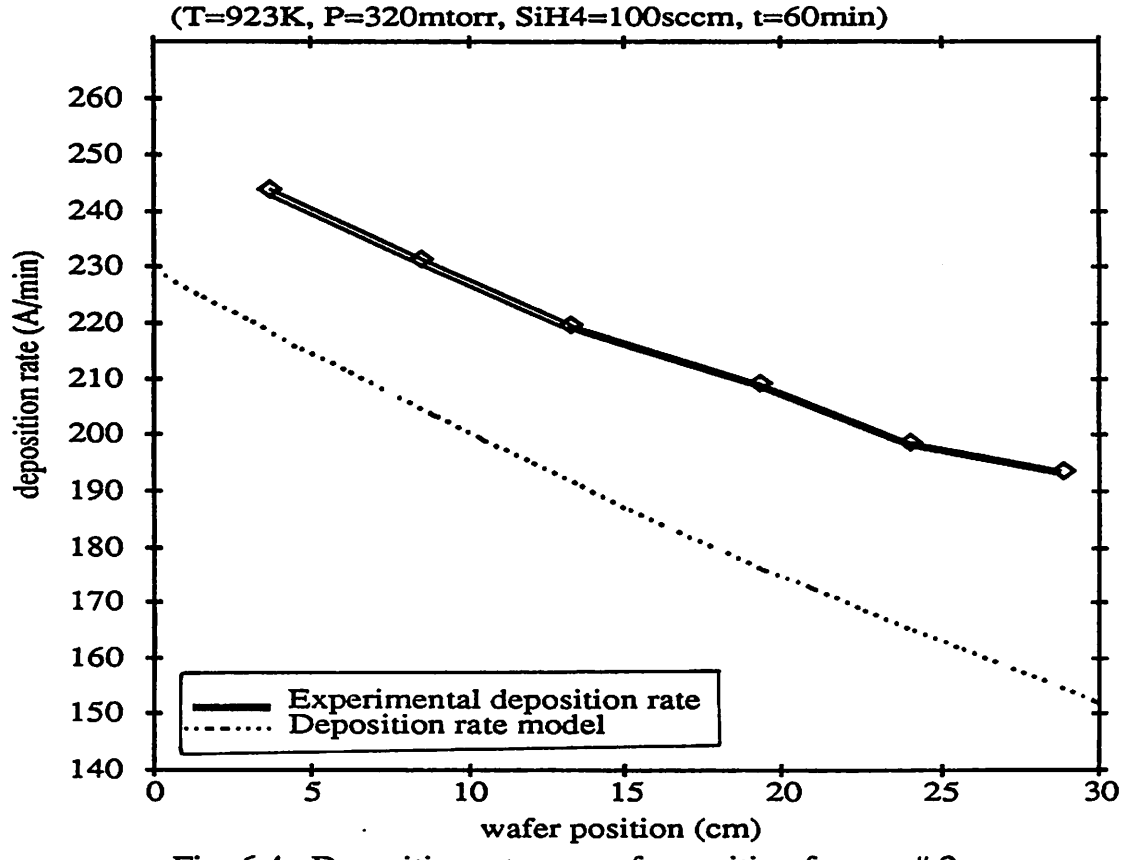


Fig. 6.4 Deposition rate vs. wafer position for run # 2

The deposition rate model, as shown in Eq. (5.6), has significant implications for process design and control. A high deposition rate R_0 is needed for manufacturing throughput, which can be achieved through a high temperature T or a higher pressure P or both. However, in order to improve the "wafer-to-wafer" uniformity in the batch deposition process, η (the mole depletion of SiH_4) must be minimized by decreasing the temperature T or the pressure P or both. A lower temperature T or a lower pressure P will cause a decrease in the deposition rate R_0, which should therefore be compensated by an increase in the silane flow to improve both the throughput and the uniformity.

The deposition rate model is further validated by comparing it to experimental results performed six to twelve months after the 24 characterization runs. The physically-derived, empirically calibrated model shows a good match with the data within the replication error of the equipment. Two deposition rate profiles, with both the experimental and model-predicted rates, are shown in Fig. 6.5 and Fig 6.6.

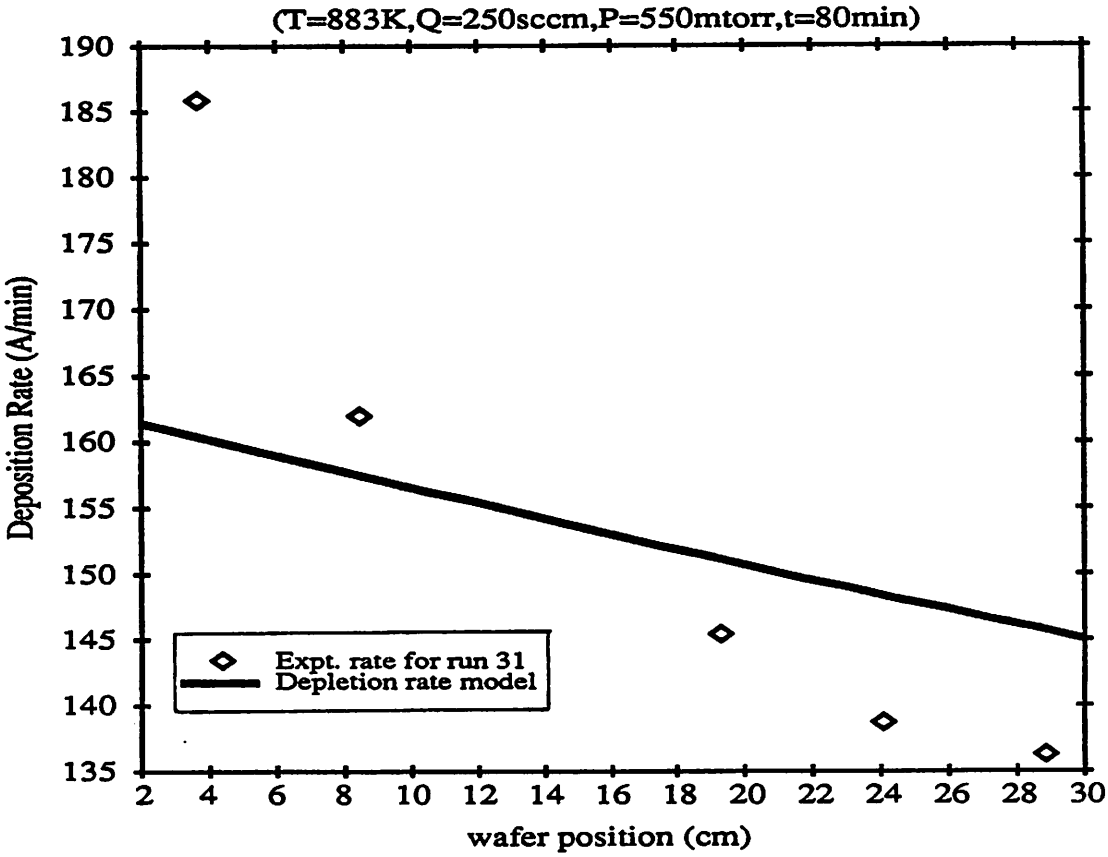


Fig. 6.5 Deposition rate profile for run # 31
(8 months after original characterization)

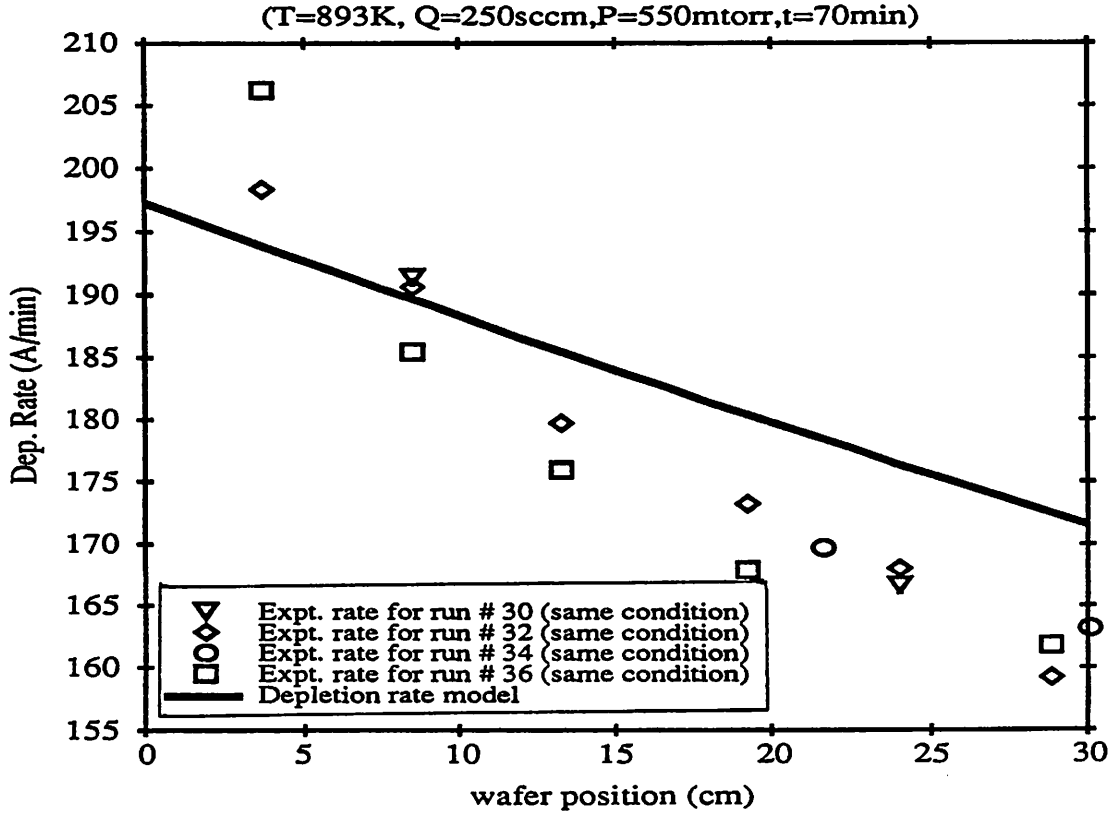


Fig. 6.6 Deposition rate profile for run # 30,32,34,36
(8-12 months after original characterization)

6.3. The Residual Stress Model

The scatter plot in Fig. 6.7 compares the average stress of the sample wafers in each run with the average stress model. The agreement between the actual values and the model values is within the replication error of the equipment.

A contour plot of the film stress, as a function of deposition time t and deposition temperature T (with pressure P and SiH_4 flow fixed at typical settings) is shown in Fig. 6.8. This plot reveals that in order to achieve a 'zero' film stress, as required in many VLSI micromachinery applications [6,7], the temperature T should be in the range of 885 to 905 K, with appropriate matching deposition time t. In the interest of repeatability, the best temperature T and time t combination is approximately 903 K and 105 minutes, where the stress is least sensitive to parameter variations.

The stress model is also validated by experimental data after the initial characterization runs. The comparison (Table 6.1) shows good match between the model and the data, given the replication error of the equipment.

Run#	expt. stress (10^9 dyne/cm^2)	lower 3 σ	model stress (10^9 dyne/cm^2)	upper 3 σ
31	5.71	4.15	5.19	6.24
32	1.31	0.14	1.03	1.91
36	0.52	0.14	1.03	1.91

Run# 31 was performed 8 months after the initial characterization.
Run# 32 was performed 8 months after the initial characterization.
Run# 36 was performed 12 months after the initial characterization.

Table 6.1 Comparison between model and experimental stress.

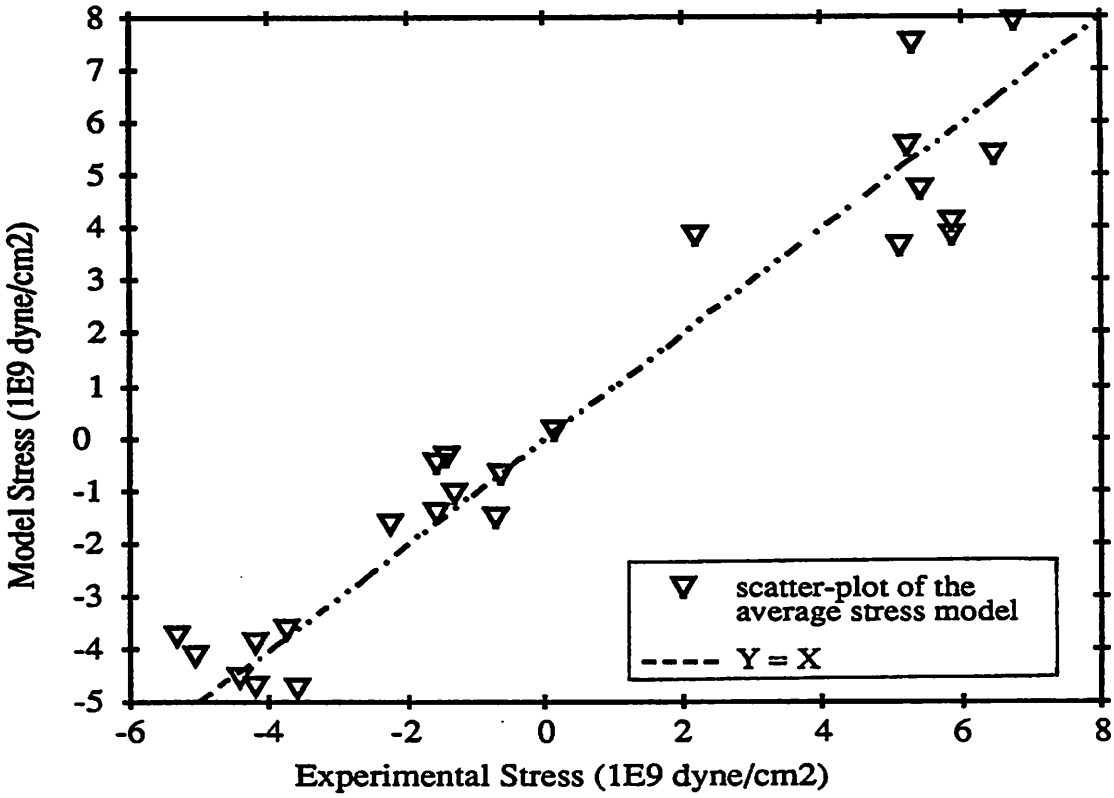


Fig. 6.7 Model stress vs. experimental stress

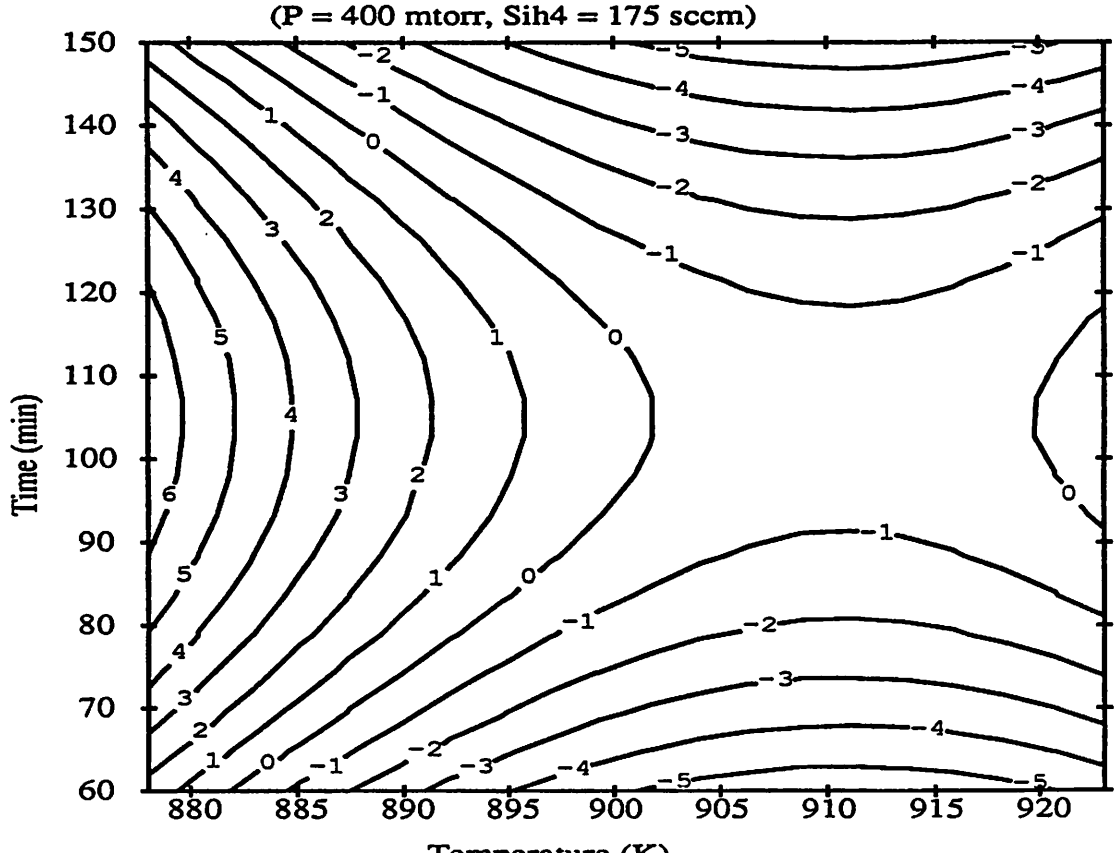


Fig. 6.8 Contour plot of stress

6.4. The "Within-wafer" Deposition Rate Uniformity Model

Fig. 6.9 shows a scatter plot that compares the average experimental "within-wafer" deposition rate variation (in %) with the model. All the discrepancies are consistent with the high replication error of the equipment.

Overall, the model has captured the basic process trends and is in agreement with the results reported in the literature [1-5]. The contour plot in Fig. 6.10 shows that the "within-wafer" deposition rate variation increases with high pressure and high SiH_4 flow, and decreases with lower pressure and lower SiH_4 flow. To enhance the "within-wafer" deposition rate uniformity, it is desirable to work in the low pressure region, with high silane flow at the upper left corner of the contour plot. This is consistent with the process limitations in Table 3.3.

Experimental data on "within-wafer deposition rate uniformity" were also gathered after the initial characterization runs. The match (Table 6.2) between the model and the later data is within the replication error of the furnace.

Run#	expt. unif. (%)	lower 3 σ	model unif. (%)	upper 3 σ
31	3.82	2.38	3.11	3.85
32	2.64	2.70	3.46	4.23
36	1.83	2.70	3.46	4.23

Run# 31 was performed 8 months after the initial characterization.
Run# 32 was performed 8 months after the initial characterization.
Run# 36 was performed 12 months after the initial characterization.

Table 6.2 Comparison between model and experimental uniformity.

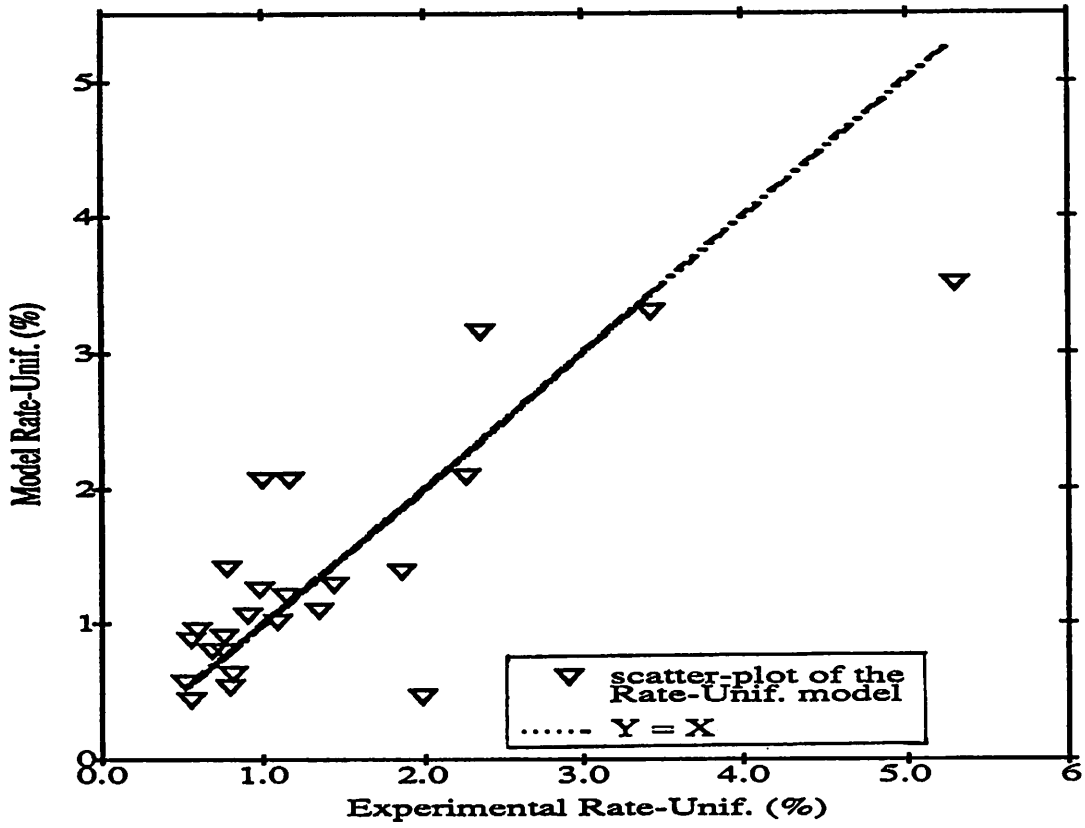


Fig. 6.9 Model rate unif. vs. experimental rate unif.

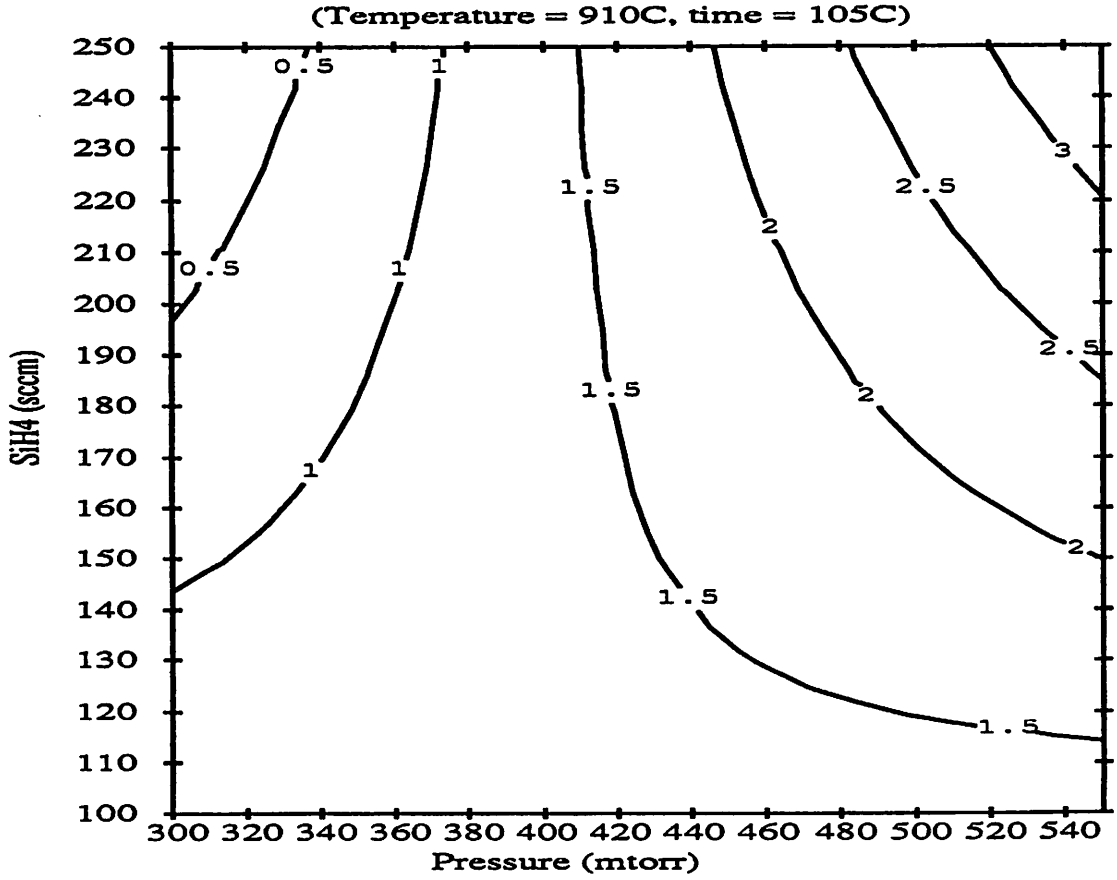


Fig. 6.10 Contour plot for deposition rate uniformity

6.5. The Residual Stress Uniformity Model

Fig. 6.11 is a scatter plot that compares the average experimental stress variation (in $\log_{10}$ %) with the model. The agreement between the actual data and the model is again consistent with the replication error of the film stress in the furnace.

In Fig. 6.12, the stress uniformity model is plotted by varying the temperature T and the silane flow Q for a deposition time of 105 minutes. The variation of the stress peaks between 895 and 915 K. In order to minimize the average stress variation of each run, it is best to operate at the upper left corner of the contour plots (i.e. the process window) where the temperature is low and the SiH_4 is high.

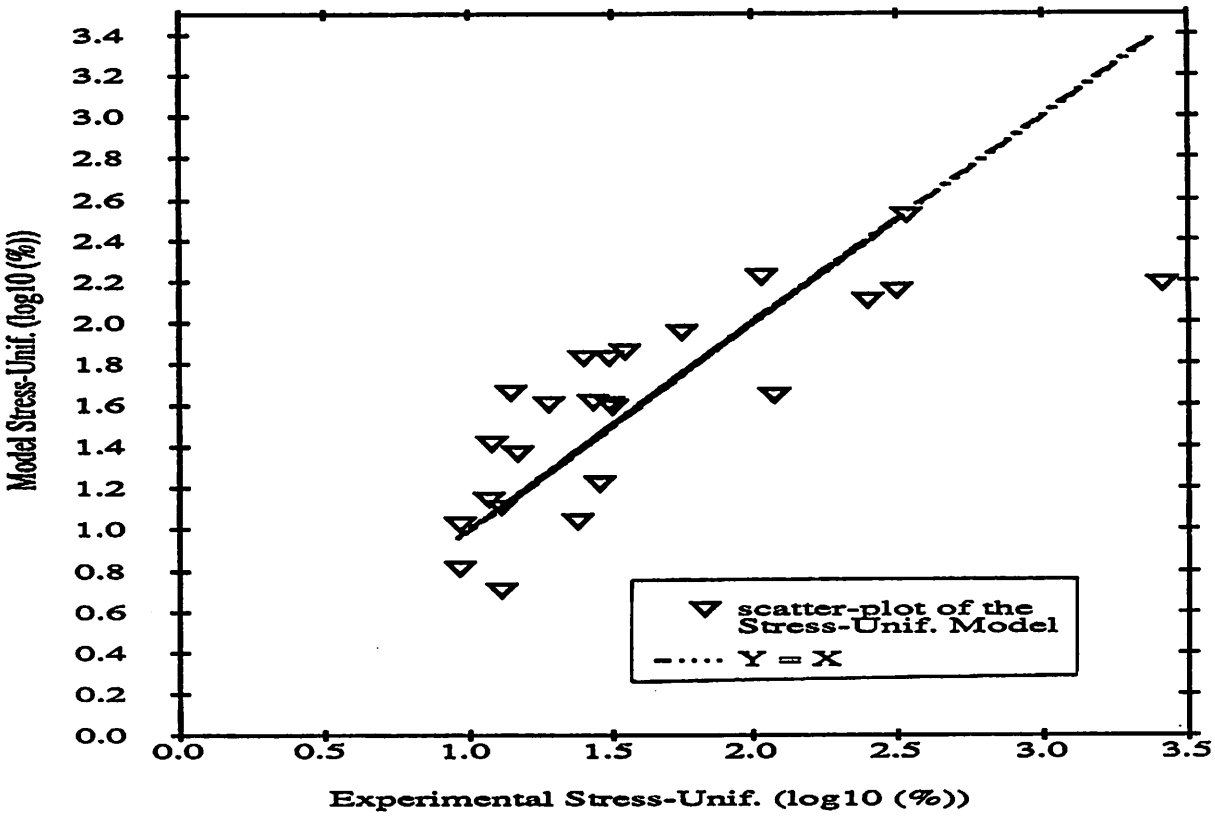


Fig. 6.11 Model stress unif. vs. experimental stress unif.

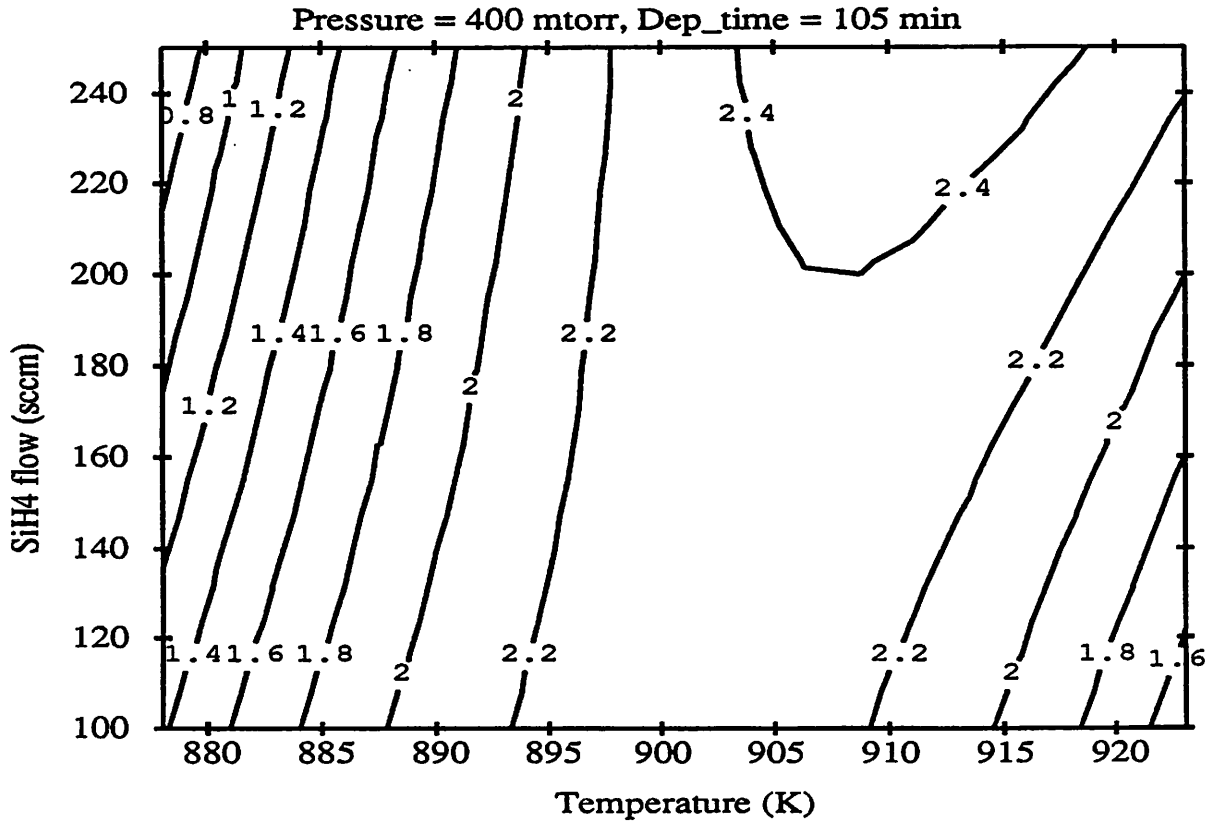


Fig. 6.12 Contour plot of stress uniformity

REFERENCES

[1] S. Wolf, R.N. Tauber, "Silicon Processing; for the VLSI Era; Volume 1 - Process Technology", *Chapter 3 on Chemical Vapor Deposition of Amorphous and Polycrystalline Film*, Lattice Press, pp. 161-197.

[2] R.J. Gieske, J.J. McMullen, "Low Pressure Chemical Vapor Deposition of Polysilicon". Materials and Molecular Research Division, Lawrence Berkeley Laboratory and Department of Chemical Engineering, University of California, Berkeley, 1977.

[3] W.A. Brown, T.I. Kamins, "An Analysis of LPCVD System Parameters for Polysilicon, Silicon Nitride and Silicon Dioxide Deposition", *Solid State Technology*, July 1979, pp. 51-57, 84.

[4] T. I. Kamins, "Polycrystalline Silicon for Integrated Circuit Applications", Kluwer Academic Publishers, 1988.

[5] N. Cheung, "Thin Film Technology for IC Fabrication", EECS 290P course notes, University of California, Berkeley, Spring 1988.

[6] H. Guckel, et. al., "Fine-Grained Polysilicon Films with Built-in Tensile Stress", *IEEE Transaction on Electron Devices*, Vol. 35, No.6, June 1988, pp.800-801.

[6] H. Guckel, D.W.Burns, H.A.C. Timans, C.C.G. Visser, D.W. DeRoo, T.R. Christenson, P.J. Klomberg, J.J. Sniegowski and D.H. Jones, "Processing Conditions for Polysilicon Films with Tensile Strain for Large Aspect Ratio Microstructures," IEEE Solid-State Sensors and Actuators Workshop, Hilton Head, SC (6-9 June 1988).

CHAPTER 7
COMPUTER-AIDED RECIPE GENERATION FOR LPCVD REACTORS

7.1. Introduction

In a multi-product manufacturing environment, equipment utilization tends to be low due to the time needed to design and fine-tune process recipes. This task is complicated by the fact that recipes must often balance ill-defined process objectives and equipment limitations that are not clearly understood. For high volume production, a recipe must also yield results that are insensitive to process and equipment variations (Fig. 7.1).

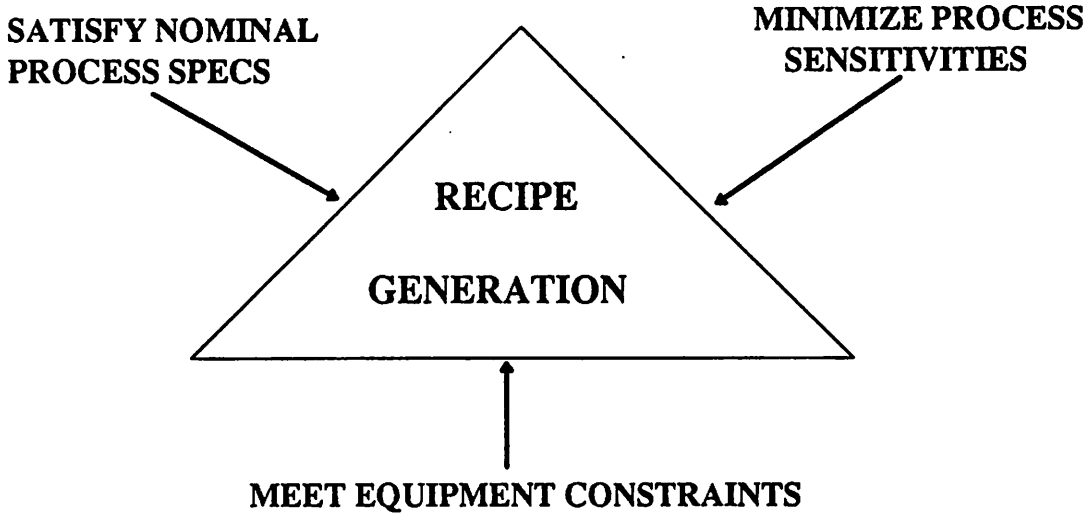


Fig. 7.1 Multiple objectives in recipe generation.

In this chapter we describe a formal, systematic methodology that facilitates the task of recipe design. This is achieved by employing statistical equipment-specific

models [1,2] in conjunction with interactive response surface exploration [3] and automatic numerical optimization [4]. A specially designed interface helps the users deal with quantitative and qualitative objectives and lets them explore the process design space. If the original objectives cannot be achieved, the recipe generator guides the user to resolve the conflicts and find settings that offer the best compromise between process objectives and equipment limitations.

This methodology has been incorporated in the Berkeley Computer Aided Manufacturing (BCAM) system and it is currently being used to synthesize recipes for LPCVD of undoped polysilicon. The system generates recipes that meet objectives related to the average values of film properties and also to their uniformities and repeatabilities. The generator is implemented on top of a powerful numerical optimizer [5,6], a flexible contour plot program and a graphic user interface built using PICASSO [7], a generic graphic toolkit that runs on X11 window primitives. Experimental use of the recipe generator has resulted in efficient, robust recipes generated in a fraction of the time required to create recipes by trial and error. This methodology is now being extended to plasma etching processes.

7.2. Defining Process Objectives for LPCVD of Undoped Polysilicon

In order to generate meaningful recipes, clear and concise objectives should be stated both at the process and equipment levels. The objectives are then used to either form a numerical optimization problem, or to simply explore the response surfaces[1] of the process. A good set of objectives should deal not only with the nominal process specifications, but also with the process uniformity and consistency. The operational

[1] The *response surface* of a process depicts geometrically the characteristics of the process with respect to its input variables. Response surfaces in more than three dimensions must be projected into three-dimensional (3D) or two-dimensional (2D) spaces.

limitations of the equipment should also be taken into account. A set of process and equipment objectives is defined below for LPCVD of undoped polysilicon:

Satisfy the Quantitative Process Specifications:
(1) Film thickness - average thickness - wafer-to-wafer thickness uniformity - within-wafer thickness uniformity
(2) Film stress - average stress - wafer-to-wafer stress uniformity - within-wafer stress uniformity
(3) Reproducibility of the above

Comply with Equipment Limitations
(1) Minimum deposition pressure attainable by the mechanical pump at maximum pump speed.

Satisfy the Qualitative Process Specifications:	
(1)	film quality
(2)	grain size
(3)	refractive index
(4)	texture
(5)	step coverage

7.3. Recipe Generation via Interactive Response Surface Exploration

An optimal recipe can be generated by exploring response surfaces of the process. In general, the understanding of a process with multiple inputs and outputs can be greatly enhanced by their response surfaces, and in particular by the use of *contour diagrams*. There are two types of contour diagrams: the *contour surfaces* and *contour lines*. The contour surfaces are obtained by projecting the response surfaces into a three-dimensional (3D) space, while the contour lines are obtained by projecting the response surfaces into a two-dimensional (2D) space. While a 3D projection is useful, the 2D contour lines lend themselves better to precise interaction for process optimization.

The 2D contour lines are extremely useful for process optimization, especially for process with a few (3 or 4) process parameters. For a process with a single output response, the 2D contour diagram can quickly provide information for locating the approximate optimal response by varying two process parameters at a time.

For processes with multiple output responses, the contour diagrams of all the process models can be superimposed in a common design space [3]. The advantages of

overlaying the contours are that it: (1) clearly shows whether and how complex requirements can be met for every response, (2) allows new conditions to be determined if objectives change, and (3) indicates in what direction process changes should be made in order to meet the objectives.

The contour diagrams, drawn according to process models fitted by statistical regression analysis, represent the average process response surfaces of a noisy system. To better depict the true response surfaces with account of the process/equipment capabilities, the prediction intervals around the average response surfaces should be computed from the ANOVA tables (Ch.5) [8] and presented together with the contour diagrams.

For processes with many (more than 3 or 4) process parameters and output responses, contour diagrams are not as effective for process optimization [3,9]. An automatic numerical optimization method can be employed to simultaneously satisfy many desired process objectives, as will be shown later in Section 7.4.

7.3.1. Functionality of the Response Surface System

During the interactive response surface exploration, the user selects any two process parameters as variables on the x- and y-axis, and fixes the values of the remaining process parameters. Once the cross-sectional projection of the response surfaces is defined, a maximum of four contour diagrams will be shown on the screen to represent four different process objectives: the *deposition rate of the first wafer*, the *within-wafer deposition rate uniformity*, the *average film stress*, and the *film stress uniformity*.

Other quantitative objectives, such as the *wafer-to-wafer deposition rate uniformity*, the *process sensitivities*, the *equipment capability*, as well as the qualitative objectives are more difficult to visualize with contour diagrams. Hence, they are not shown

on the current version of the recipe generator.

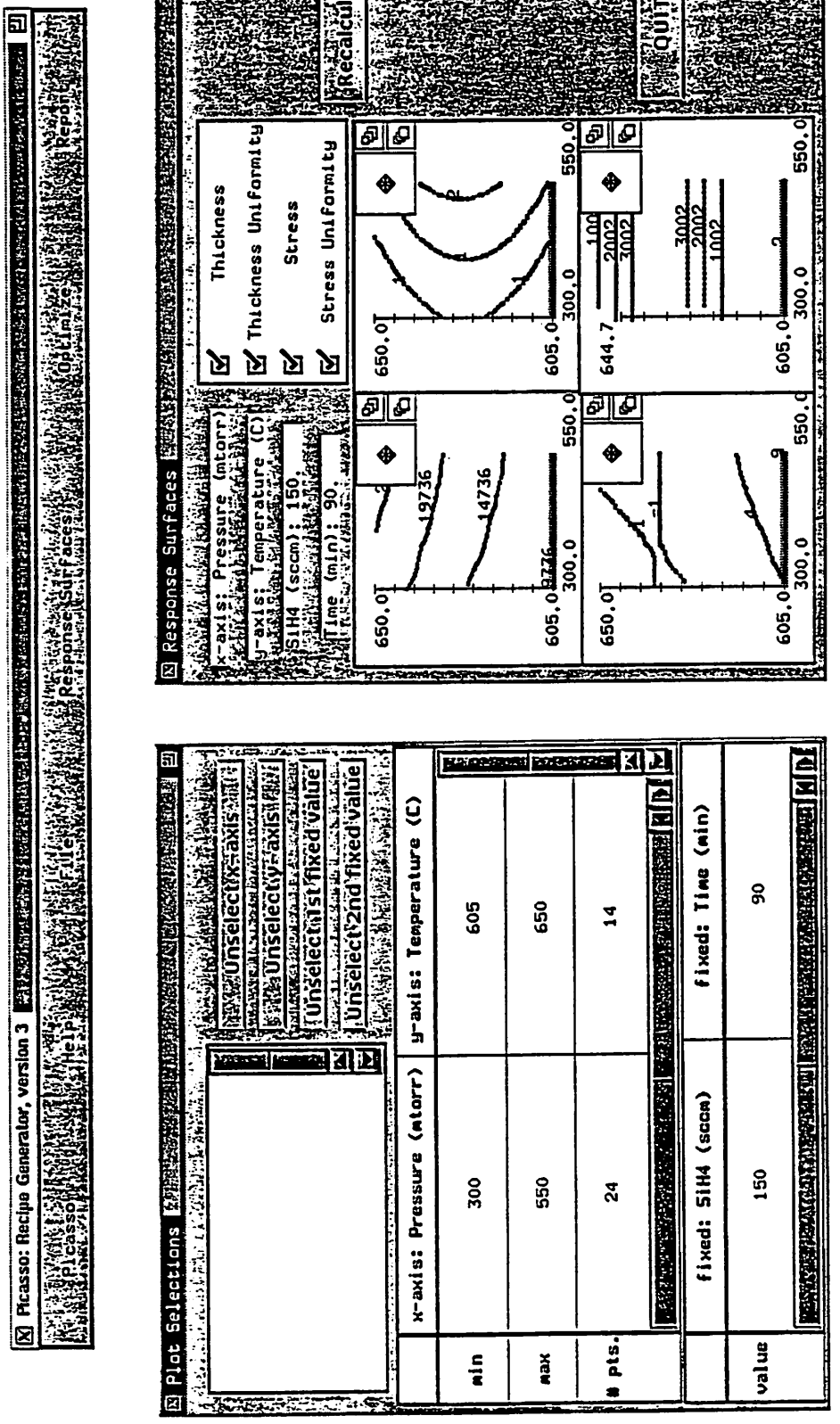
7.3.2. A Sample Session of Response Surface Exploration

A sample screen of the interactive response surface exploration is shown in Fig. 7.2. In the upper window, the user defines the 2D projection of the process response surfaces. In this example, "temperature" is selected as a variable on the y-axis with a range from 605 °C to 650 °C; "pressure" is selected as a variable on the x-axis with a range from 300 *mtorr* to 550 *mtorr*. The rest of the process parameters, "SiH_4 flow" and "deposition time", are fixed at 150 *sccm* and 90 *minutes* respectively.

A set of 4 contour plots, representing the response surfaces of average *polysilicon thickness*,, the *within-wafer thickness uniformity*, the *polysilicon stress*, and the *stress uniformity*, are drawn according to the defined projection. A user can then "navigate" through the design space simultaneously for all the four contour diagrams using the *panning* option at the upper left corner of the "thickness" plot. Once an interesting region is identified, a user can use the *zooming* option to exam in detail a particular process window. A tentative "recipe" is created when the user finds a setting that yields satisfactory results for all four process objectives.

In a future version of the recipe generator, estimations of the qualitative objective (Ch. 3.5) and the wafer-to-wafer deposition rate profile will be presented after a tentative recipe is created. A recipe will be final when all the qualitative and quantitative objectives are either met or selectively relaxed by the operator.

Fig. 7.2 Sample screen of response surface exploration session.



7.3.3. The Implementation of the Response Surface System

The interactive recipe generator is implemented by integrating the following two software modules:

(1) **PICASSO:** PICASSO is an object-oriented toolkit for developing graphical user interfaces to general application programs.

The toolkit includes a collection of predefined generic *gadgets* and *widgets*[2], such as simple bit-map and pix-map widgets, complex scrolling text, scrolling table, 2D and 3D plotting widgets. In addition, a variety of other components are provided that allow users to interact with an application program using buttons, check boxes, and menus (pull-down, pop-up, etc.). The parameters (e.g. the layout, font-size, and etc.) of these graphical components can be specified and modified in an *object-oriented* fashion via the corresponding property slots [10]. The use of a high-level interface toolkit, such as PICASSO, accelerated the implementation of the recipe generator. PICASSO is written using the Common Lisp Object-Oriented System (CLOS), Common Lisp (CL) and the X Window primitives.

(2) **CONTOUR:** A C program, refined from the program CONTOUR [11], is used to generate the Cartesian coordinates of 2D contour diagrams of the projected process response surfaces. The implementation details of the program are outlined in Appendix 2 at the end of this chapter.

(3) **The Lisp-C foreign function interface:** Four Lisp functions are written to allow foreign function interactions between the PICASSO user interface, (implemented in Common Lisp), and the CONTOUR program (implemented in C). The

[2] *Widgets and gadgets* are user interface abstractions. A gadget is an abstraction for an output port behavior (e.g. text-gadget). A widget is the abstraction for input port behavior (e.g. menu-widget) [7].

implementation details are summarized in Appendix 2.

7.4. Automatic Recipe Generation via Numerical Optimization

The second approach to recipe generation is to cast the process and equipment objectives into an optimization problem, which in turn can be solved by a numerical optimizer. This approach offers an effective alternative to process optimization when there are many process parameters and responses. In automatic recipe synthesis, values or ranges of the desired process results (objectives) are used to generate the values of the process parameters. One advantage of the automatic synthesis of recipes is the automation of the search towards meeting a set of multiple and possible conflicting process objectives. Using this method, a user can quickly decide whether the optimization problem is feasible. If it is not, the user can iteratively adjust the process specifications until the problem becomes feasible.

7.4.1. Functionality of the Optimization-based System

During the automatic synthesis of LPCVD recipes, the user defines constraints for the quantitative objectives such as *film thickness, stress*, and their respective *uniformities*. These quantitative objectives will be cast as the cost function of the optimization problem. The user also defines qualitative goals, such as *film smoothness, grain size, refractive index, texture* and *step coverage*. System recommendations, based on the qualitative models developed in Ch. 3, will be issued in order to help meet the qualitative objectives.

The equipment limitations, as well as the acceptable ranges of the process parameters, are cast as numerical constraints in the optimization problem. The outcome of the optimization, i.e. the optimized values of the equipment settings and responses, are

reported back to the user. A thickness profile versus the wafer position in the reactor is then plotted according to the generated recipe. Statistical confidence intervals (at the 90% level) are shown for all output results.

7.4.2. A Sample Session of Automatic Recipe Generation

A sample screen on the automatic recipe synthesis session is shown in Fig. 7.3. In the upper left window, the user specifies bounds for the quantitative objectives. The ranges of the objectives are summarized below:

objectives	low	high
thickness	5000Å	7000Å
thickness uniformity	0%	5%
stress	-10^9 dyne/cm^2	10^9 dyne/cm^2
stress uniformity	0%	5%

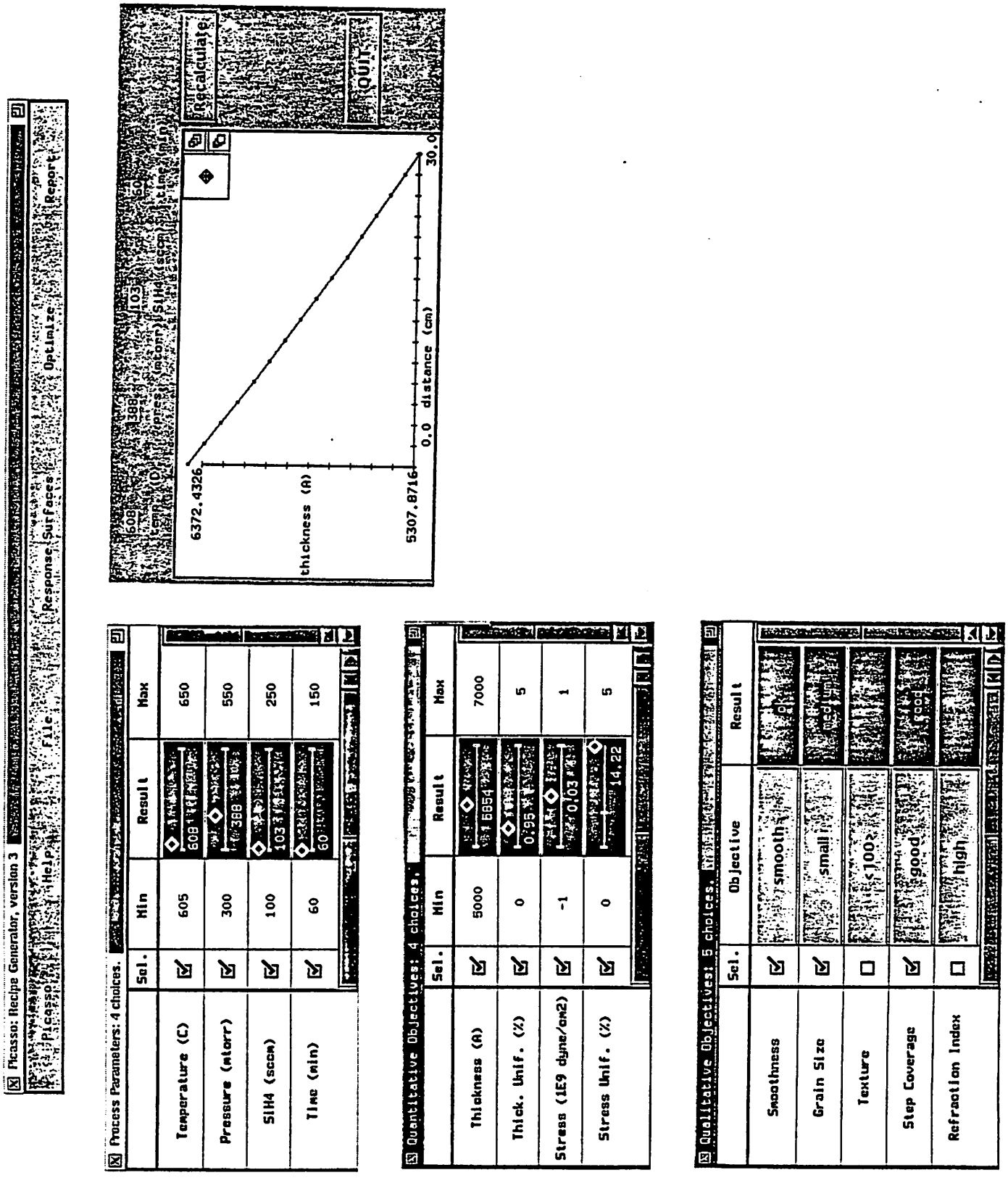
In the middle left window, the user specifies the qualitative objectives. The goals of *film smoothness, grain-size, and step-coverage* are set to *smooth, small,* and *good* respectively. The limits of the process parameters, such as *temperature, pressure,* SiH_4 *flow, and time,* are set at the default values shown in the lower left window. After the solution of the numerical optimization problem, the results of the generated recipe are displayed. The quantitative results, based on the optimization of the statistical models, are shown by means of a sliding scale pointer between the bounds. The qualitative results, based on the simulation of the qualitative models, are shown in the "result"

column of the window. A thickness profile is shown as a function of wafer position in the upper right window of the system.

A first iteration of the recipe generation indicates that the *thickness*, *thickness uniformity*, and *stress* objectives are easily satisfied. The *stress uniformity* is however very sensitive and difficult to control with respect to the equipment settings. With this piece of information regarding the feasibility of the process objectives, the user can move the acceptable bounds on the *stress uniformity* from 0 to 20 %. At the same time, the acceptable bounds on *thickness* can be set from 5000 to 6000 Å, *thickness uniformity* from 0 to 3%, and *stress* from -0.5 and 0.5×10^9 dyne/cm^2. The bounds on the process parameters remain the same. All the process objectives are satisfied easily in the second attempt. The stress uniformity has drastically improved from around 15% to 9%. Usually a recipe can be finalized after a few iterations.

The system recommendations regarding the qualitative objectives, as well as the statistical confidence intervals, will be incorporated in a future version of the recipe generator.

Fig. 7.3 Sample screen of automatic recipe generation session.



7.4.3. The Implementation of the Optimization-based System

The automatic recipe synthesis system consists of the following two software modules:

(1) **PICASSO**: Different generic components of the PICASSO toolkit are used to construct the user interface of automatic recipe generation. The "table-matrix" widgets are used to place the process objectives and parameters on a scrollable table for effective use of the display screen area. The "meter" widgets are used to implement sliding scale pointers between the limits of quantitative objectives and parameters. The "qual" widgets are used to implement pull-down menus for selecting the goals of qualitative objectives. A "plot" widget is used to display the final thickness profile versus the wafer position in the furnace.

(2) **NAG**: In order to solve the minimization problem, we employ the general numerical optimizer E04UCF from the public-domain NAG (Numerical Algorithm Group) library [5,6]. E04UCF is a Fortran routine designed to minimize an arbitrary smooth function subject to constraints, including simple bounds on the variables, linear constraints and smooth nonlinear constraints. E04UCF may also be used for unconstrained, bound-constrained and linearly constrained optimization. The user must provide subroutines that define the objective and constraint functions and as many of their first partial derivatives as possible. Unspecified derivatives are approximated by finite differences. Since all matrices are treated as dense, E04UCF is not intended for large sparse problems. E04UCF uses a sequential quadratic programming (SQP) [12] algorithm in which the search direction is the solution of a quadratic programming (QP) problem. The Fortran subroutine, which utilizes EO4UCF to solve the minimization problem with the given boundary conditions, is outlined in Appendix 2.

The cost function to be minimized in Eq. (7.1) is the weighted sum of the squared differences between the process objectives and their target values (Fig. 7.4). The limits of each process objective implicitly reflect the relative importance of the objective in the cost function [4, 13]. Hence, the difference of the limits is used as an appropriate weighting/normalizing factor w_i for each objective in the cost function.

$$\text{COST FUNCTION} = \min_{\mathbf{p}} \sum_{i} w_i \, (R_i(\mathbf{p}) - \hat{R}_i)^2 \tag{7.1}$$

where $\mathbf{p}$ is the process parameter vector, R_i is the i-th process response, $\hat{R}_i$ is the target value of the i-th process responses, and w_i is the weighting/normalizing factor for the respective process response.

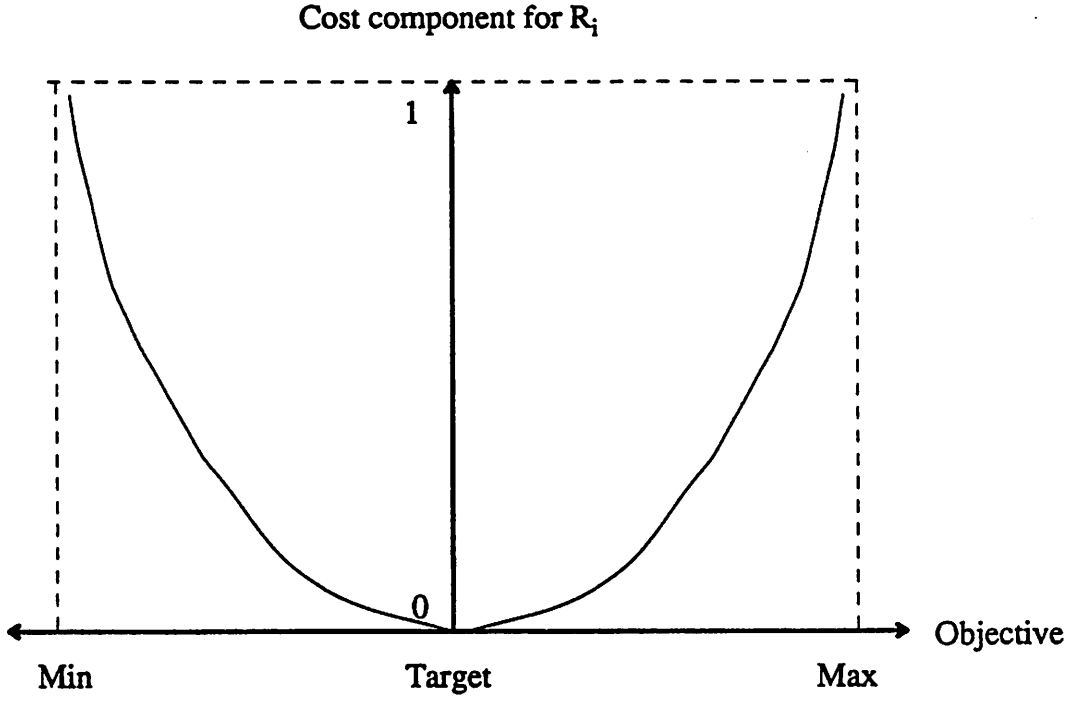


Fig. 7.4 A normalized quadratic cost component for R_i.

For the automatic recipe generation of LPCVD of undoped polysilicon, the cost function is defined as follows:

$$\text{COST FUNCTION} = \left[\frac{\text{Thickness of first wafer} - \text{Max. thickness}}{0.5 \times (\text{Max. thickness} - \text{Min. thickness})}\right]^2 + \left[\frac{\text{Thickness of last wafer} - \text{Min. thickness}}{0.5 \times (\text{Max. thickness} - \text{Min. thickness})}\right]^2 + \left[\frac{\text{Thickness unif.} - 0.5 \times (\text{Max. thickness unif.} + \text{Min. thickness unif.})}{0.5 \times (\text{Max. thickness unif.} - \text{Min. thickness unif.})}\right]^2 + \left[\frac{\text{Stress} - 0.5 \times (\text{Max. stress} + \text{Min. stress})}{0.5 \times (\text{Max. stress} - \text{Min. stress})}\right]^2 + \left[\frac{\text{Stress unif.} - 0.5 \times (\text{Max. stress unif.} + \text{Min. stress unif.})}{0.5 \times (\text{Max. stress unif.} - \text{Min. stress unif.})}\right]^2 \quad (7.2)$$

Due to the depletion effect of the reactant SiH_4 gas (Ch.3), the maximum deposition rate is usually observed on the wafer that is closest to the gas inlet. Similarly, the minimum deposition rate is observed on the wafer that is furthest from the gas inlet. Hence, the target value for the *thickness* of the first wafer is the upper bound of *thickness*; and the target value for the *thickness* of last wafer is the lower bound of *thickness*. The targets for *thickness uniformity*, *stress*, and *stress uniformity*, are the averages of their specified bounds. The difference of the lower and upper bounds of the process objectives are used as their corresponding weighting/normalizing factors in the cost function.

In order to improve the convergence of the optimization problem, the process objectives are normalized to similar numerical ranges [14]. For example, the linearized thickness model of Eq. (5.5) is used for optimization in place of the full model of Eq. (5.6).

In order to improve the repeatibility of the process, we must operate in a region that will be least sensitive to random equipment variations. The cost factor, to be added to the cost function in Eq. (7.2), is shown below:

$$\text{COST FACTOR} = \min_{\mathbf{p}} \sum_{i} \left[\mathbf{w}_i \cdot \mathbf{R}_i'(\mathbf{p}) \right]^2 \tag{7.3}$$

where $\mathbf{w}_i$ and $\mathbf{R}_i'(\mathbf{p})$ are vectors defined as follow:

$$\mathbf{w}_i = \left[\frac{\Delta T}{0.5 \times (R_i^{max} - R_i^{min})}, \frac{\Delta P}{0.5 \times (R_i^{max} - R_i^{min})}, \frac{\Delta Q}{0.5 \times (R_i^{max} - R_i^{min})}, \frac{\Delta t}{0.5 \times (R_i^{max} - R_i^{min})} \right]$$

$$\mathbf{R}_i'(\mathbf{p}) = \left[\frac{\partial R_i(\mathbf{p})}{\partial T}, \frac{\partial R_i(\mathbf{p})}{\partial P}, \frac{\partial R_i(\mathbf{p})}{\partial Q}, \frac{\partial R_i(\mathbf{p})}{\partial t} \right]$$

Here ($\cdot$) is the inner product operator and $\mathbf{p}$ is the vector containing the process parameters T, P, Q, t. ΔT, ΔP, ΔQ, Δt are the estimated standard deviations of the equipment settings. R_i^{max},R_i^{min} are the limits of a process responses. Finally, $\partial R_i/\partial T$, $\partial R_i/\partial P$, $\partial R_i/\partial Q$, $\partial R_i/\partial t$ are the partial derivatives of a process response with respect to the process parameters.

The cost factor in Eq. (7.3) tries to minimize the sensitivity of the process objectives with respect to the equipment noise. The weighting/normalizing factor $\mathbf{w}_i$ reflects the significance of each objective in the cost function.

The feasibility of optimization depends strongly on the definition of the numerical constraints [6,14]. With this in mind, only the necessary boundary conditions of equipment settings and equipment performances should be included as numerical constraints. This consideration will make it more likely that the problem has a feasible solution.

For the recipe generation of LPCVD, the numerical constraints are defined as follows:

(1) Process parameter (equipment setting) constraints

$$T_{min} < T < T_{max}$$

$$P_{min} < P < P_{max}$$

$$Q_{min} < Q < Q_{max}$$

$$t_{min} < t < t_{max}$$

(2) Equipment performance constraints (due to pump speed limitations)

$$P > P_{min}(T, Q)$$

where:

$$P_{min}(T, Q) = 413.5 - 0.9(T-883) - 1.5(Q-175) - 0.005(T-883)(Q-175) \text{ in } \textit{mtorr}.$$

An initial estimate of the solution is required by the optimization algorithm. The estimated solution does not have to be feasible. At the beginning of a session, the system uses the center points of the equipment settings as the first estimate. In subsequent iterations, the initial guess is the solution of the last feasible problem.

(3) **Qualitative Constraints**: The system issues recommendations to let the users manually adjust the boundary conditions of the equipment settings in order to satisfy the qualitative objectives. The recommendations are based on the qualitative models discussed in Ch.3. For example, in order to achieve "small" grain-size, "low" values for the temperature (T), silane flow (Q), time (t) and a "high' value for pressure (P) should be used. As another example, in order to achieve smoother film quality both the temperature and pressure should be kept at "low" settings.

The qualitative models need to be further refined for integration with the numerical optimization problem. Possibilities for future research are discussed in the last chapter.

7.5. A Simple Algorithm for Deposition Rate Uniformity

The statistical models, which we developed for use in recipe generation, are valid under the condition of a uniform deposition temperature, a uniform deposition pressure and a front-end injected reactant SiH_4 gas flow. If the temperature is constant, the depletion of the reactant SiH_4 will result in a decrease in deposition rate along the length of the furnace tube. A common solution to this "depletion effect" is to use a rising temperature profile from the gas inlet (load zone) to the gas outlet (exhaust zone) by using several heating elements at different zones of the furnace [15]. The increased reaction rate resulting from the temperature differential is intended to compensate for the SiH_4 depletion. This temperature ramp, however, should be used with caution as it will affect other polysilicon film properties (such as the residual stress, grain size and etc.), which strongly depend on the reaction temperature.

We describe below a simple and general method for improving the wafer-to-wafer uniformity of the deposition rate. The focus is on specifying the required temperature ramp at any deposition furnace with a front-end reactant gas inlet, such as those for undoped polysilicon, phosphorus-doped glass and silicon nitride. The uniformity algorithm estimates the required temperature ramp using a simplified physical deposition rate model, process heuristics, and data from a single calibration experiment. The algorithm has shown satisfactory experimental results.

7.5.1. The Simplified Deposition Rate Model

The deposition rates for most LPCVD deposition processes are exponentially dependent on temperature. These deposition rate models, such as the LPCVD model of Eq. (3.20), take the form of an Arrhenius equation [15]:

$$R = C(Q,P,z)\exp(\frac{-\Delta E}{kT}) \tag{7.4}$$

where R is the deposition rate and C(Q,P,z) is the *frequency factor* that depends on the SiH_4 flow, the deposition pressure and the wafer position. ΔE is the *activation energy* of the process in *eV or kcal/mole*, which can be estimated using experimental data (Ch. 4, 5). Finally, R is the universal gas constant given by 1.98719 *cal/mol-K* and T is the deposition temperature in K, which can be position-dependent.

The temperatures of the heating elements, located along the furnace, are the control parameters in achieving the deposition rate uniformity. All other process parameters, such as the deposition pressure and the reactant gas flow, are fixed at the values of the calibration experiment.

7.5.2. The Calibration Experiment

The calibration experiment should use a uniform deposition temperature. The wafer-to-wafer deposition rate uniformity of the calibration experiment can be improved by applying a temperature ramp, with no change in the deposition pressure and reactant gas flow.

7.5.3. Estimation of the Temperature Ramp

The following procedure estimates the center-zone temperature (T_C) of a three-zone deposition furnace. Estimation of the temperatures for the load and exhaust zones, (T_L and T_E), can be carried out in a similar manner.

(1) First the wafer-to-wafer deposition rate of the calibration experiment (or any other run with a uniform temperature profile) is plotted against the wafer position as in Fig. 7.5. Next, the desired deposition rate with the required wafer-to-wafer uniformity is plotted against the wafer position as in Fig. 7.6.

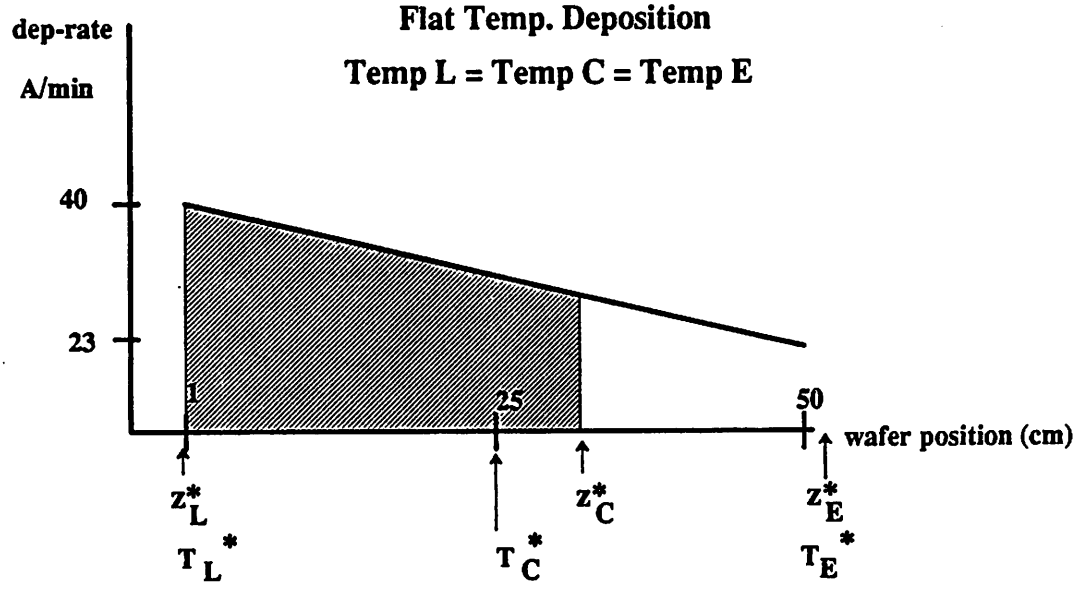


Fig. 7.5 Deposition rate profile under flat temperature distribution.

(2) Since the center-zone temperature (T_C) is currently being estimated, the "relative deposition area" up to the center-zone (z_C) of Fig. 7.6 must be evaluated. The "relative deposition area" is defined as the area under the deposition rate curve (shaded area of Fig. 7.6). Next, the wafer position (z_C^*) with the same "relative deposition area" in the calibration experiment of Fig. 7.5 is found.

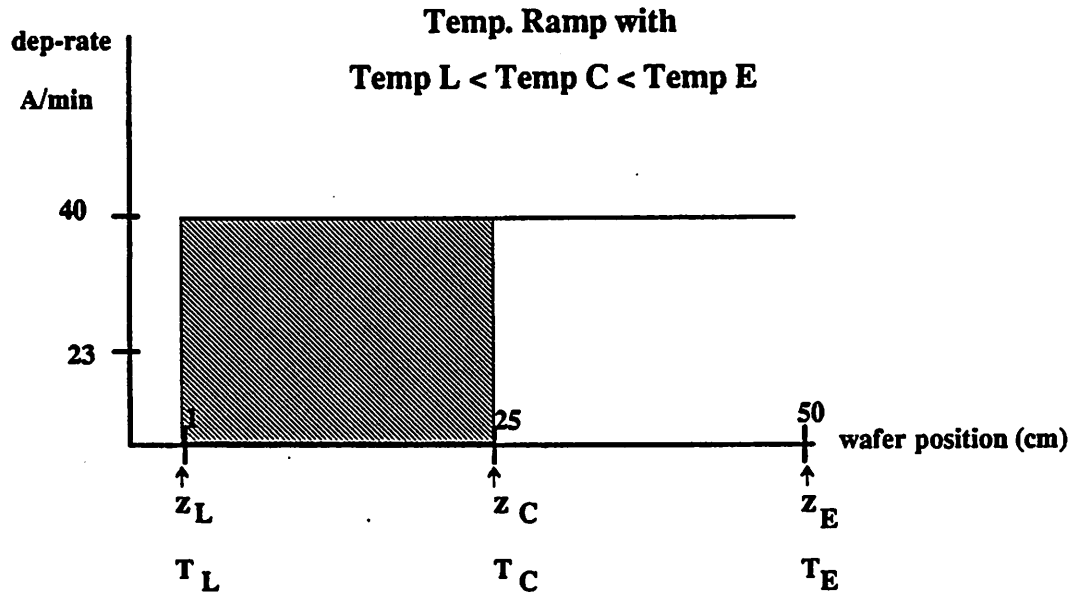


Fig. 7.6 Deposition rate profile under ramped temperature distribution.

Wafer positions with the same "relative deposition area" can be thought as having approximately "identical reacting environments", i.e. equal *frequency factors* C(Q,P,z).

(3) Now we calculate the ratio of the deposition rates $R(z_C^*)$ and $R(z_C)$ for wafer positions with "identical reacting environments" at locations z_c and z_c^*.

$$\frac{R(z_C^*)}{R(z_C)} = \frac{C(Q,P,z_C^*)\exp(\frac{-\Delta E}{kT_C^*})}{C(Q,P,z_C)\exp(\frac{-\Delta E}{kT_C})} \tag{7.5}$$

where:
$R(z_C^*)$ is the deposition rate at z_C^*, which can be estimated from the calibration experiment
T_C^* is the uniform temperature of the calibration experiment
$R(z_C)$ is the desired uniform deposition rate
T_C is the center-zone temperature to be estimated.
Q is the reactant gas flow, fixed at the value of the calibration experiment
P is the deposition pressure, fixed at the value of the calibration experiment

(4) Since the ratio is taken at two wafer positions with approximately "identical reacting environments", the $C(Q,P,z_C^*)$ of the numerator is canceled by the $C(Q,P,z_C)$ of the denominator in Eq. (7.5). Since the center zone temperature T_C is now the only unknown variable in Eq. (7.5), it can be estimated. Similar estimation of T_L and T_E will specify a rising temperature profile for uniform deposition rates across the furnace.

7.5.4. Experimental Verification

Two sets of experiments were performed to verify the algorithm. The first experiment involving the deposition of polysilicon was done at Advanced Semiconductor Materials/America, Inc. (ASM). The three-zone furnace consisted of 6 boats with a total of 150 wafers. A calibration run at a constant temperature of 610 °C with 24 sample test wafers was performed first. The wafer-to-wafer deposition rate of the calibration run was plotted against the wafer position in Fig. 7.7. The temperature ramp (610, 617, 623 °C) calculated from this algorithm agrees well with the actual manufacturing temperature ramp (610, 616, 620 °C) that the process engineers have empirically found and used to achieve a ± 5% wafer-to-wafer thickness uniformity. The overall process conditions for the ASM experiment are listed in Table 7.1.

No. of sample wafers:	24 (4" wafer)		
Wafer load:	150 (25/boat, 6 boats)		
Time:	39 mins		
Pressure:	600 mtorr		
SiH_4:	300 sccm		
Cali. temp:	610°C (load)	610°C (center)	610°C (exhaust)
temp. ramp:	610°C (load)	616°C (center)	620°C (exhaust)

Table 7.1 Processing condition at ASM

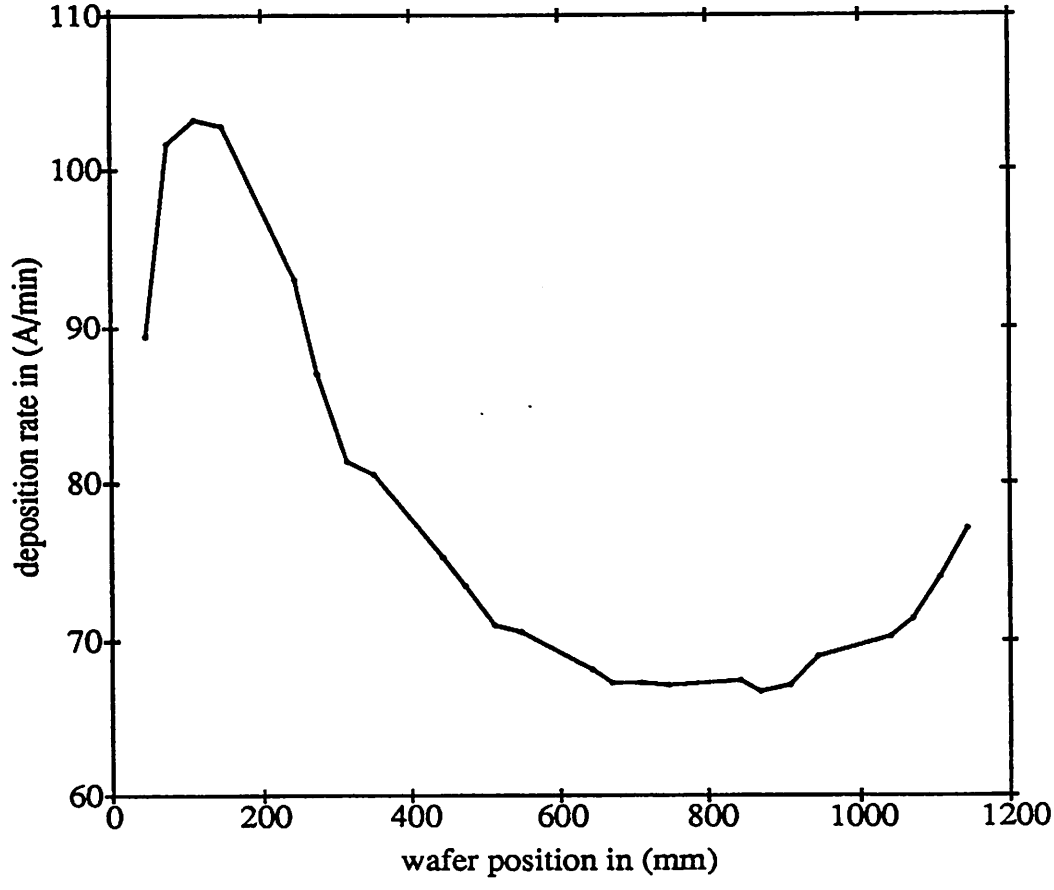


Fig. 7.7 Deposition rate profile for the ASM calibration experiment

A second set of experiments involving the deposition of amorphous silicon was done at our microfabrication laboratory in Berkeley (UCB). The three-zone furnace consisted of 2 boats with a total of 26 wafers. Again a calibration run at a constant temperature of 595 °C with 8 sample test wafers, was performed. The temperature and deposition rate profiles of the calibration run, plotted against the wafer position, are shown in Fig. 7.8 and 7.9. The predicted temperature ramp were 595, 597, and 600 °C across the 2 boats to achieve ± 5% wafer-to-wafer thickness uniformity. An in-situ twelve-point thermocouple was installed to accurately monitor the temperature profile. The processing conditions of the UCB experiment are listed in Table 7.2. The desired temperature profile (Fig. 7.10) was obtained across the 2 boats. A plot of polysilicon growth rate versus wafer distance (Fig. 7.11) showed that the thickness varied about 5%. The variation of thickness could be as high as 20% without any temperature ramp.

No. of sample wafers:	8 (4" wafer)		
Wafer load:	26 (13/boat, 2 boats)		
Time:	60 mins		
Pressure:	300 mtorr		
SiH_4:	120 sccm		
Cali. temp:	595°C (load)	595°C (center)	595°C (exhaust)
temp. ramp:	595°C (load)	597°C (center)	600°C (exhaust)

Table 7.2 Processing condition at UCB

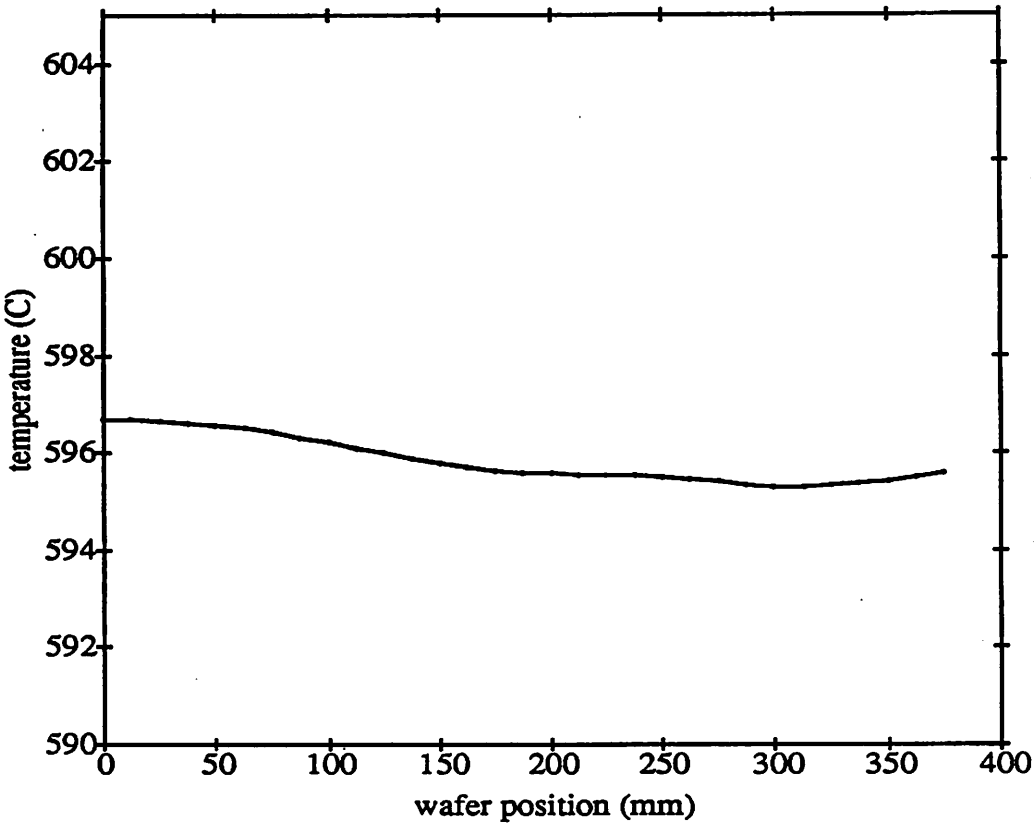


Fig. 7.8 Temperature profile of the UCB calibration experiment

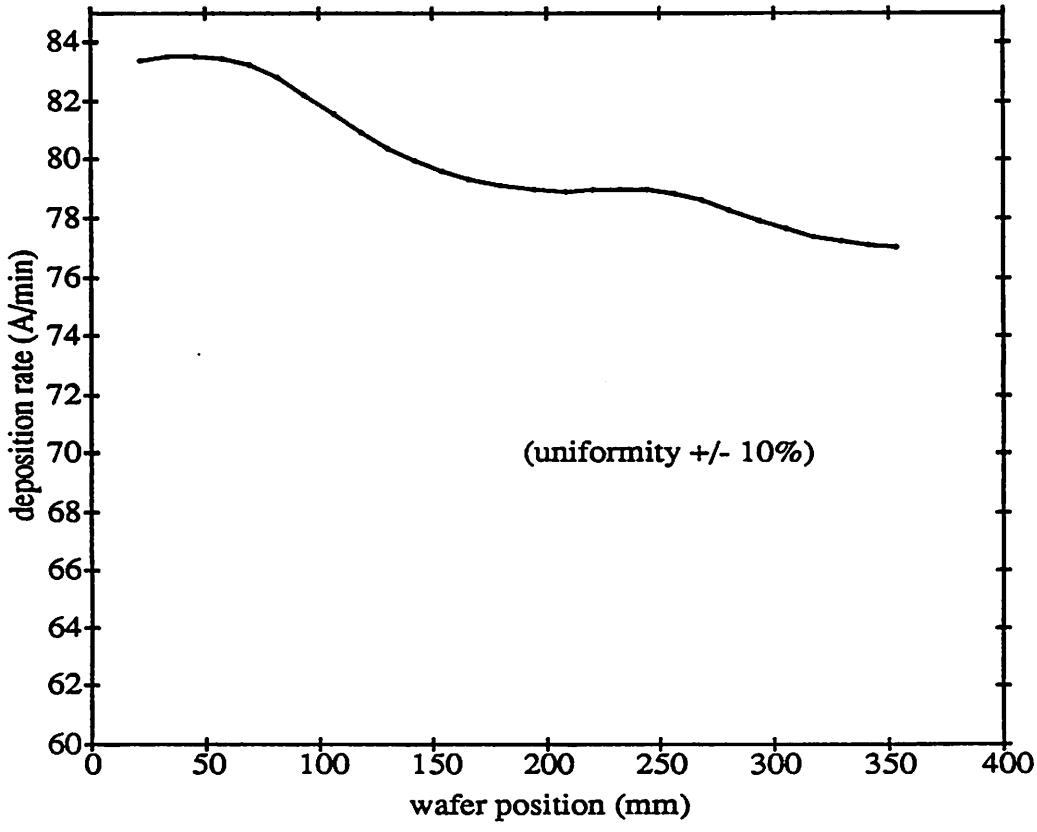


Fig. 7.9 Deposition rate profile of the UCB calibration experiment

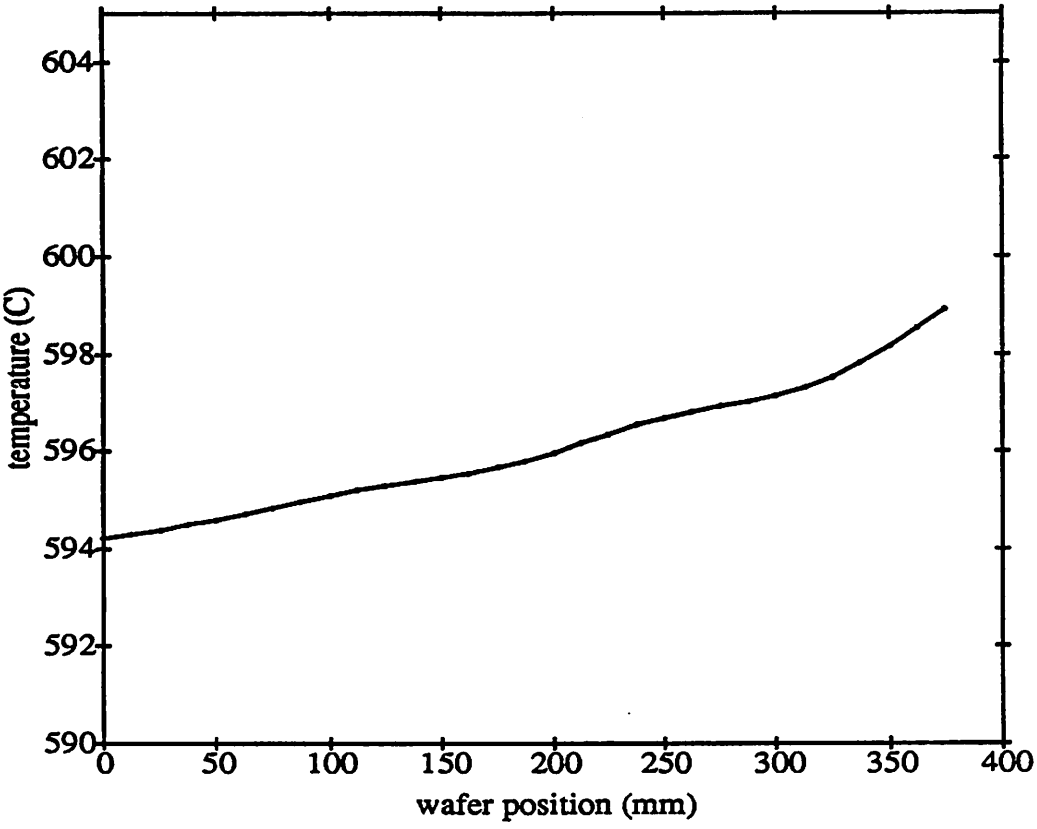


Fig. 7.10 Temperature profile of the UCB verification experiment

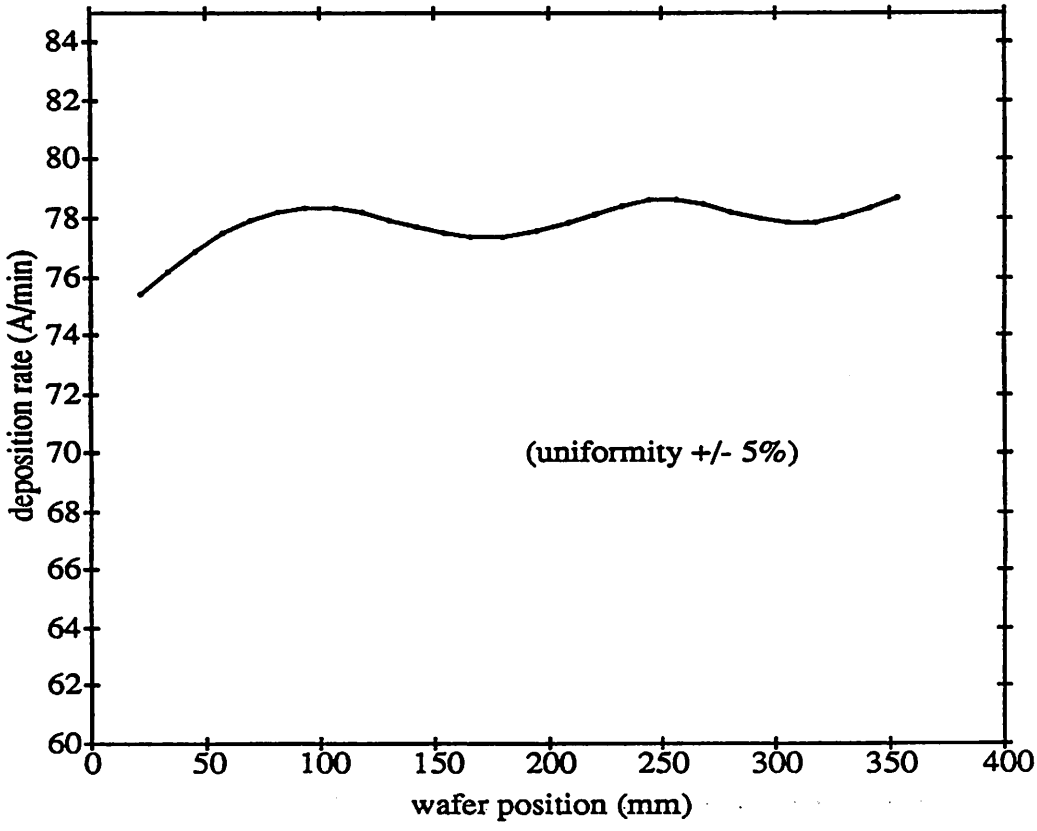


Fig. 7.11 Deposition rate profile for the UCB verification experiment

7.5.5. Other Uniformity Algorithms

Other algorithms are available to resolve the wafer-to-wafer thickness uniformity problem. They include:

(1) Using the deposition rate model described in Eqs. (3.18) to (3.21) in a *piecewise-linear* fashion [16-18] for simulation under a particular temperature profile. The desired temperature ramp can be estimated after several simulations.

(2) Expressing the temperature ramp T(z) as a function of position, which is subsequently included in the integral of Eq. (3.12) and (3.13). An analytical deposition rate model with the temperature ramp function can be computed and used to estimate the thickness uniformity.

(3) Employing sophisticated mass transport and chemical kinetics [19] to simulate the reacting environment of the furnace.

(4) Applying systematic statistical experimental design methods to minimize the process variation [20].

(5) Designing new furnaces to incorporate distributed gas injection, or vertical gas flow, to eliminate the mass depletion effect [15].

7.5.6. Summary for the Uniformity Algorithm

A simple and effective algorithm is presented to resolve the wafer-to-wafer deposition rate uniformity problem of the horizontal front-end injected LPCVD furnace. In this algorithm, a simplified physical deposition rate model is combined with several process heuristics and one calibration experiment to estimate the required temperature ramp. This algorithm works under the mild assumption that depletion of SiH_4 is less than 20% at the outlet of the furnace, which is usually a reasonable assumption for most

LPCVD systems. The accuracy of the algorithm is limited by the ability of the furnace to realize the designed temperature profile. Lack of accurate control due to thermal flow between zones and the limited number of independent heating zones can limit the effectiveness of this approach.

The uniformity algorithm should be used carefully as the designed temperature ramp can affect other polysilicon film properties (such as the residual film stress, grain size, film quality, and etc.), which are highly dependent on the deposition temperature [15]. This algorithm will be incorporated in a future version of the recipe generator.

REFERENCES

[1] K.K. Lin, J. Huang, and C. Spanos, "Statistical Equipment Modeling for VLSI Manufacturing", *Proceedings of the Fifth Symposium on Automated Integrated Circuits Manufacturing*, pp. 127-138, Proceedings Volume 90-3, The Electrochemical Society, October 1989.

[2] C. Spanos "Special Issues in Semiconductor Manufacturing", *ERL Memorandum No. UCB/ERL M90/8*, Department of EECS, University of California, Berkeley, January 13, 1990.

[3] G. E. Box, W. G. Hunter, J. S. Hunter, "Statistics for Experimenters - An Introduction to Design, Data Analysis, and Model Building", John Wiley & Sons, 1978.

[4] J.M. Shyu, "Performance Optimization of Integrated Circuits," *ERL Memorandum No. UCB/ERL M88/74*, Department of EECS, University of California, Berkeley, November 22, 1988.

[5] P. Gill, W. Murray, M. Saunders, M. Wright, "User's Guide for SOL/NPSOL: A Fortran Package for Nonlinear Programming," Report SOL 83-12. Dept of Operations Research, Stanford University, CA 94305, 1983.

[6] The NAG Fortran Library Routine Document for E04UCF, 1986.

[7] L. Rowe, and et. al., "PICASSO Reference Manual," Computer Science Division - EECS, University of California, Berkeley, CA 94720, December 1989.

[8] N. R. Draper, H. Smith, "Applied Regression Analysis", John Wiley & Sons, Second Edition, 1981.

[9] G. Box, N. Draper, "Empirical Model-Building and Response Surfaces", John Wiley & Sons, 1987.

[10] Sonya E. Keene, "Object-Oriented Programming in Common Lisp: A Programmer's Guide to CLOS", Addison-Wesley Inc., 1988.

[11] K. Toh, User manual for "Contour plot program for X-windows, HP2648s and Postscript", University of California, Berkeley, January 1990.

[12] Mordecal Avriel, Nonlinear Programming: Analysis and Methods, Prentice-Hall, Inc., 1976.

[13] William T. Nye, "Delight: An Interactive System for Optimization-Based Engineering Design", *Memorandum No. UCB/ERL M83/33*, Univ. of California, Berkeley, CA, 31 May 1983.

[14] P. Gill, W. Murray, M. Wright, "Practical Optimization", Academic Press, London, 1981.

[15] S. Wolf, R.N. Tauber, "Silicon Processing; for the VLSI Era; Volume 1 - Process Technology", *Chapter 3 on Chemical Vapor Deposition of Amorphous and Polycrystalline Film*, Lattice Press, pp. 161-197.

[16] N. Cheung, "Thin Film Technology for IC Fabrication", EECS 290P course notes, University of California, Berkeley, Spring 1988.

[17] K. F. Roenigk, K. F. Jensen, "Analysis of Multicomponent LPCVD Processes", *J. of Electrochem. Soc.: SOLID STATE SCIENCE AND TECHNOLOGY*, Vol. 132 No. 2, p. 448, Feb. 1985.

[18] K. F. Jensen, D. B. Graves, "Modeling and Analysis of Low Pressure CVD Reactors", *J. of Electrochem. Soc.: SOLID STATE SCIENCE AND TECHNOLOGY*, Vol. 130 No. 9, p. 1950, Sept. 1983.

[19] FLUENT manual, version 2.9, Technical Note, Creare Inc., Feb. 1987.

[20] E. Sachs, R-S. Guo, S. Ha, and A. K. Hu, "Process Control System for VLSI Fabrication", *Proceedings of the Fifth Symposium on Automated Integrated Circuits Manufacturing*, Proceeding Volume 90-3, pp. 78-93, The Electrochemical Society, Inc., 1990.

CHAPTER 8
CONCLUSIONS AND FUTURE RESEARCH

8.1. Conclusions

The subject of this dissertation is a methodology for building *equipment-specific* process models. The effectiveness of this methodology is demonstrated through the application of a physically-derived analytical LPCVD model, in conjunction with a D-optimal statistical experimental design and standard regression analysis. The required number of characterization runs is small and it can be scheduled and completed in stages for the iterative refinement of the model. The final models describe both the nominal process responses and the associated process uniformities, and show good predictive capability.

The statistical equipment-specific process models have been incorporated in a system used for interactive response surface exploration and automatic numerical optimization. A friendly user interface helps generate equipment recipes that meet multiple process objectives. This methodology has been incorporated in a software module of the Berkeley Computer Aided Manufacturing (BCAM) system. The system generates recipes that meet objectives related to the average values of film properties, and also to their uniformities and repeatabilities. The generator employs a powerful numerical optimizer, a flexible contour plot C program, and a graphic user interface built in PICASSO.

8.2. Future Research

The methodology that we have developed so far can benefit from several improvements and extensions. Some of these are described below.

8.2.1. Modeling and Characterization

(1) **Qualitative models:** The use of fuzzy set and fuzzy logic for qualitative modeling is briefly demonstrated at the end of Ch.3. Given the rules may have different degrees of uncertainty, the use of *evidential reasoning* [1] should also be investigated.

(2) **Experimental design techniques:** Other optimal experimental design techniques, such as the G- and I-optimal designs, can be used to characterize a given model [2,3]. In an I-optimal design, characterization runs are geared to minimize the expected deviation[1] of the fitted model from the actual model. In a G-optimal design, characterization runs are generated to minimize the prediction error of the fitted model [4].

In our work, non-linear process models (such as the nominal deposition rate model) are linearized[2] in order to simplify the optimal experimental design for the characterization runs. An *ad hoc* sequential experimental design plan, which divided the second stage experiment into two sub-stages, was used to optimize the statistical accuracy of both the deposition rate and stress models. Hence, more advanced experimental techniques for handling situations with non-linear process models, and/or multiple models (responses) should also be explored [5].

---

[1] The deviation is expressed as the sum of the squared differences over the experimental points.

[2] *Linear* to the model coefficients. In general, our models are non-linear to the process parameters. Please refer to Ch. 4.

8.2.2. Recipe Generation

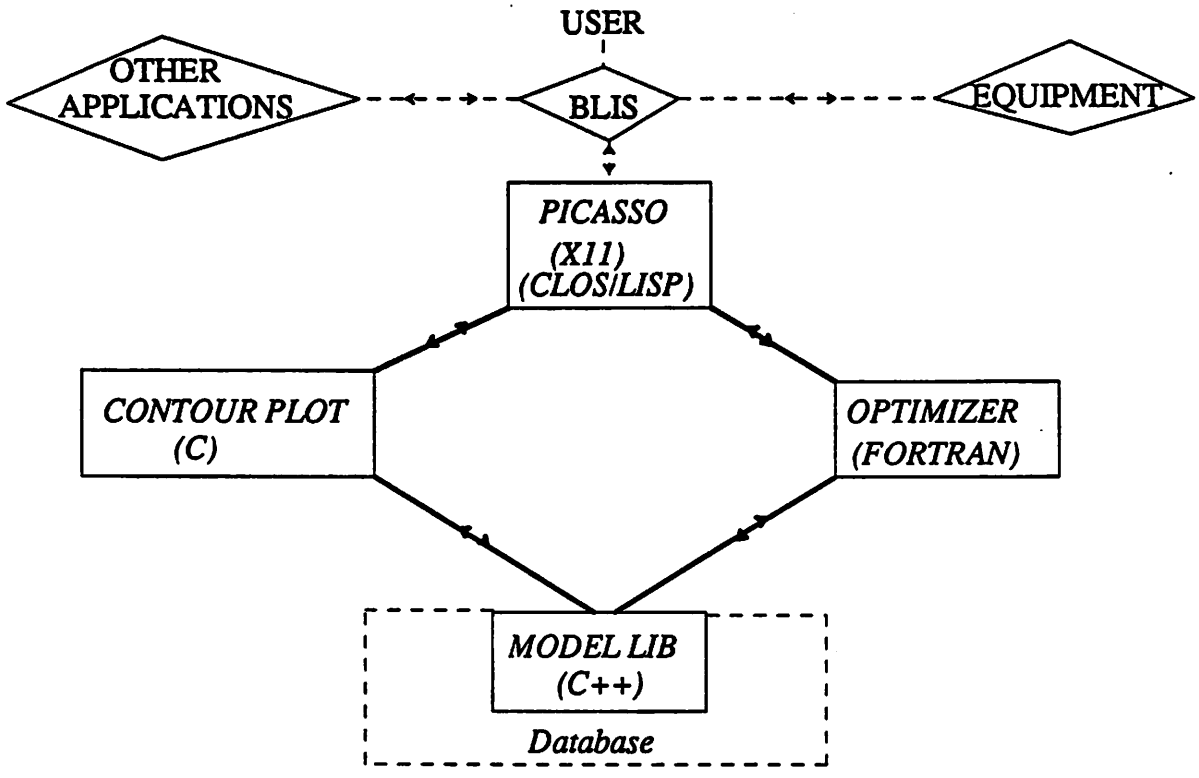
(1) **Acceptance in the clean room:** A prototype version has been released to potential users in the Berkeley Microfabrication Laboratory for evaluation. The system is favorably received by the users as a useful process design tool. A list of helpful user comments has been compiled and will be used for the improvement of future versions.

(2) **Qualitative constraints:** The qualitative models described in Ch.3 are solved in reverse (backward reasoning) to generate recommendations on meeting the qualitative process objectives. The backward reasoning rules for recipe generation are currently represented separately from the forward reasoning rules for simulation described in Ch.3. The backward reasoning rules are represented by simple symbolic matching between the specified process goals and the required qualitative process constraints. A unified representation of the rules for both forward and backward reasonings is thus desirable. Also, a strategy for translating qualitative process constraints to quantitative constraints is also of great importance for numerical optimization.

8.2.3. Tool Integration

Fig. 8.1 depicts the important issues for successful tool integration in an CAD/CAT/CAM environment.

(1) **Process model library:** A model library, serving as a central repository of the equipment-specific process models for different pieces of equipment, must be used to support various manufacturing applications (Ch. 2). The library contains both the models and the generic methods that operate on the models. The library is currently implemented with the object-oriented C++ language. The model library should interact with the INGRES relational database in order to support multiple model versions.



BLIS: Berkeley Laboratory Information System

Fig. 8.1 Tool integration for recipe generator.

(2) **Process flow language and interchange format**: At present, the recipe generator is driven solely by commands from the user interface and is thus isolated from other applications. In view of the integration with other applications in a CIM environment (Ch. 2), future research should consider driving the recipe generator through a more general inter-application interface such as the Berkeley Process Flow Language (BPFL) [6]. BPFL has been designated as a driver for various applications in the BCIM framework. Other future research should also consider using the PIF (Process Interchange Format) [7], for specifying the structural requirements and interfacing to other process CAD tools.

(3) **Process equipment**: Due to hardware difficulties, the present recipe generator cannot down-load recipes directly to the process equipment. This functionality should be included in future implementations.

(4) **Programming environment**: Three different languages (Lisp, Fortran, C) have been used for the rapid prototyping of the recipe generator. This has not only increased the complexity of the implementation, but also affected the performance of the system. A coherent programming environment with one programming language (such as C++) should be the target of future implementations.

8.2.4. Other Applications of the Models

The developed equipment-specific process models have been or will be applied to many different manufacturing applications (Fig. 2 of Ch. 2). Some of these applications are outlined below:

(1) **Process diagnosis**: The deviation between the predicted values of process models and the actual in-line measurements of process results is used as "evidence" for automated process/equipment diagnosis [8].

(2) **Statistical process control (SPC)**: The statistically derived equipment-specific process models can be used in SPC analysis. The 3-σ prediction bounds of the models, which can be computed at various recipe settings, can serve as the "lower" and "upper" control limits (LCL and UCL) for the equipment [9,10]. SPC can be combined with the knowledge-based system for process/equipment diagnosis [8].

(3) **Process monitoring**: The significant process parameters, as identified in the process models, should be collected from the equipment during the real-time process monitoring [8].

(4) **Link to process CAD tools**: Realistic prediction of the process responses from equipment settings can be accomplished by making the equipment models available to a Technology CAD tool [10].

(5) **Feed-back/forward control**: Statistically characterized process models are incorporated in a supervised workcell controller for photolithography [10]:

(6) **Temporal models**: The temporal behavior, such as aging, of the process equipment has been investigated with the LPCVD models [9].

REFERENCES

[1] L. Zadeh, "A Simple View of the Dempster-Shafer Theory of Evidence and its Implication for the Rule of Combination", *The AI Magazine*, pp.86-89, Summer 1986.

[2] S. Crary, "A Simulated Annealing Approach to Optimal Design of Experiments", *submitted to Lecture Notes in Computer Sciences*, Springer-Veriag, 1990.

[3] S. Crary, "A Differential Approach to Optimal Design of Experiments", *submitted to the American Journal of Physics*, 1990.

[4] G. E. Box, N. R. Draper, "Empirical Model-Building and Response Surfaces", John Wiley & Sons, 1987.

[5] R. H. Myers, et. al., "Response Surface Methodology: 1966 - 1988", *Technometrics, Vol. 31, No. 2*, May 1989, pp. 137-159.

[6] C. Williams, "Berkeley Process Flow Language - BPFL", *ERL Research Summary*, EECS, University of California, Berkeley, 1990.

[7] S. G. Duvall, "An Interchange Format for Device and Process Simulation", *IEEE Trans. Computer-Aided-Design*, Vol. CAD-7, pp. 741-754, July 1988.

[8] N. H. Chang, C. Spanos. "Chronological Equipment Diagnosis with Evidence Integration", *Submitted to the SPIE Conference on Applications of Artificial Intelligence VIII, 1990.*

[9] S. Lee, "A Strategy of Adaptive Regression Modeling of LPCVD Reactors", from "Special Issues in Semiconductor Manufacturing", *Memorandum No. UCB/ERL M90/8*, pp. 69-80, January 1990.

[10] H. Guo, T. Luan, Z-M. Ling, *Private Communication*, University of California, Berkeley, 1990.

APPENDIX 1
THE ANOVA TABLES

Source	DF	Sum of Squares	Mean Square	F-Ratio	Significance
Total	11	2.104	0.191		
Regression	3	2.079	0.693	221.600	0.000
Residual	8	0.025	0.003		
Lack of Fit	8	0.025	0.003		
Random Error	0				

Term	Coefficient	Standard Error	T-Value	Significance
1	21.51	0.92	23.27	0.0001
1/T	-15414.13	632.83	24.36	0.0001
log(P)	0.22	0.08	2.87	0.0207
1/Q	-61.92	8.34	7.43	0.0001

Table A1.1 ANOVA table of the first-stage deposition rate model

Source	DF	Sum of Squares	Mean Square	F-Ratio	Significance
Total	11	314.637	23.603		
Regression	5	310.898	62.180	99.780	0.000
Residual	6	3.739	0.623		
Lack of Fit	6	3.739	0.623		
Random Error	0				

Term	Coefficient	Standard Error	T-Value	Significance
1	164.59	36.67	4.49	0.0042
(T-900.5)/22.5	191.19	41.22	4.64	0.0035
(t-120)/60	182.02	40.83	4.46	0.0043
$((T-900.5)/22.5)^2$	31.65	5.01	6.32	0.0007
$((T-900.5)/22.5) \times ((t-120)/60)$	192.29	40.48	4.75	0.0032
$((t-120)/60)^2$	-17.17	3.35	5.13	0.0022

Table A1.2 ANOVA table of the first-stage stress model

Source	DF	Sum of Squares	Mean Square	F-Ratio	Significance
Total	22	3.867	0.176		
Regression	3	3.803	1.268	380.400	0.000
Residual	19	0.063	0.003		
Lack of Fit	19	0.063	0.003		
Random Error	0				

Term	Coefficient	Standard Error	T-Value	Significance
1	20.65	0.73	28.46	0.0001
log(P)	0.29	0.06	5.33	0.0001
1/T	-15189.21	515.92	29.44	0.0001
1/Q	-47.97	5.24	9.15	0.0001

Table A1.3 ANOVA table of the second-stage deposition rate model

Source	DF	Sum of Squares	Mean Square	F-Ratio	Significance
Total	24	422.222	17.593		
Regression	6	393.605	65.601	41.260	0.000
Residual	18	28.617	1.590		
Lack of Fit	18	28.617	1.590		
Random Error	0				

Term	Coefficient	Standard Error	T-Value	Significance
(898-T)/22.5	3.856	0.443	8.709	0.0001
(120-t)/45	3.445	0.676	5.096	0.0001
((898–T)/22.5) × ((400–P)/150)	-2.521	0.478	5.273	0.0001
((400–P)/150) × ((120–t)/45)	-1.764	0.605	2.915	0.0090
$((898–T)/22.5)^2$	3.376	0.437	7.731	0.0001
$((120–t)/45)^2$	-5.140	0.723	7.105	0.0001

Table A1.4 ANOVA table of the second-stage stress model

Source	DF	Sum of Squares	Mean Square	F-Ratio	Significance
Total	23	27.659	1.203		
Regression	3	18.259	6.086	12.950	0.000
Residual	20	9.400	0.470		
Lack of Fit	20	9.400	0.470		
Random Error	0				

Term	Coefficient	Standard Error	T-Value	Significance
1	1.49	0.31	4.89	0.0001
(P-400)/150	0.99	0.25	3.97	0.0007
((P−400)/150) × ((Q−175)/75)	1.06	0.23	4.64	0.0001
$((T-900.5)/22.5)^2$	-0.71	0.29	2.43	0.0244

Table A1.5 ANOVA table of the "within-wafer" deposition rate uniformity model

Source	DF	Sum of Squares	Mean Square	F-Ratio	Significance
Total	23	8.504	0.370		
Regression	4	5.180	1.295	7.401	0.001
Residual	19	3.324	0.175		
Lack of Fit	19	3.324	0.175		
Random Error	0				

Term	Coefficient	Standard Error	T-Value	Significance
1	2.31	0.19	12.14	0.0001
(T-900.5)/22.5	0.44	0.17	2.57	0.0188
((T−900.5)/22.5) × ((Q−175)/75)	0.38	0.16	2.33	0.0308
((Q−175)/75) × ((t−105)/45)	-0.44	0.13	3.41	0.0029
$((T-900.5)/22.5)^2$	-0.88	0.22	4.01	0.0008

Table A1.6 ANOVA table of the second-stage stress uniformity model

APPENDIX 2
IMPLEMENTATION DETAILS OF THE RECIPE GENERATOR

1. The Contour Plot Program

The contour plot program, written in C, generates the cartesian coordinates that define the contour levels of the response surface. The main function "gen_contour" is given below:

```
char      *gen_contour(xmin, ymin, xl, yl, nx, ny, zls_head, cstep);
int       nx, ny;
double    xmin, ymin, xl, yl, cstep;
zls_ptr   zls_head;
```

The input arguments to "gen_contour" are defined as follows:

```
xmin: the lower limit of the x variable.
ymin: the lower limit of the y variable.
xl:   the range (length) of the x variable.
yl:   the range (length) of the y variable.
zls_head: a C pointer to the head of link-list structures,
          that contain the values of the contour levels
          (z values) within the defined x-y cartesian space.
cstep:    The step size for displaying the contour levels
          (z values). The step size decides the number of
          contour curves to be plotted.
```

The output of "gen_contour" is a character string containing the cartesian coordinates of the contour curves. The algorithm of "gen_contour" first generates many rectangular grids, according to the ranges and intervals specified for the x and y-axis of the projected space (shown in Fig. A2.1). Next, the values of the contour levels $z = f(x, y)$ are computed at the corners of each grid, as well as the center of the grid by interpolation. Each grid is further "clipped" into four smaller triangles. A line segment will join any two points in a triangle, if a desired contour level intersects that triangle. All the line segments, with the connecting cartesian coordinates, will be re-organized and stored in linked-list structures. The linked-list structures, representing the contour curves, are later stored and returned as a character string.

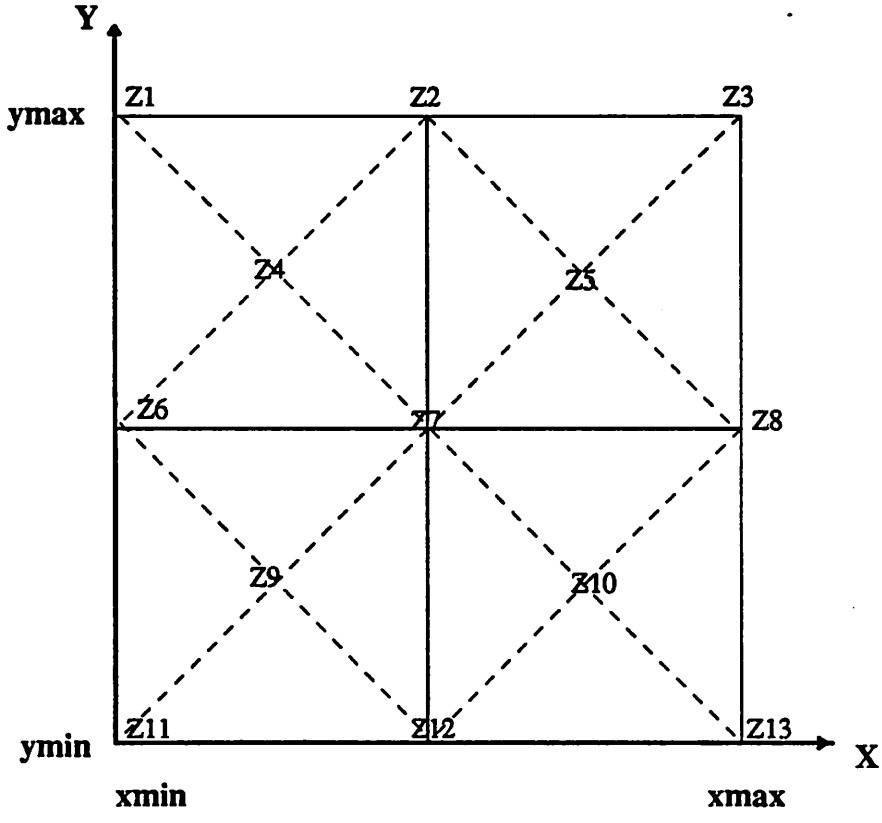


Fig. A2.1 The contour plot algorithm.

2. The Lisp-C Foreign Function Interface

Four Lisp functions serve as intermediate agents between the PICASSO user interface and the contour plot program "gen_contour". Each Lisp function will gather information from the user interface regarding the names, ranges and intervals of the free variables; as well as the names and values of the fixed variables. According to this information, each Lisp function will call the appropriate model to construct a linked-list structure containing the contour levels at the corners of each grid. A foreign function call [1] can finally be issued from each Lisp function to the C program "gen_contour" using the available foreign function interface routines in Common Lisp. The returned character strings from the C program will then be transformed to Lisp arrays, which are suitable as inputs to PICASSO 2D plotting widgets for displaying the contour curves .

3. The Fortran-NAG Interface

A Fortran subroutine "rg" is written to pass the following arguments to the NAG routine E04UCF for numerical optimization:

```
subroutine rg(objmin, objmax, conmin, conmax, x, obj, ifail, istate)

double precision objmin(4), objmax(4), conmin(5), conmax(5),
                 x(4), obj(4)

integer          ifail, istate(5)
```

The input arguments for "rg" are:

```
objmin: an input array of lower limits for the process objectives
objmax: an input array of upper limits for the process objectives
conmax: an input array of lower limits for the process constraints
conmin: an input array of upper limits for the process constraints
     x: an input/output array of initial/final the process parameters
   obj: an output array of the optimized process objectives
 ifail: a diagnostic parameter
istate: a diagnostic array
```

4. The Lisp-Fortran Foreign Function Interface:

A Lisp function is written to serve as an intermediate agent between the PICASSO user interface and the Fortran routine "rg". The Lisp function gathers information from the user interface regarding the limits of the quantitative process objectives and parameters as well as the goals of the qualitative objectives. With this information, the Lisp function constructs the appropriate numerical Lisp arrays. A foreign function call [1], with the Lisp arrays as arguments, is issued from Lisp to the Fortran "rg" routine, which in turn calls E04UCF. The optimization results are returned from Fortran to Lisp via the same set of arrays[1]. The Lisp function will evaluate the resulting arrays and display them appropriately on the PICASSO user interface.

[1] Arguments in Fortran are passed by address.

REFERENCES

[1] The Allegro Common Lisp User Guide, Release 3.0, June 1988.