THE INSTITUTE FOR SYSTEMS RESEARCH

ISR TECHNICAL REPORT 2007-25

A Comparative Study of Reactor Designs for the Production of Graded Films with Applications to Combinatorial CVD

Ramaswamy Sreenivasan, Raymond A. Adomaitis, Gary W. Rubloff

A Comparative Study of Reactor Designs for the Production of Graded Films with Applications to Combinatorial CVD

Ramaswamy Sreenivasan, Raymond A. Adomaitis[1]

Department of Chemical and Biomolecular Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742

Gary W. Rubloff

Department of Materials Science and Engineering and Institute for Systems Research, University of Maryland, College Park, MD 20742

Accepted for publication in the Journal of Crystal Growth, 2007.

Abstract:

Segmented CVD reactor designs enabling spatial control of across-wafer gas phase composition were evaluated for depositing graded films suitable for combinatorial studies. Specifically two reactor designs were constructed and evaluated with experiments and response surface model (RSM) based analysis to quantify the reactor performance in terms of film thickness uniformity, sensitivity to adjustable reactor operating conditions, range of thickness over which uniformity could be achieved and each reactor's ability to control the thickness gradient across the wafer surface. Design features distinguishing the two reactor systems and their influence on gradient control versus deposition rate performance are summarized. RS models relating wafer state properties to process recipes are shown to be effective tools to quantify, qualify and compare different reactor designs.

I. Introduction

The semiconductor industry constantly innovates and improves process and tool designs in an effort to keep up with Moore's Law. Chemical vapor deposition (CVD) tools are prevalent in every semiconductor fabrication facility as an efficient method for depositing

[1] Corresponding author. Email: adomaiti@umd.edu; Tel: (301)405-2969; Fax: (301)314-9920; Address: 2267 A.V. Williams Building, University of Maryland, MD 20742

nonvolatile solid films with good film conformality. However, conventional CVD systems are designed for a narrow range of operating conditions and do not offer much flexibility for improving process recipes and optimizing process development cycles for new materials. Also, most designs do not allow for controlling precursor concentration gradients over a wafer surface during a deposition run allowing for combinatorial capabilities.

There are relatively few examples of chemical vapor deposition reactor systems designed with combinatorial capabilities. Those that do exist, however, all demonstrate the capability to produce films with graded properties over a portion of the substrate surface. For example, the CVD reactor design of Gladfelter [1-2] features three feed tubes in a triangular arrangement across the substrate; a different single-source precursor is fed through each tube, generating compositional spreads of three metal dioxides over the substrate. In Wang [3-5], thickness graded films of hydrogenated silicon were deposited in a hot-wire CVD system featuring a mask and motorized shutter; control of the shutter spead was used to create strips of graded films over the substrate. Finally, in Taylor and Semancik [6], microhotplate devices were used to control the temperature in an array of micro-scale substrate samples; it was found that temperature gradients in the microhotplate supports resulted in a microstructurally graded film on the support legs.

Earlier work [7-9] by the authors of this study describes the preliminary construction and testing of a spatially programmable chemical vapor deposition (SP-CVD) system that was developed at the University of Maryland. The original SP-CVD reactor design (henceforth denoted as design A), construction, operation and preliminary evaluation experiments are described in the cited references. Figure 1 depicts a schematic diagram of design A comprising the individually controllable segmented showerhead with segments S1, S2 and S3 arranged over the wafer surface. For this and the previous studies, we consider blanket tungsten by H_2 reduction of WF_6 as the model deposition system; the overall deposition reaction is:

$WF_6\,(g) + 3H_2\,(g) \rightarrow W(s) + 6HF\,(g)$.

The results from the earlier work cited demonstrated for the first time the SP-CVD system's ability to be reprogrammed, effectively reconfiguring the reactor solely in

software between deposition runs to intentionally induce spatially non-uniform thickness deposition patterns on a single wafer. In [7], a relatively simple linear model was used to relate average film thickness under each of the segments to the feed gas recipe of each segment. Because this model did not account for segment-to-segment interactions, a more accurate modeling approach is developed in this paper to enable modeling of those interactions. The purpose of this paper is to demonstrate the use of Response Surface Models (RS models) to predict film thickness response over the entire wafer to adjustable process parameters enabling control to a specified thickness spatial function, such as a linear thickness gradient across a patch of wafer surface. This model is used to quantify the reactor's performance and examine the relative merits of different reactor designs. This approach is applied to evaluate two reactor designs: the "original" SP-CVD reactor (design A) and a modification (design B) motivated by an attempt to reduce the chamber volume.

III. Modeling for design A

Key to this study is the development of an accurate model of the full wafer response to adjustable process operating conditions; the model is necessary to compute process recipes that optimize a wafer profile objective function. The model, while physically motivated, will be identified from a set of experiments.

A. The response surface approach

The response surface modeling approach comprises of the following three steps [10]:

1) Systematic experiments: This step entails setting up a series of experiments that generate a range of reliable measurements of the desired output or response variable. The input variables/predictor variables are varied systematically to generate the range of measurements of the response variable by running experiments on the process tool. For the reactor designs discussed in this paper, we selected a subset of experiments based on our intuitive understanding of segment-to-segment interactions based on the results from preliminary experiments in [7], followed by a statistical analysis of the estimated parameters.
2) Identify a mathematical model relating the response variables (wafer thickness profile) to the input variables. The model form (linear vs. quadratic for example)

is based on our physical and intuitive understanding of the process. The model is tested for accuracy and validated. The derivation of the RS model is discussed in detail the next section.

3) The RS model is used to optimize the settings of the input variables to minimize the value of an objective function, based on our film gradient control criterion, solving a constrained non-linear optimization problem. This optimization is discussed in section IV, topic C in this paper.

In this paper, the input variables of the reactor system are defined by the recipe of the SP-CVD tool. This recipe comprises the flow rate of H_2 to each segment, the flow rate of WF_6 to each segment (H_2: WF_6 flow ratio is fixed at 4:1), and the showerhead-wafer gap size. The desired response variable is the film thickness of deposited tungsten defined at a specific spatial resolution over the wafer surface.

B. Derivation of the model form

Under isothermal processing conditions, the overall reaction rate can be expressed as the following surface reaction expression [11]:

$$R_{kin} = k_o [P_{WF_6}]^0 [P_{H_2}]^{1/2} \quad (1)$$

where,

R_{kin} is the rate of deposition of tungsten

$[P_{WF_6}]$ is the partial pressure of WF_6

$[P_{H_2}]$ is the partial pressure of H_2

According to this reaction kinetics model, the reaction rate does not depend on WF_6 partial pressure when sufficient WF_6 is present. However, the reaction rate is assumed to be proportional to the square root of the hydrogen precursor concentration X_{H_2}, so as a first order approximation we have:

$$X_{H_2} = \frac{P_{H_2}}{Total\ pressure\ P} \approx \frac{flow\ of\ H_2\ (sccm)}{Total\ flow\ of\ precursor\ (sccm)}$$

Furthermore, in our experiments we should expect a linear relationship between the deposition rate of the W film and $\sqrt{X_{H_2}}$ when precursor conversion rates are low, and so the square root of H_2 flow to each segment is used as input to our model.

The SP-CVD reactor has a showerhead with three segments which interact with one another by the following two gas transport mechanisms:

(1) Inter-segment gap diffusion: In this mechanism, process gases diffuse from one segment to the other segments though the gap between the wafer surface and the bottom of the segments owing to the concentration gradients between the segments when different recipes are used in neighboring segments.

(2) Inter-segment back diffusion: In this mechanism, process gases diffuse from the common exhaust volume (CEV) back into the segments owing to gas composition differences between the CEV and individual segments; these differences are attributable to different precursor recipes in the different segments or depletion at high deposition rates.

The showerhead-wafer gap is a process parameter that controls segment-to-segment interaction in the gap region and is included in the RS model. We derive a model which will predict the entire wafer film thickness profile $W_{pred}(r,\theta)$(in nm, for a fixed deposition time) based on the following input variables:

$x_i : \sqrt{H_2\, flow(sccm)}$ *to segment i*

g *: wafer – showerhead gap, mm*

Based on the deposition rate expression, and the mechanism of segment-to-segment interaction through inter-segment gap diffusion and back diffusion from CEV, we intuitively define the properties of the model to satisfy the following requirements:

1) The model should be such that it predicts the local thickness under segment *i* to be proportional to x_i and, to a lesser extent x_j for $i \neq j$ because of back diffusion.

2) Segment *i* film thickness dependency on x_j is modulated by *g* for transport to that region by inter- segment gap diffusion.

3) No deposition should take place when all x_i=0 and the deposition rate should not change with *g* alone.

Under these assumptions, we arrive at the RS model form to be:

$$W_{pred}(r,\theta) = b_1(r,\theta)x_1 + b_2(r,\theta)x_2 + b_3(r,\theta)x_3 + b_{1,4}(r,\theta)x_1 g + b_{2,4}(r,\theta)x_2 g + b_{3,4}(r,\theta)x_3 g \quad (2)$$

where the subscripts i=1,2,3 denote the segment number, and the double subscripts (i, 4), i=1,2,3 denote the segment feed/gap interaction terms (i.e., 4 represents the 4th model input which is the gap size).

To understand this model, consider a spatial point (r',θ') under segment 1; the terms in the expression $b_1(r,\theta)x_1 + b_2(r,\theta)x_2 + b_3(r,\theta)x_3$ are 'designed' to satisfy requirement 1, i.e., $W_{pred}(r',\theta')$ would be primarily dependent on x_1 and the coefficient b_1 quantifies this dependency. The terms $b_2(r',\theta')x_2$ and $b_3(r',\theta')x_3$ account for the contribution of back diffusion to the point thickness $W_{pred}(r',\theta')$.

The terms in the expression $b_{1,4}(r,\theta)x_1 g + b_{2,4}(r,\theta)x_2 g + b_{3,4}(r,\theta)x_3 g$ are designed to satisfy requirement 2, i.e., if the spatial point (r',θ') is under segment 1, $W_{pred}(r',\theta')$ also will depend on the inter-segment gap diffusion which is captured by the terms $b_{1,4}(r',\theta')x_1 g$, $b_{2,4}(r',\theta')x_2 g$, and $b_{3,4}(r',\theta')x_3 g$. Finally, the absence of a constant term in equation (2) satisfies requirement 3, i.e., when $x_i=0$ and $g=0$, $W_{pred}(r,\theta) = 0$.

The six spatially varying coefficients $b_i(r,\theta)$ and $b_{i,j}(r,\theta)$ are computed from the solution of the least squares procedure using the N experimentally determined thickness maps and corresponding process recipes, where N>6. The unique computational approach necessary to compute the spatially varying coefficients will be discussed in a separate publication.

C. Data set to build RSM for design A

25 wafers were processed for creating the data set from which we derived the RS model. Each wafer was dipped into 10% HF solution to remove native silicon-oxide film and impurities that block the nucleation of tungsten crystals; after cleaning, the wafers were immediately loaded onto the substrate heater in the reaction chamber. For all experiments described in this article, the heater temperature is set at 400C giving an approximate wafer temperature of 380C. Deposition time was 900 seconds for all wafers. All experiments were carried out at a reactor pressure of 1 torr maintained by a downstream throttle valve.

Table 1 summarizes a set of experiments which were carried out to generate films of varying thicknesses under different segments by varying the flow rates of the precursor

gases and the showerhead-wafer gap sizes. After each deposition process, film thickness was measured using a 4 point probe (4PP) ex-situ metrology station. The 4-point probe measurements result in a rectangular grid of measurements over the wafer surfaces with an approximate spatial resolution of 3.45 mm generating 900 measurement points. Numerical analysis of these wafer maps begins by interpolating the thickness data to a numerical quadrature grid defined on a computational domain that has the same physical dimensions as the wafer (see e.g., [12] for the underlying numerical methods, and [13] for another CVD application). This quadrature grid also is used for numerical interpolation of film thickness in each segment to give a finer (higher resolution) representation of film thickness under each segment.

D. RS model identification and validation for design A

The six spatially varying coefficients $b_i(r,\theta)$ and $b_{i,j}(r,\theta)$ are computed from the solution of the least squares procedure using the 25 experimentally determined thickness maps and corresponding process recipes.

Figure 2 illustrates the comparison between the model's prediction and true measurement for wafers No. 6, 8 and 23 (Table 1). These wafers were processed with the reactor operating in the non-uniformity mode. The RS model was used to predict the segment averaged values which show a good agreement with the true segment averaged values as shown by the bar charts. 10 wafers were processed with the same recipe, operating the reactor in the uniformity mode. This recipe (40 sccm of H_2 in S1 and S3 and 20 sccm of H_2 in S2) was calculated to be the recipe required to produce a thickness of 660 nm in each segment using a linear model. See Ref. 7 for details regarding this linear model. The average thickness of these 10 wafers are calculated and illustrated as a wafer map in Fig. 2 and compared with the RS model's prediction for the same. The RS model predicts the uniformity in agreement with the measured values to an accuracy of 8% with a standard deviation of 5%. Thus RS models can be effectively used to predict the thickness maps produced by the reactor when operated in both the non-uniformity mode and the uniformity mode.

IV. Performance analysis for design A

The validated RS model was used to evaluate the original reactor design using the following three criteria:

A. Sensitivity of film thickness profile to gas flow rate and gap

Differentiating equation (2) with respect to x_1 we obtain:

$$\frac{\partial W_{pred}(r,\theta)}{\partial x_1} = b_1(r,\theta) + b_{1,4}(r,\theta)g \tag{3}$$

Similarly we obtain:

$$\frac{\partial W_{pred}(r,\theta)}{\partial x_2} = b_2(r,\theta) + b_{2,4}(r,\theta)g \text{ and} \tag{4}$$

$$\frac{\partial W_{pred}(r,\theta)}{\partial x_3} = b_3(r,\theta) + b_{3,4}(r,\theta)g \tag{5}$$

These sensitivity maps with respect to the reactant concentration of gas fed to each segment are calculated for all (r,θ) over the patch of the wafer surface under each segment. Color plots of these patches are found in Fig. 3 showing the sensitivity of $W_{pred}(r,\theta)$ in each segment to each x_i with increasing gap size. Two effects are observed here:

1) The thickness of film $W_{pred}(r,\theta)$ in each segment patch i is most sensitive to the corresponding x_i. This observation is physically intuitive.
2) The sensitivity decreases with increasing gap size because as gap size increases, precursor gases 'escape' into the external chamber volume causing reactant depletion over the wafer surface.

Because of the second of these two observations, design A performs poorly when gap sizes are greater than 3 mm. For large gap sizes, the conversion rates are reduced and programmability of the reactor cannot be exploited for desired uniformity/non-uniformity profiles.

B. Range of segment-to-segment uniformity

With the potential limitations in operating performance for large gap sizes in mind, the RS model is used to predict the range of uniform film thicknesses that can be produced across all segments (i.e., same thicknesses in all three segments) given the fixed range of flow rates allowed by the mass flow controllers (MFCs) for varying gap size. The MFCs

for WF_6 have a range from 0 sccm to 12 sccm limiting the H_2 flow range from 0 sccm to 48 sccm. The gap size ranges from 0 to 5 mm.

We define:

W_i^s to be the average thickness for segment i

x_i^s to be the recipe for segment i to achieve W_i^s

S_i to be the segment area , $i = 1,2,3$

We then write out the expression for W_i^s using (2) and the above defined terms in compact matrix form as:

$$\begin{bmatrix} W_1^s \\ W_2^s \\ W_3^s \end{bmatrix} = \mathbf{B} \begin{bmatrix} x_1^s \\ x_2^s \\ x_3^s \end{bmatrix} + \mathbf{DG} \begin{bmatrix} x_1^s \\ x_2^s \\ x_3^s \end{bmatrix} \tag{6}$$

where the segment-averaged RS model coefficients are

$$B_{i,j} = \int_{Si} b_j(r,\theta) ds / \int_{Si} ds_i; \;\; i, j = 1,2,3, \tag{7}$$

$$D_{i,j} = \int_{Si} b_{j,4}(r,\theta) ds / \int_{Si} ds_i; \;\; i, j = 1,2,3 \tag{8}$$

and the influence of gap size g is included as the matrix

$$\mathbf{G} = \begin{bmatrix} g & 0 & 0 \\ 0 & g & 0 \\ 0 & 0 & g \end{bmatrix} \tag{9}$$

To calculate the range of uniformity that can be achieved using the reactor, we use the following steps:

1) Set g=0 mm, $W^s \rightarrow 0$ (desired uniform thickness)
2) Set $W_1^s = W_2^s = W_3^s = W^s$
3) Solve equation (6) for unknowns x_1^s , x_2^s and x_3^s .
4) If x_i^s are valid (positive and below the MFC upper limit), then set $W^s = W^s + \varepsilon_w$ (increment thickness) and return to step 2 i.e., if the unknowns x_1^s , x_2^s and x_3^s are within the acceptable flow range, we return to step 2, increase the value of the desired uniform thickness and recalculate the unknowns in step 3.

5) If x_i^s are not valid, then set $g = g + \varepsilon_g$ (increment gap), $W^s \to 0$ i.e., if the unknowns do not lie between 0 and $\sqrt{48}$ (sccm)$^{0.5}$, we conclude that the desired uniform thickness cannot be achieved with the current gap size and mass flow constraints and return to step 1 and increment the gap size provided it is less than 5 mm.
6) If g < 5 mm, return to step 2.

We thus calculate the maximum value of W^s for a given gap size given the mass flow constraints of the reactor. Figure 4 contains a plot of max W^s vs. gap size. The plot indicates that this reactor can be used to produce uniform films ranging from 0 to 800 nm across all three segments, given the above mentioned range of flow rates. As gap size exceeds ~3.4mm, max W^s reduces to zero. Thus design A is limited to depositing uniform films of thickness < 800 nm and cannot produce segment-to-segment uniform films for g>3.4mm. We hypothesize that the curve does not gradually tail off as g grows, but ends abruptly because of segment region asymmetries, such as non-uniform CEV concentration or heater hot spots.

C. Gradient control performance

The programmable reactor can be used to produce wafers with a deliberate thickness gradient across segment regions. To demonstrate this ability, we define a set point film gradient over a subset of the wafer by defining a line of length s_m on the wafer surface starting at point P1 in S1 and ending on P3 in S3 (Fig. 5a). The segment wall separating S1 and S3 bisects this line. We define the desired thickness gradient along this line $W_{set}(s)$ by the following equation:

$$W_{set}(s) = \frac{2s - s_m}{s_m} W_m \delta_m + W_{avg} \tag{10}$$

where

$0 \le s \le s_m$, W_{avg} is the mean thickness along the gradient defined by the user in nm, and W_m is the difference between the values of the thickness at the two extremities of the gradient, defined by the user in nm.

δ_m is a tuning parameter that varies from -1 to 1. This parameter is used so that equation (10) represents all possible linear thickness gradients from P1 to P3 over the length s. Consider, for example,

$$W_{set}(s) = -(\frac{2W_m}{s_m})s + (W_{avg} + W_m) \text{ when } \delta_m = -1 \quad \text{(maximum negative gradient)} \tag{11}$$

$$W_{set}(s) = W_{avg} \text{ when } \delta_m = 0 \text{ (flat profile)} \tag{12}$$

$$W_{set}(s) = (\frac{2W_m}{s_m})s + (W_{avg} - W_m) \text{ when } \delta_m = +1 \text{ (maximum positive gradient)} \tag{13}$$

Equation (11) is the equation of a line with negative slope implying that the desired film thickness $W_{set}(s)$ decreases from P1 to P3 in a linear fashion. Equation (12) is the equation of a horizontal line, implying that the desired film thickness remains constant at W_{avg} nm from P1 to P3. Finally, equation (13) is an equation of a line with a positive slope, implying that the desired film thickness increases from P1 to P3 in a linear fashion. All other values of δ_m between -1 and 1 represent the remaining gradients between $-(\frac{2W_m}{s_m})$ and $+(\frac{2W_m}{s_m})$ respectively.

In each of the above cases, we set the thickness gradient over any desired target circular patch on the wafer surface, as shown in Fig. 5. The gradient along one axis of the patch is defined by (10), while along the orthogonal axis the gradient is set to zero resulting in a flat tilted circular set point patch, $W_{set}(r,\theta)$. Our objective is to calculate a recipe that when inserted into the RS model gives $W_{pred}(r,\theta)$ that matches $W_{set}(r,\theta)$ over the entire target patch as accurately as possible. This can be stated as the following optimization problem:

$$\min_{g,x} \left\| W_{pred}(r,\theta) - W_{set}(r,\theta) \right\| \tag{14}$$

$$\text{subject to } 0 \le x_i \le 7(sccm)^{0.5} \quad i = 1,2,3$$

$$0 \le g \le 5mm$$

We define the objective function by numerically computing the norm of the difference between the computed wafer profile in the patch region and the set point. We assume that the WF_6 flow to each segment can be controlled to within 1 sccm resolution (based on the specifications on the mass flow controllers used for gas delivery) and that the gap size g

can be adjusted in 1mm increments. This means an exhaustive search over all possible segment recipes and gap sites is possible to guarantee that we will determine a global minimum to solve the constrained optimization problem (14). Figures 5b, 5c and 5d illustrate the $W_{set}(s)$ and $W_{pred}(s)$ as thickness maps over the patch area when δ_m equals -1, 0 and 1 respectively. The parameter O_v is the value of the expression $\left\| W_{pred}(s) - W_{set}(s) \right\|$ at the end of the optimization.

Figure 6a illustrates the gradient across the circular patch extending from S1 to S3 obtained for different values of δ_m. Plots of the x_i, g, and O_v computed as solutions to (14) as a function of δ_m are shown in plots 6b and 6c respectively. When δ_m = -1 our set point corresponds to a film profile that is thickest under S1 and thinnest under S3 on the defined circular patch. Intuitively, we would predict that x_1 should be higher than x_3. The optimization routine computes a recipe which confirms our intuition. Figure 6b illustrates this recipe with x_1~6 $(\text{sccm})^{0.5}$, with x_2~0 $(\text{sccm})^{0.5}$ and x_3~0 $(\text{sccm})^{0.5}$. Intuitively, to maintain a steep thickness gradient we would expect to use a small gap size. The optimization routine arrives at gap size ~1 mm (Fig. 6b) to achieve this desired gradient.

When δ_m = 0, our set point corresponds to a film profile that is flat from S1 to S3. Intuitively, we would predict that x_1~x_3. Figure 6b illustrates this recipe with x_1, x_3 ~ 3 $(\text{sccm})^{0.5}$ and x_2~4 $(\text{sccm})^{0.5}$. However the gap size is 2 mm. Intuitively we would expect a larger gap size for flat profiles, but because design A yields very poor conversion rates with large gap sizes, the RS model used in the optimization routine 'recommends' a small gap size even for flat profiles.

When δ_m = 1, the target film profile is thickest under S3 and thinnest under S1 on the defined circular patch. Intuitively, we would predict that x_3 should be higher than x_1 using a smaller gap size. The optimization routine computes this recipe to be x_1~0 $(\text{sccm})^{0.5}$, x_2~6 $(\text{sccm})^{0.5}$, and x_3~5 $(\text{sccm})^{0.5}$ (Fig. 6b) with a gap size of ~1 mm.

Thus the RS model was effectively be used to identify recipes to achieve desired thickness gradients on wafer. O_v as low as ~1nm when δ_m equaled 0, with a mean of 19 nm over all δ_m was achieved. Roughly speaking, this is approximately 5% error in achieving our set point profile, implying a very good gradient control.

V. The mini reactor (Design B)

The analysis using the RS model for design A revealed the following drawbacks:

1) Gap sizes > 3 mm cannot be used because precursor gases escape into the larger chamber volume resulting in poor conversion rates. Smaller precursor flow rates would further lower the conversion rate. The use of design A for combinatorial Atomic Layer Deposition (ALD), a future research direction, requires minute quantities of precursors pulsed into the reactor. Most of these pulses would escape into the larger chamber volume resulting in a very poor growth rate. This drawback calls for a smaller chamber volume.
2) The maximum thickness that can be uniformly deposited in all three segments is ~800 nm. A smaller chamber volume would improve conversion rates and thicker uniform films can be deposited.

Owing to the above drawbacks of design A, it was decided to design, construct and implement a smaller chamber (called a mini chamber in Figure 7) volume. In this article, we refer to the SP-CVD reactor with the mini chamber as design B. The mini chamber was constructed from aluminum and comprised of two parts. 1) The main mini chamber and 2) the lid with appropriately shaped holes through which the segments pass. Figure 7 illustrates a schematic of the front view of the design B. The mini chamber seats around the heater and the wafer. The lid rests on small screws drilled horizontally into the segment walls 120 degrees apart. The clearance between the outer segment walls and the inner walls of the mini chamber is 0.38 mm**.** After a wafer is transferred to the wafer lifter by a wafer holder from the load lock chamber and lowered onto the heater, the segments are lowered. The lid of the mini chamber then rests on the upper lip of the mini chamber while the segments continue to be lowered closer to the wafer. With the lid resting on the mini chamber, the segment-wafer gap can be varied from a minimum of 0 mm to a maximum of 10 mm. The mini chamber together with the lid enclosed the wafer in a cylindrical volume of diameter ~ 106 mm and a height of ~ 10.5 mm which is considerably smaller than the chamber volume in design A. [15] and [16] provide video links for design B.

VI. Modeling for design B

A. Data set to build RSM for design A

28 wafers were processed to create the data set from which we derived the RS model. Table 2 summarizes this data set. Pre-process cleaning, process temperature, pressure, and post process metrology and numerical interpolation techniques remained the same as they were for design A.

B. RS model identification and validation for design B

The six spatially varying coefficients $b_i(r,\theta)$ and $b_{i,4}(r,\theta)$ are computed in the same manner as they were computed for design A using the 28 experimentally determined thickness maps and corresponding process recipes.

Figure 8 illustrates the comparison between the model's prediction and true measurement for wafers No 2, 11, 13 and 22 (Table 2). These wafers were processed with the reactor operating in the non-uniform mode**.**. The RS model predicts the uniformity in agreement with the measured values to an accuracy of 14% with a standard deviation of 8%. Compared to the design A, the films deposited by design B are 3-4 times thicker. Design B confines more precursor gases over the wafer surface and improves reactant conversion by as much as 400%. However model accuracy appears to be lower than design B.

VII. Performance analysis for design B

The validated RS model was used to evaluate design B for the same three performance criteria used to evaluate the reactor design A.

A. Sensitivity to gap size

The RS model captures the sensitivity of the mini reactor to x_i and gap through the color plots in Fig.9. The inferences from the plots are:

1) As with design A, the thickness of film $W_{pred}(r,\theta)$ in each segment is most sensitive to the corresponding x_i for that segment.

2) This sensitivity does not decrease significantly with increasing gap size because the mini chamber in design B confines the gases preventing their escape to the main chamber as in design A. The slight decrease in sensitivity

is attributed to the inter-segment diffusion that is facilitated by increasing gap size.

B. Range of segment-to-segment uniformity

The range of uniformity that can be achieved using design B was calculated using the same procedure used for design A. Figure 10 shows a plot of max W^s vs. gap size for both design A and the design B. The plot indicates that the design B can be used to produce uniform films across all three segments ranging from 0 to 1800 nm, given the earlier mentioned range of flow rates. Design B can thus be used to produce uniform films at a rate 2 to 3 times that of design A.

C. Gradient control performance

We defined and solved the gradient optimization problem for design B using the same approach used for design A. Figures 11b, 11c and 11d illustrate the $W_{set}(s)$ and $W_{pred}(s)$ when δ_m equals -1, 0 and 1 respectively while Figure 12 illustrates optimized profiles and plots between δ_m and x_i, gap and O_v.

When δ_m = -1 the optimization routine computes a recipe (Fig. 12b) with x_1~2(sccm)$^{0.5}$, x_2~0 (sccm)$^{0.5}$ and x_1~0 (sccm)$^{0.5}$. The optimization routine recommends a gap size ~2 mm to achieve this desired gradient.

When δ_m = 0, the optimization routine computes a recipe (Fig. 12b) with x_1~0 (sccm)$^{0.5}$, with x_2~2 (sccm)$^{0.5}$ and x_3~0 (sccm)$^{0.5}$ with a gap size ~4 mm. This confirms our intuition that we would expect a larger gap size for flat profiles.

When δ_m = 1, the optimization routine computes this recipe to be, x_1~0 (sccm)$^{0.5}$, x_2~0 (sccm)$^{0.5}$ and x_3~3 (sccm)$^{0.5}$ (Fig. 12b) with a gap size of ~1 mm. Comparing Fig.6b with Fig.12b, we see that design B requires ~ 50% smaller flow rates of precursor gases (defined by x_i) than design A for the same thickness gradient.

VIII. Concluding Remarks

The RS modeling approach was used successfully to compare the processing capabilities of two CVD reactor designs and to assess their ability to produce controlled graded films

over a sub-section of the wafer surface. The following table compares the merits/demerits of the two designs:

Serial No	Criteria of comparison	Design A	Design B
1	RS Model fidelity reported as $\left\|\frac{W_{predicted} - W_{measured}}{W_{measured}}\right\|$	8 ± 5%	14 ± 8%
2	Programmability (Uniformity/Non uniformity control)	Good	Good
3	Sensitivity to gap and flow rate	Sensitivity declines rapidly with increasing g	Exhibits good sensitivity with increasing g
4	Range of segment-to-segment uniformity for the given flow constraints	0-800nm with gaps 0<g<3mm	0-1800nm with gaps 0<g<2.5mm
5	Gradient control	Good	Good

We conclude that design A could be effectively used to deposit uniform and non-uniform films at low gap sizes accurately and with good repeatability. Design B could be used to deposit thicker uniform and non-uniform films. The gap size could be effectively used as a knob to control inter-segment diffusion in the case of design B.

Atomic Layer Deposition (ALD) processes for binary and ternary systems control film composition by adjusting the pulsing and purging frequencies of the individual precursors. Film compositions can be varied from one wafer to the next using this approach. However, deliberate composition gradient control within a single wafer deposition run has not been demonstrated for ALD. We are currently studying reactor designs for combinatorial ALD that enable gas composition gradient control over the wafer surface to deposit varying compositions over a single wafer. Design B, because of its smaller volume and higher conversion rates could prove useful for this purpose.

Acknowledgements

The authors acknowledge the support of the National Science Foundation (NSF) through Grant No. CTS0219200. They would also like to acknowledge Mr. Russ Wood at the Physics Machine Shop for his expert guidance while machining the mini chamber.

Figure Captions

Figure 1: Schematic of the SP-CVD reactor assembly: design A

Figure 2: True wafer maps (data) of wafers No.6, No.8 and No.23 (Table 1), and averaged profiles of 10 wafers processed with the same recipe, obtained from 4 point probe measurements and numerically interpolated in MATLAB are shown in the top row. The interpolated data are compared to the maps predicted by the RS model shown in the second row. The third row compares the average thickness for each segment through bar charts. The recipe is written in the format**:**

$$\left\lfloor \sqrt{H_2\ flow(sccm)\ in\ S1}, \sqrt{H_2\ flow(sccm)\ in\ S2}, \sqrt{H_2\ flow(sccm)\ in\ S3}, gap \right\rfloor$$

Figure 3: Sensitivity of the reactor to recipe and gap size as predicted by the RS model for design A. The greater the redness of the plot within a segment, the more sensitive that segment is to H_2 flow in that segment. As gap size increases, sensitivity to H_2 flow decreases because with increasing gap, the precursor gases 'escape' into the external volume of the chamber.

Figure 4: The range of uniformity control possible for the design A as predicted by the RS model. This plot conveys that this reactor design could be used to deposit uniform films ranging from 0 to 800 nm using gap sizes ranging from 0 to 3 mm, with WF_6 flow rates ranging from 0 to 12 sccm in each segment (limited by the MFCs) and H_2 flow rates in each segment ranging from 0 to 48 sccm (to maintain the stoichiometric ratio of 1:4 WF_6:H_2). Ar flow in each segment is 60-(H_2 flow+ WF_6flow) sccm.

Figure 5: Gradient control for design A for three cases of δ_m values (-1, 0 and 1). The value of the minimized objective function O_v at the end of the optimization routine is shown below the plots in each case. W_{avg}, W_m and O_v have units of nm.

Figure 6: Evaluation of gradient control across segments 1 and 3 as a function of δ_m using the RS model for design A.

Figure 7: Schematic front view of the SP-CVD reactor with the mini chamber (design B). When the segments are lowered, the lid is stopped by the wall of the mini chamber while the segments continue to be lowered to the desired segment-wafer gap. This design renders a chamber with a reduced volume and overcomes drawbacks of design A.

Figure 8: True wafer maps (data) of wafers No.2, No.11, No.13 and No.22 (Table 2), obtained from 4 point probe measurements and numerically interpolated in MATLAB are

shown in the top row. They are compared with the predicted maps by the RS model for the same shown in the second row. The third row compares the average thickness for each segment through bar charts. The recipe is written in the format: $\left\lfloor \sqrt{H_2\ flow(sccm)\ in\ S1}, \sqrt{H_2\ flow(sccm)\ in\ S2}, \sqrt{H_2\ flow(sccm)\ in\ S3}, gap \right\rfloor$.

Figure 9: Sensitivity of the reactor to recipe and gap size as predicted by the RS model for design B. The greater the redness of the plot of the plot within a segment the more sensitive that segment is to H_2 flow in that segment. As gap size increases, sensitivity to H_2 flow does not decrease as in design A because the mini reactor design minimizes the gases escaping into the larger chamber volume**.**

Figure 10: The range of uniformity control possible for the design A vs. design B as predicted by the RS model. This plot indicates that design B could be used to deposit uniform films ranging from 0 to 1800 nm using gap sizes ranging from 0 to 2.5 mm, with WF_6 flow rates ranging from 0 to 12 sccm in each segment (limited by the MFCs) and H_2 flow rates in each segment ranging from 0 to 48 sccm (to maintain the stoichiometric ratio of 1:4 WF_6:H_2). Ar flow in each segment is 60-(H_2 flow+ WF_6flow) sccm.

Figure 11: Gradient control for design B for three cases of δ_m values (-1, 0 and 1). The value of the minimized objective function O_v at the end of the optimization routine is shown below the plots in each case**.** W_{avg}, W_m and O_v have units of nm.

Figure 12: Evaluation of gradient control across segments 1 and 3 as a function of δ_m using the RS model for design B.

Table Captions

Table1: Wafers 1 to 25 were used to deposit films from the above recipes (varying flow rates and showerhead-wafer gaps sizes) for generating the data to obtain the RS model for reactor design A. WF_6:H_2 flow ratio in each segment is 1:4. Ar flow in each segment is 60-(H_2 flow+ WF_6 flow) sccm.

Table 2: Wafers 1 to 28 were used to deposit films from the above recipes (varying flow rates and showerhead-wafer gaps sizes) for generating the data to obtain the RS model for design B. WF_6:H_2 flow ratio in each segment is 1:4. Ar flow in each segment is 60-(H_2 flow+ WF_6 flow) sccm.

References

[1]. R. C. Smith, N. Hoilien, J. Roberts, S. A. Campbell, and W. L. Gladfelter, Combinatorial chemical vapor deposition of metal dioxides using anhydrous metal nitrates, Chem. Mater. (2002) **14**, 474-476.

[2]. B. Xia, R. C. Smith, T. L. Moersch, and W. L. Gladfelter, Balancing reactor fluid dynamics and deposition kinetics to achieve compositional variation in combinatorial chemical vapor depositions, Appl. Surf. Sci. (2004) **223**, 14-19.

[3]. Q. Wang, J. Perkins, H. M. Branz, J. Alleman, C. Duncan, and D. Ginley, Combinatorial synthesis of solid state electronic materials for renewable energy applications, Appl. Surf. Sci. (2002) **189**, 271-276.

[4]. Q. Wang, Combinatorial hot-wire CVD approach to exploring thin-film Si materals and devices, Thin Solid Films, (2003) **430**, 78-82.

[5]. Q. Wang, F. Liu, and D. Han, High-throughput chemical vapor deposition system and thin-film silicon library, Macromol. Rapid Commun. (2004) **25**, 326-329.

[6]. C. J. Taylor and S. Semancik, Use of microhotplate arrays as microdeposition substrates for materials exploration, Chem. Mater. (2002) **14**, 1671-1677.

[7]. R. Sreenivasan, R. A. Adomaitis, G. W. Rubloff, A Demonstration of Spatially Programmable Chemical Vapor Deposition: Model-Based Uniformity/Non-uniformity Control., J. Vac. Sci. Technol. (2006) B **24**, 2706.

[8]. R. A. Adomaitis, J. Kidder, and G. W. Rubloff, U.S. Patent No. 6,821,910 (23 November 2004).

[9]. J. O. Choo, R. A. Adomaitis, L. Henn-Lecordier, Y. Cai, and G. W. Rubloff, Development of a Spatially Controllable Chemical Vapor Deposition Reactor with Combinatorial Processing Capabilities, Rev. Sci. Instrum. (2005) **76**, 062217.

[10]. Response Surfaces: design and analyses, A. J. Khuri, J. A. Cornell, 2nd ed. ISBN: 0-8247-9741-8.

[11]. C. R. Kleijn, C. J. Hoogendoorn, A. Hasper, J. Holleman, and J. Middelhock, Transport Phenomena in tungsten LPCVD in a single-wafer reactor, J. Electrochem. Soc. (1991) **138**, 509-517.

[12]. R. A. Adomaitis, Objects for MWR, Computers & Chem. Eng (2002) **26**, 981-998.

[13]. R. P. Parikh, R. A. Adomaitis, J. D. Oliver, and B. H. Ponczak, Implementation of a Geometrically-based Criterion for Film Uniformity Control in a Planetary SiC CVD Reactor System, J. Process Control (2007) **17**, 477-488.

[14]. Video link 1 for design B: http://www.doflick.com/ViewVideo.aspx?vId=166

[15]. Video link 2 for design B: http://www.doflick.com/ViewVideo.aspx?vId=165

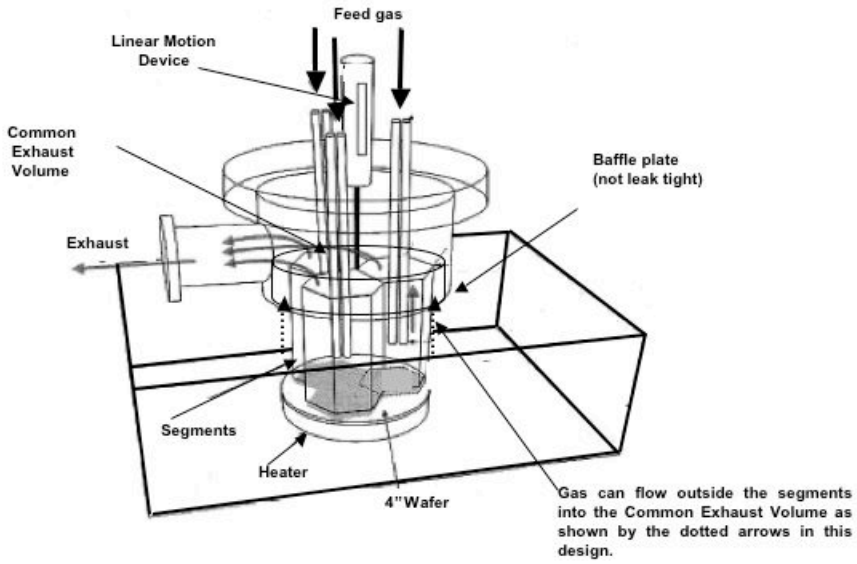


Figure 1

Design A data set

Wafer number	H2 flow S1(sccm)	H2 flow S2(sccm)	H2 flow S3(sccm)	Gap (mm)
1	24	36	48	1
2	48	24	36	1
3	24	36	48	1
4	48	24	36	1
5	36	48	24	1
6	24	36	48	1
7	48	24	36	1
8	36	48	24	1
9	48	24	36	1
10	40	20	40	1
11	40	20	40	1
12	40	20	40	1
13	40	20	40	1
14	40	20	40	1
15	40	20	40	1
16	40	20	40	1
17	40	20	40	1
18	40	20	40	1
19	40	20	40	1
20	24	36	48	3
21	48	24	36	3
22	36	48	24	3
23	24	36	48	3
24	48	24	36	3
25	36	48	24	3

Table1

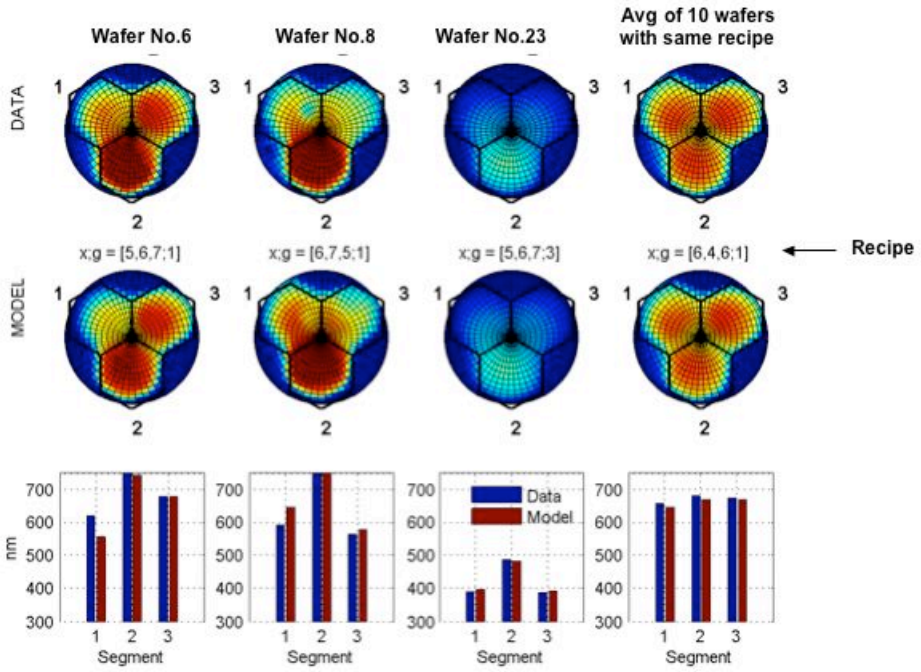


Figure 2

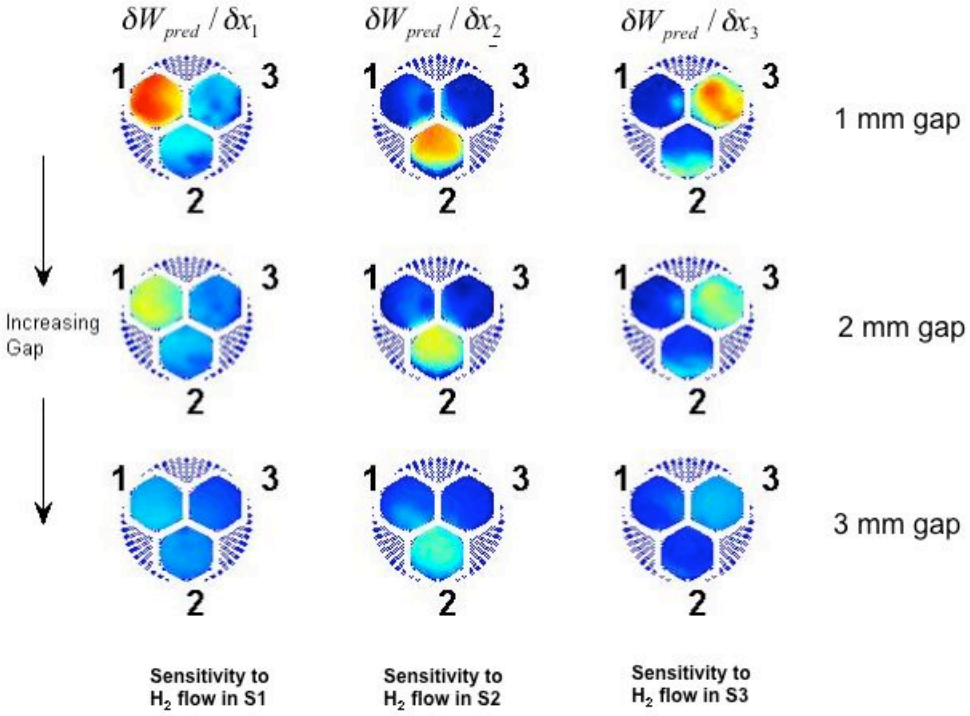


Figure 3

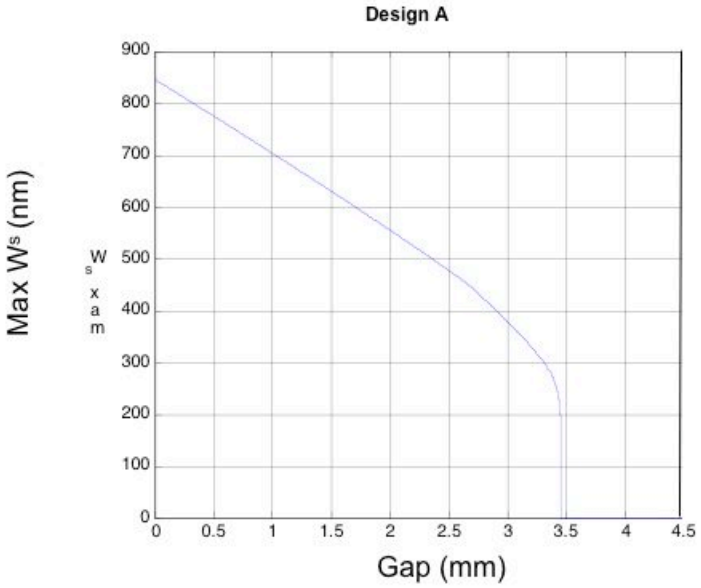


Figure 4

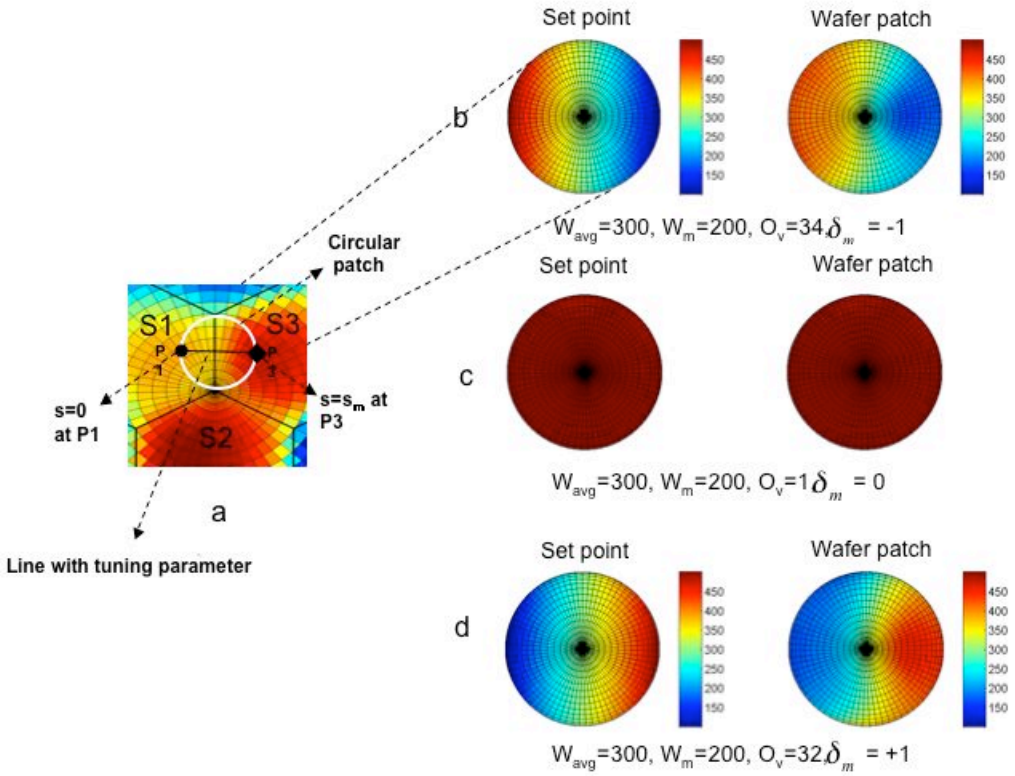


Figure 5

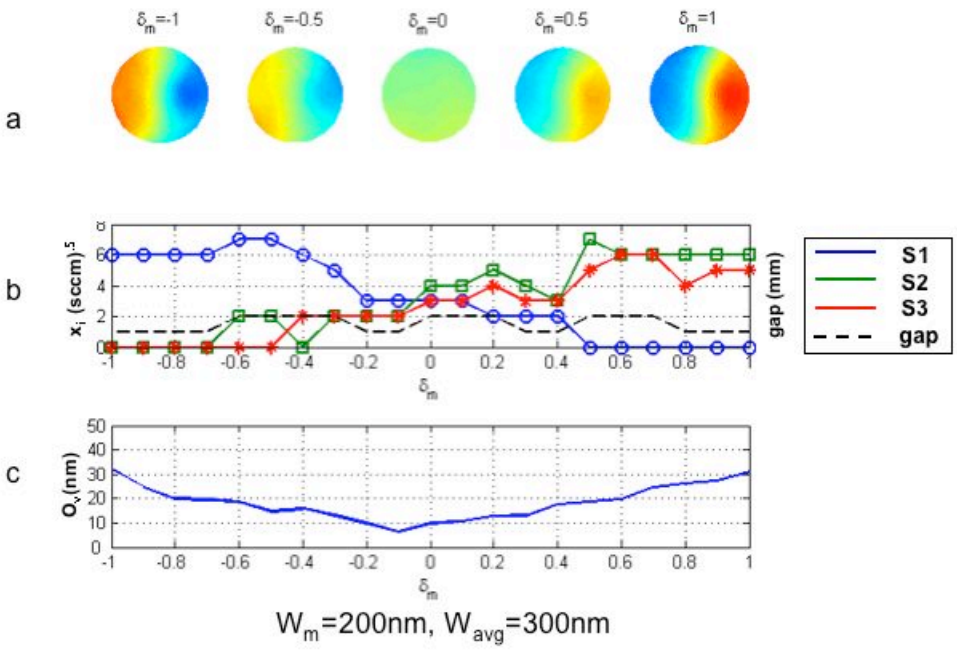


Figure 6

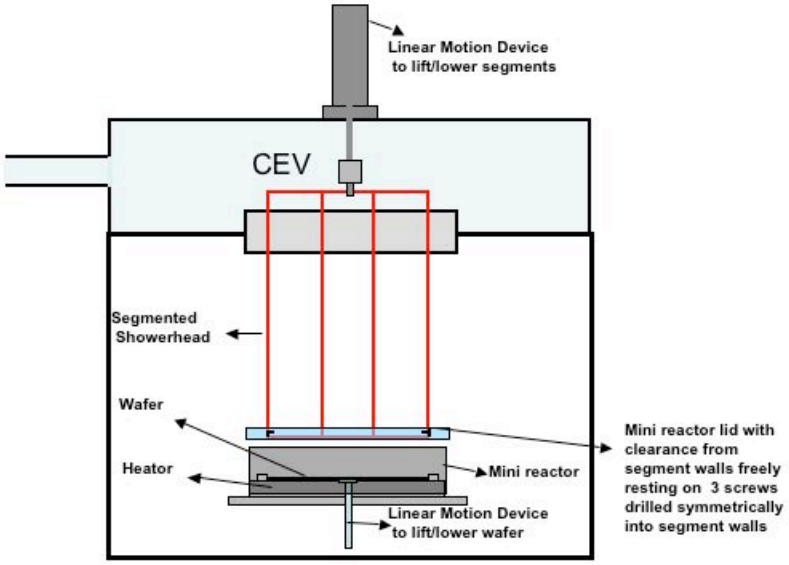


Figure 7

Design B data set

Wafer number	H2 flow S1(sccm)	H2 flow S2(sccm)	H2 flow S3(sccm)	Gap (mm)
1	16	32	48	3
2	48	16	32	3
3	32	32	32	3
4	32	0	0	3
5	0	32	0	3
6	0	0	32	3
7	0	0	32	1
8	0	32	0	1
9	32	0	0	1
10	48	16	32	1
11	32	48	16	1
12	16	32	48	1
13	16	32	48	5
14	32	48	16	5
15	32	32	32	5
16	32	0	0	5
17	0	32	0	5
18	0	0	32	5
19	0	0	32	3
20	0	32	0	3
21	32	0	0	3
22	32	32	32	3
23	48	16	32	3
24	32	48	16	3
25	16	32	48	3
26	0	0	32	2
27	0	32	0	2
28	32	0	0	2

Table 2

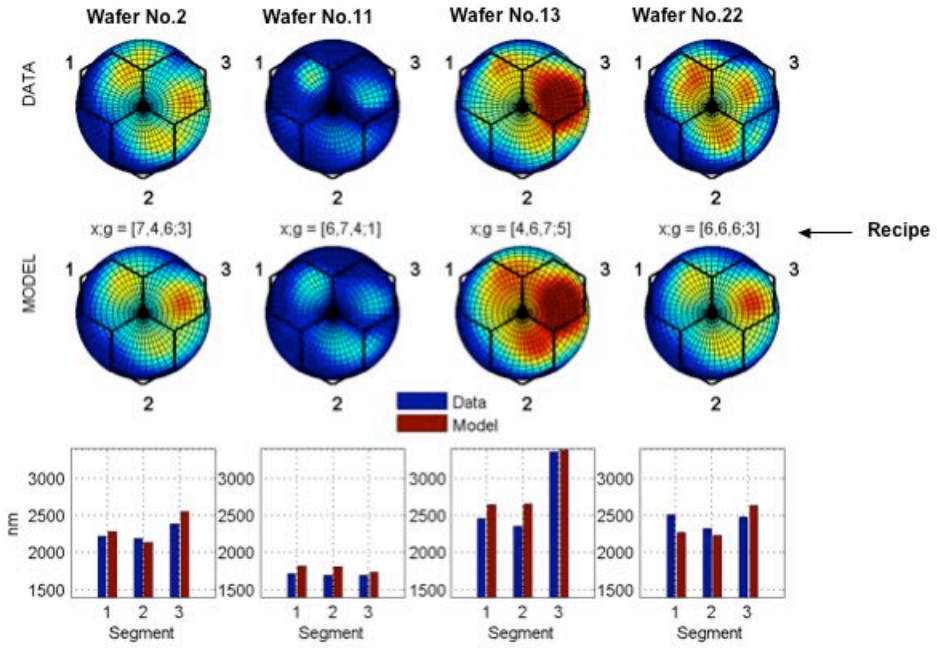


Figure 8

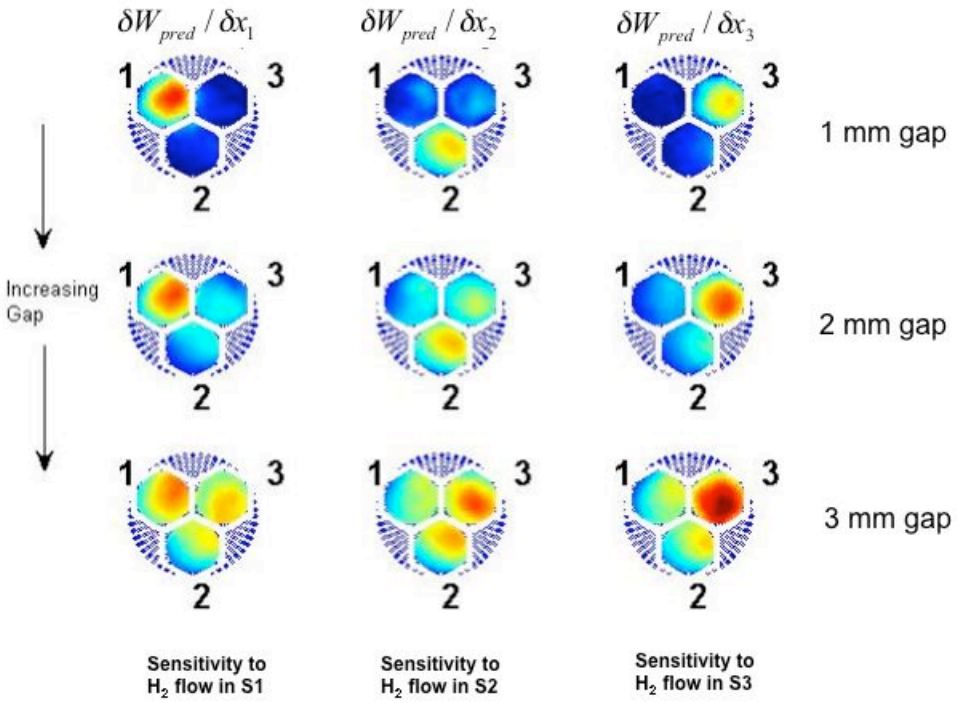


Figure 9

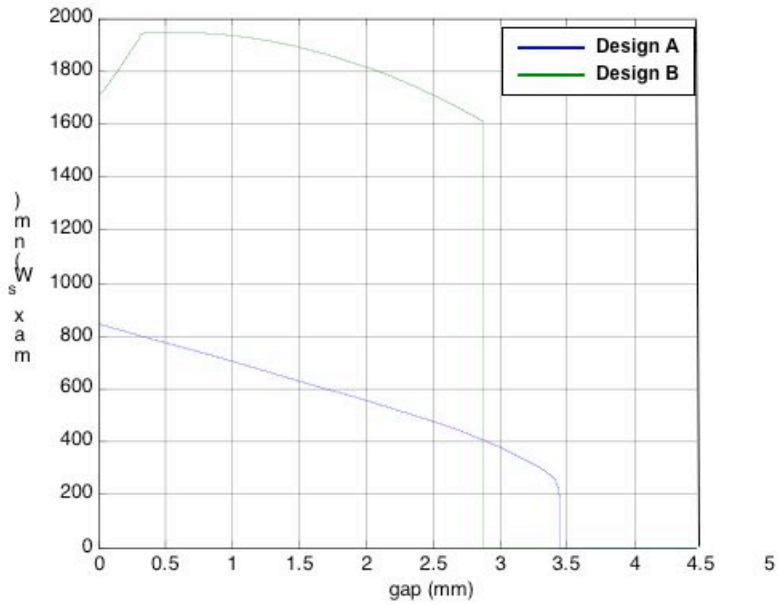
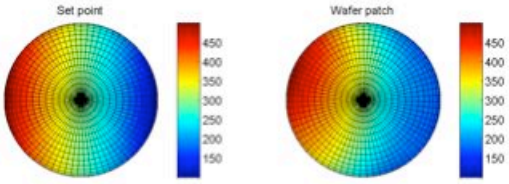
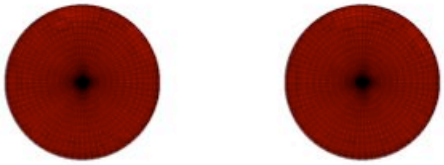


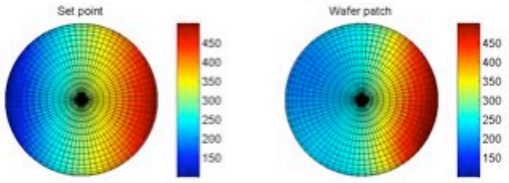
Figure 10



W_{avg}=300, W_m=200, O_v=15 δ_m = -1

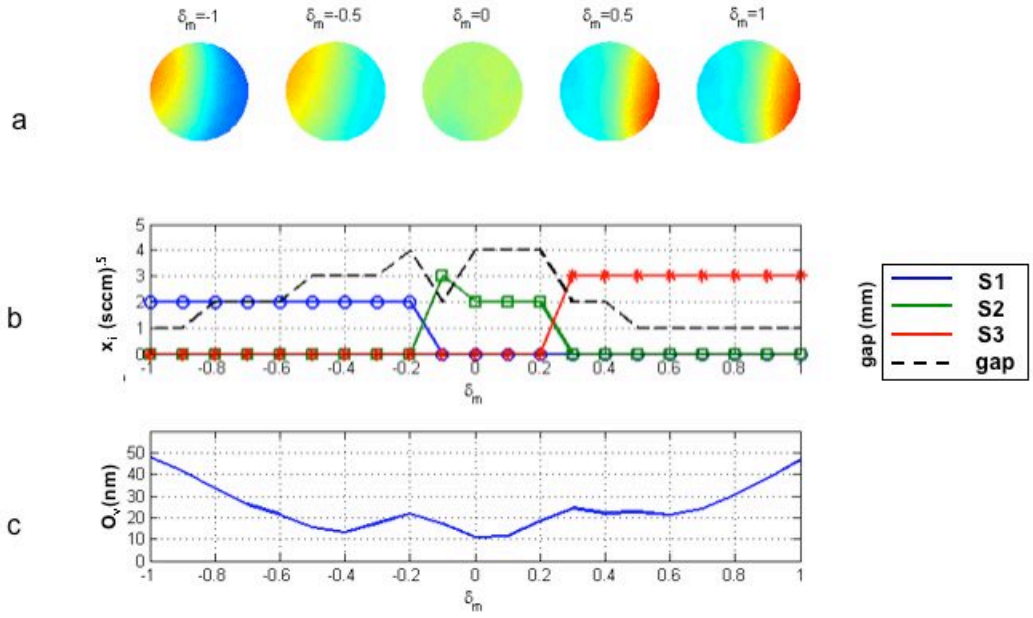


W_{avg}=300, W_m=200, O_v=5 δ_m = 0



W_{avg}=300, W_m=200, O_v=38 δ_m = +1

Figure 11



W_m=200nm, W_{avg}=300nm

Figure 12