ELSEVIER

Thin Solid Films 402 (2002) 280–289

thin solid films

www.elsevier.com/locate/tsf

Application of inverse problem algorithm for temperature uniformity in rapid thermal processing

Senpuu Lin[a,*], Hsin-Sen Chu[b,1]

[a]*Department of Mechanical Engineering, National Lien Ho Institute of Technology, Miaoli 360, Taiwan, ROC*
[b]*Department of Mechanical Engineering, National Chiao Tung University, Hsinchu 300, Taiwan, ROC*

Received 20 April 2001; received in revised form 24 August 2001; accepted 6 September 2001

Abstract

This article presents a finite-difference-method formulation to the application of inverse problem algorithms for uniform temperature tracking of several different linear ramp-up rates in rapid thermal processing. A one-dimensional thermal model and temperature-dependent thermal properties of silicon wafers are used. The required incident-heat-flux profiles for temperature uniformity across 300-mm-diameter 0.775-mm-thick silicon wafer were intuitively evaluated. Our numerical results indicate that temperature non-uniformity occurring during the ramp increase with the ramp-up rate. Although a linear ramp-up rate of 300 °C/s was used and random errors did reach 3.864 °C, the temperature over the wafer was maintained within 0.665 °C of the wafer center if the incident-heat-flux profiles were dynamically controlled according to the inverse results. These temperature non-uniformities could be acceptable in the advanced rapid thermal processing system.

Keywords: Temperature uniformity; Rapid thermal processing; Inverse problem algorithm

1. Introduction

The semiconductor-manufacturing trend toward ultra-large-scale-integration has led to a dramatic decrease in feature size with new generation. Rapid thermal processing (RTP) is an emerging and promising technology for microelectronic processes such as annealing, oxidation, chemical vapor deposition that provides numerous advantages over conventional furnace-based batch processing. RTP will be indispensable to meet the process requirements of future devices. In RTP, wafers are processed one at a time in a small, cold-walled reaction chamber. The treated wafer is heated from room temperature to as high as 1100 °C, in a few tenths of seconds, then processed and cooled. In general, the wafer is heated by means of radiation supplied by one or more lamp banks [1]. Because of radiative heat losses from the wafer edges, uniform irradiation does not guarantee uniform radial temperature profiles. If not corrected, this radial temperature gradient is usually large enough to generate thermal stresses and produce an unacceptable variation in film thickness, which affect the processing results and wafer-to-wafer uniformity [2]. Consequently, maintaining a uniform temperature across the wafer is a process requirement [3]. Generally, the maximum temperature difference across the wafer must be maintained within 2 °C during rapid thermal processing [4,5]. The stringent process uniformity and repeatability requirements demand a continuous improvement in wafer temperature control [6].

Many approaches including utility of a patterned susceptor [7] and model-based control [2–4,6,8–11] have been proposed for achieving temperature uniformity. This paper primarily deals with the determination of incident heat flux from heaters to eliminate radial

* Corresponding author. Tel.: +886-3-7323152, ext. 25; fax: +886-3-7371190.
E-mail addresses: spuulin@mail.nlhu.edu.tw (S. Lin), hschu@cc.nctu.edu.tw (H.-S. Chu).
[1] Tel.: +886-3-5712121, ext. 55115; fax: +886-3-5727930.

0040-6090/02/$ - see front matter
PII: S0040-6090(01)01630-3

temperature gradients across wafers [5,12–21]. Hill and Jones [13] investigated thermal uniformity with a uniform intensity field and one in which the intensity was linearly enhanced to a maximum of 8% vertically over the last 15 mm of a 150-mm wafer. Kakoschke et al. [14] evaluated enhanced illumination intensities at wafer peripheries vertically and laterally for a compensation of edge heat losses during processing. Gyurcsik et al. [15] introduced a two-step procedure for solving an inverse optimal-lamp-contour problem to achieve temperature uniformity in steady state. Sorrel et al. [5] applied power-law (first-, second- and seventh-degree) irradiation profiles to study the increases required in perimeter irradiation to maintain a wafer at an approximately uniform temperature. Norman [16] presented a technique based on linear programming for minimization of worst case error during temperature trajectory following. Zöllner et al. [17] compensated for radial temperature decreases using an adjustable lamp arrangement with optimized power settings calculated from wafer heat losses. Riley and Gyurcsik [18] presented a nodal analysis of wafer-edge to determine the amount of lateral heating needed to counteract edge cooling. Cho et al. [19] optimized the incident heat flux profile over a wafer by determining the heat loss profiles using Lord's thermal model [12], which simulates radial temperature gradients by assuming uniform temperature through the wafer thickness. Following the work of Riley and Gyurcsik [18], Perkins et al. [20] used wafer-edge node analysis to determine the idealized intensity profiles required for maintaining thermal uniformity both during transient and steady state. Jan and Lin [21] derived a necessary and sufficient condition on the irradiation of the wafer surface for their lamp configuration design. Some of these approaches rely largely on trial-and-error, which can be quite expensive and time-consuming. Some approaches, such as Cho et al. [18] and Perkins et al. [19], are systematic design methods for knowing whether a design satisfying given specifications exists and whether a given approach has an optimal design that satisfies given specifications. With microelectronic device specifications becoming tighter, it is imperative that the wafer temperature trajectory must be controlled precisely to meet processing requirements. Recently, computer-aided design of thermal processing based on fundamental models has become more attractive [2,3]. Balakrishnan and Edgar [2] discussed the development of temperature control algorithms for the RTP process. Janicki et al. [22] had successfully applied an inverse problem algorithm, which was implemented as a digital filter, for integrated circuit temperature estimation. Kurz and Müller [23] presented an inverse-modeling algorithm that is able to calculate the powers of arbitrary numbers of heaters in a crystal growth configuration in order to obtain a prescribed temperature distribution in a growing crystal. The inverse problem deals with

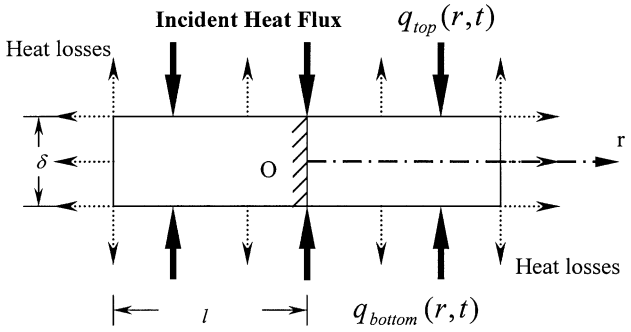

Fig. 1. Schematic representation of energy flux in a silicon wafer under two-sided incident radiation and radiant loss emitted from all surfaces.

determining crucial parameters in analyses such as those for internal energy sources, surface heat fluxes, thermal properties, etc., and has been widely applied to many design and manufacturing problems [24,25]. We [26,27] applied a one-dimensional thermal model to study the temperature uniformity of 300-mm silicon wafers subjected to a uniformly distributed heat flux during RTP using the inverse-source method. It was discovered that the resulting maximum temperature differences were only 0.326 °C during RTP. In the present article, a finite-difference-method formulation to the application of inverse problem algorithms for uniform temperature tracking of several different linear ramp-up rates is studied in rapid thermal processing systems. A one-dimensional thermal model and temperature-dependent thermal properties of silicon wafers are adopted. The required incident-heat-flux profiles for temperature uniformity across 300-mm-diameter 0.775-mm-thick silicon wafers were intuitively evaluated. The effects of random error of input data on temperature uniformity are also investigated.

2. Thermal model

Consider the thin axially symmetrical circular silicon wafer shown in Fig. 1. Let l and δ be the radius and thickness, respectively. T_0 is the initial uniform wafer temperature, and the ambient temperature is T_a. Symmetric heating on both sides of the wafer is adopted. The total incident heat fluxes on the top and the bottom surfaces of the wafer are denoted by q_{top} and q_{bottom}, respectively. Assume that the heat losses occur at all surfaces and that the temperature is uniform through the wafer thickness. A description of the one-dimensional thermal model is used here.

The governing equation for an axially symmetric cylindrical coordinate system with its origin at the wafer

center is

$$\rho c(T)\frac{\partial T}{\partial t}=\frac{1}{r}\frac{\partial}{\partial r}\left[k(T)r\frac{\partial T}{\partial r}\right]+\frac{1}{\delta}[G_{\text{top}}+G_{\text{bottom}}]$$
$$0<r<l \quad (1)$$

with

$$G_{\text{top}}=\alpha_{\text{top}}q_{\text{top}}(r,t)-\varepsilon_{\text{top}}\sigma_s(T^4-T_a^4)$$
$$G_{\text{bottom}}=\alpha_{\text{bottom}}q_{\text{bottom}}(r,t)-\varepsilon_{\text{bottom}}\sigma_s(T^4-T_a^4)$$

where $\sigma_s=5.672\times10^{-12}$ W cm^{-2} K^{-4} is the Stefan–Boltzmann constant, the wafer temperature T is a function of radius r, and time t; ρ, $k(T)$, $c(T)$, α_{top}, ε_{top}, α_{bottom}, and $\varepsilon_{\text{bottom}}$ are the wafer density, thermal conductivity, specific heat capacity, absorptivity of the top side, emissivity of the top side, absorptivity of the bottom side and emissivity of the bottom side, respectively. Here, because of the large temperature variations that occur during the processing, the temperature dependence of wafer thermal conductivity as well as specific heat capacity must be considered as follows [1]:

$$k(T)=802.99T^{-1.12}\quad (\text{W cm}^{-1}\ \text{K}^{-1})\quad 300\text{–}1683\ \text{K} \quad (2a)$$

$$c(T)=0.641+2.473\times10^{-4}T\quad (\text{J g}^{-1}\ \text{K}^{-1})\quad >300\ \text{K}, \quad (2b)$$

while the wafer density is assumed to be constant and equal to 2.33 g cm^{-3}. Since the silicon wafer is considered to be homogeneous in the present study, the dependence of $k(T)$ on spatial position is introduced only implicitly by the spatial dependence of the temperature. Because $k(T)$ is weakly dependent on temperature (see Fig. 2), spatial temperature variations across wafers at a certain time are expected to be small enough ($\leq$200 K) so that spatial variations in thermal conductivity may be ignored [14]. Thus, Eq. (1) may be reduced to

$$\rho c(T)\frac{\partial T}{\partial t}=k(T)\left[\frac{\partial^2 T}{\partial r^2}+\frac{1}{r}\frac{\partial T}{\partial r}\right]+\frac{1}{\delta}[G_{\text{top}}+G_{\text{bottom}}].$$
$$0<r<l \quad (3)$$

The initial and boundary conditions for the system described above are

$$T(r,t)=T_0 \quad \text{at } t=0 \quad (4)$$

$$\frac{\partial T}{\partial r}=0 \quad \text{at } r=0 \quad (5)$$

$$-k(T)\frac{\partial T}{\partial r}=\varepsilon_{\text{edge}}\sigma_s(T^4-T_a^4), \quad \text{at } r=l \quad (6)$$

where $\varepsilon_{\text{edge}}$ is the emissivity for the radiant heat losses at the wafer edges. We may assume without loss of generality that the incident heat fluxes on both sides during processing are equal, i.e. $q_{\text{top}}(r,t)=q_{\text{bottom}}(r,t)=q(r,t)$, and that the absorptivity of all wafer surfaces equals the emissivity of those surfaces. For simplicity, the emissivity of all surfaces is assumed to be the same and simply temperature-dependent as described by Virzi [28]:

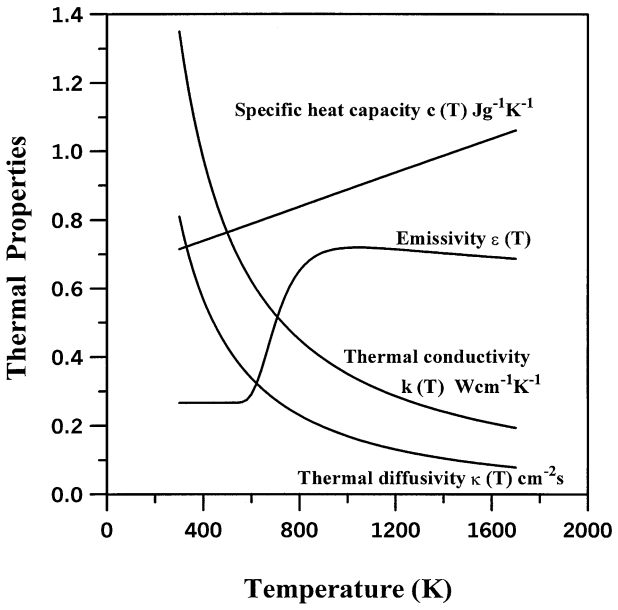

Fig. 2. Temperature dependence of silicon wafer thermal properties [1,28].

$$\alpha_{\text{top}}=\varepsilon_{\text{top}}=\varepsilon_{\text{bottom}}=\alpha_{\text{bottom}}=\varepsilon_{\text{edge}}=\varepsilon(T)$$
$$=0.2662+1.8591T^{-0.1996}e^{[(-1.0359\times10^{25})/T^{8.8328}]} \quad (7)$$

Thus, Eqs. (3) and (6) may be rewritten, respectively, as

$$\rho c(T)\frac{\partial T}{\partial t}=k(T)\left[\frac{\partial^2 T}{\partial r^2}+\frac{1}{r}\frac{\partial T}{\partial r}\right]+\frac{2}{\delta}\varepsilon(T)[q(r,t)-\sigma_s(T^4-T_a^4)]$$
$$0<r<l \quad (8)$$

and

$$-k(T)\frac{\partial T}{\partial r}=\varepsilon(T)\sigma_s(T^4-T_a^4) \quad \text{at } r=l \quad (9)$$

Defining wafer thermal diffusivity as $\kappa(T)=\dfrac{k(T)}{\rho c(T)}$ and introducing the dimensionless temperature θ, radial position R, time τ, incident heat flux Q, thickness-to-radius ratio η, thermal conductivity $K(\theta)$, specific heat capacity $C(\theta)$, thermal diffusivity $D(\theta)$, and emissivity $\varepsilon(\theta)$ as:

$$\theta=\frac{T}{T_a},\quad R=\frac{r}{l},\quad \tau=\frac{k_a t}{\rho c_a l^2},\quad Q(R,\tau)=\frac{l}{k_a T_a}q(r,t),$$
$$\eta=\frac{\delta}{l},\quad K(\theta)=\frac{k}{k_a},$$
$$C(\theta)=\frac{c}{c_a},\quad D(\theta)=\frac{K(\theta)}{C(\theta)},$$
$$\varepsilon(\theta)=0.2662+1.8591(T_a\theta)^{-0.1996}e^{[-(1.0359\times10^{25})/(T_a\theta)^{8.8328}]},\qquad(10)$$

where $k_a=k(T_a)$ and $c_a=c(T_a)$ are the thermal conductivity and specific heat capacity of the wafer at the ambient temperature T_a, respectively, then, the energy equation of Eq. (8) becomes

$$\frac{\partial\theta}{\partial\tau}=D(\theta)\left(\frac{\partial^2\theta}{\partial R^2}+\frac{1}{R}\frac{\partial\theta}{\partial R}\right)+\frac{2}{\eta}\frac{\varepsilon(\theta)}{C(\theta)}[Q(R,\tau)+A(1-\theta^4)],\quad 0<R<1\qquad(11)$$

The initial condition Eq. (4) and boundary conditions Eqs. (5) and (9) become, respectively,

$$\theta(R,\tau)=\theta_0\quad\text{at}\quad\tau=0\qquad(12)$$

and

$$\frac{\partial\theta}{\partial R}=0\quad\text{at}\quad R=0\qquad(13)$$

$$\frac{\partial\theta}{\partial R}=\frac{\varepsilon(\theta)}{K(\theta)}A(1-\theta^4)\quad\text{at}\quad R=1\qquad(14)$$

Here the dimensionless initial temperature $\theta_0=\frac{T_0}{T_a}$ and the dimensionless constant

$$A=\frac{\sigma_s l T_a^3}{k_a}$$

The numerical solution techniques used here are from the finite-difference method. A central-difference representation of the space derivative and an implicit backward-difference representation of the time derivative are adopted. We can approximate the governing equation and the initial condition as well as the boundary conditions using $T(r,t)=\theta\big((i-1)\Delta R,n\Delta\tau\big)=\theta_i^n$ with p equidistant grid and the temporal coordinate increment $\Delta\tau$. After the non-linear radiant fourth-power terms in Eqs. (11) and (14) have been simulated using a linear scheme and the successive over-relaxation (SOR)-by-lines method has been adopted, the unknowns in the subgroups to be modified simultaneously are set up such that the matrix of coefficients will be tridiagonal in form permitting use of the Thomas algorithm as follows:

$$b_i^n\theta_{i-1}^n+d_i^n\theta_i^n+a_i^n\theta_{i+1}^n=c_i^n,\qquad(15)$$

The superscript is denoted as the index of the temporal grid and the subscript is denoted as the index of the spatial grid. Given the incident heat flux, we can obtain the wafer temperature distributions. Calculation of the temperature field mentioned above is a so-called direct problem. It is well-posed because for every given incident heat flux, the temperature distribution may be uniquely determined. In the present study, it is necessary to determine causes of processes (e.g. the amount of incident heat flux) from their effects (temperature distributions across the wafer). Problems of this kind are called inverse problems. The intrinsic characteristic of inverse problems is that they are ill-posed; hence it has been considered for many years that solutions to inverse problems do not exist. However, in the early 1960s it was proved that the solutions to inverse heat conduction problems usually exist and are unique, but obtained estimates are not always numerically stable [24,25]. Therefore, special methods must be applied to solve inverse problems.

3. Inverse problem algorithm

The inverse problem algorithm [24,25] is practical for use in determining how to achieve temperature uniformity with RTP systems in which the magnitudes of incident heat fluxes required for maintaining temperature uniformity across wafers during processing are unknown.

The finite-difference scheme in the thermal model above at $\tau=\tau^m=m\Delta\tau$ is used to construct the following matrix equation:

$$\left[F^m\right]\{\theta^m\}=\{\theta^{m-1}\}+\{S^m\}+\{V^m\}\{\varphi^m\},\qquad(16)$$

Then the wafer temperature distribution $\{\theta^m\}$ can be rewritten as follows:

$$\{\theta^m\}=[F^m]^{-1}\{\{\theta^{m-1}\}+\{S^m\}\}+[F^m]^{-1}[V^m]\{\varphi^m\}=[M^m]\{\{\theta^{m-1}\}+\{S^m\}\}+\{N^m\}\{\varphi^m\}\qquad(17)$$

where $[M^m]=[F^m]^{-1}$ and $[N^m]=[F^m]^{-1}[V^m]$.

The vector $\{\theta^{m-1}\}$ contains $p+1$ values of the initial distribution or the temperature distribution for the preceding time step. The $\{S^m\}$ vector includes any known variables of the problem. The vector $\{\varphi^m\}$ consists of the unknown incident heat fluxes φ_j^m for $j=1,2,\ldots,p$, $p+1$ (i.e. $\{\varphi^m\}=\sum_{j=1}^{p+1}\{u_j\}\varphi_j^m$). $\{u_j\}$ is the unit column vector with a unit at the j-th component. As well, j is the grid number of the location of the estimated heat flux function φ_j.

For the next time step $m+1$ we arrive at

$$\{\theta^{m+1}\}=[M^{m+1}]\{\{\theta^m\}+\{S^{m+1}\}\}+[N^{m+1}]\{\varphi^{m+1}\}=[M^{m+1}][M^m]\{\{\theta^{m-1}\}+\{S^m\}\}+[M^{m+1}][N^m]\{\varphi^m\}+[M^{m+1}]\{S^{m+1}\}+[N^{m+1}]\{\varphi^{m+1}\}\qquad(18)$$

In the same way, the temperature distribution at successive r future times, $\tau=\tau^{m+r-1}$, can be represented

as follows:

$$\begin{aligned}\{\theta^{m+r-1}\} &= [M^{m+r-1}]\left\{\{\theta^{m+r-2}\}+\{S^{m+r-1}\}\right\} \\ &\quad + [N^{m+r-1}]\{\varphi^{m+r-1}\} \\ &= [M^{m+r-1}][M^{m+r-2}]\ldots[M^{m+1}][M^{m}]\left\{\{\theta^{m-1}\}+\{S^{m}\}\right\} \\ &\quad + [M^{m+r-1}][M^{m+r-2}]\ldots[M^{m+1}][N^{m}]\{\varphi^{m}\} \\ &\quad + [M^{m+r-1}][M^{m+r-2}]\ldots[M^{m+1}][N^{m}]\{\varphi^{m}\} \\ &\quad + [M^{m+r-1}][M^{m+r-2}]...[M^{m+1}]\{S^{m+1}\} \\ &\quad + [M^{m+r-1}][M^{m+r-2}]\ldots[M^{m+2}][N^{m+1}]\{\varphi^{m+1}\}+\ldots \\ &\quad + \ldots[M^{m+r-1}]\{S^{m+r-1}\}[N^{m+r-1}]\{\varphi^{m+r-1}\} \end{aligned} \tag{19}$$

A temporary assumption that the incident heat flux is constant over r future time steps is used to stabilize the estimated results in the inverse algorithms:

$$\varphi_j^{m+1} = \varphi_j^{m+2} = \ldots = \varphi_j^{m+r-1} = \varphi_j^{m} \quad \text{for} \quad j = 1,2,\ldots,p,p+1. \tag{20}$$

Then, the temperatures at each i-spatial grid ($i=1, 2,\ldots, p, p+1$) for each analysis interval can be expressed as follows:

$$\theta_i^{m+k} = h_i^{m+k,0} + \sum_{j=1}^{p+1} E_{i,j}\varphi_j^{m} \quad k = 0,1,2,...,r-1, \tag{21}$$

where

$$E_{i,j}^{m+k} = \sum_{l=0}^{k} h_{i,j}^{m+k,m+l}, \tag{22}$$

here

$$h_{i,j}^{m+k,m+l} = \begin{cases} [u_i][N^{m+k}]\{u_j\}, & l=k, \quad k=0,1,2,\ldots,r-1 \\ [u_i][N^{m+k}][N^{m+k-1}]\{u_j\}, & l=k-1, \quad k=1,2,\ldots,r-1 \\ [u_i][M^{m+k}][M^{m+k-1}][N^{m+k-2}]\{u_j\}, & l=k-2, \quad k=2,\ldots,r-2,r-1 \\ \vdots & \\ [u_i][M^{m+k}][M^{m+k-1}]\ldots[M^{m+2}][N^{m+1}]\{u_j\}, & l=1, \quad k=r-2,r-1 \\ [u_i][M^{m+k}][M^{m+k-1}]\ldots[M^{m+1}][N^{m}]\{u_j\}, & l=0, \quad k=r-1, \end{cases} \tag{23}$$

and

$$h_i^{m+k,0} = \begin{cases} [u_i][M^{m}]\left\{\{\theta^{m-1}\}+\{S^{m}\}\right\}, & k=0 \\ [u_i][M^{m+1}][M^{m}]\left\{\{\theta^{m-1}\}+\{S^{m}\}\right\}+[u_i][M^{m+1}]\{S^{m+1}\}, & k=1 \\ \vdots & \\ [u_i][M^{m+r-1}][M^{m+r-2}]\ldots[M^{m+1}][M^{m}]\left\{\{\theta^{m-1}\}+\{S^{m}\}\right\} & \\ +[u_i][M^{m+r-1}][M^{m+r-2}]\ldots[M^{m+1}]\{S^{m+1}\}+\ldots & \\ +[u_i][M^{m+r-1}][M^{m+r-2}]\{S^{m+r-2}\}+[u_i][M^{m+r-1}]\{S^{m+r-1}\}, & k=r-1 \end{cases} \tag{24}$$

$[u_i]$ denotes a unit row vector (i.e. a unit at the i-component). The subscripts i and j are the grid numbers of the temperature distribution location and the unknown incident heat flux location, respectively.

When $\tau=\tau^m$, the estimated parameter vectors $\{\varphi^1\}$, $\{\varphi^2\}$,..., and $\{\varphi^{m-1}\}$ have been evaluated and the task is now to determine the unknown incident heat flux vector $\{\varphi^m\}$. We can construct the following matrix equation

$$\{\vartheta\}_{r\cdot(p+1)\times 1} = \{\Phi\}_{r\cdot(p+1)\times(p+1)}\{\Psi\}_{(p+1)\times 1}. \tag{25}$$

After the known temperature distributions have been substituted into vector ϑ, the components of vector Ψ can be found using the linear least-squares-error method. The result is:

$$\Psi = (\Phi^T\Phi)^{-1}\Phi^T\vartheta. \tag{26}$$

This equation provides a sequential algorithm flow chart as in Fig. 3, which can be used to determine the unknown incident heat fluxes by increasing the value of m by one for each time step. Thereafter, the incident heat fluxes can be obtained iteratively along the temporal coordinate. Then, the incident heat-flux profiles on the silicon wafer required for tracking uniform temperature trajectories during RTP can be estimated.

4. Results and discussion

To ensure precision and optimal computer use, the numerical solution techniques with a $p=50$ equidistant grid and a time step of $\Delta\tau=0.0001$ were used in the present study to compute radial temperature distributions and incident-heat-flux profiles on a typical 300-mm-diameter 0.775-mm-thick silicon wafer. Numerical simulations linearly ramped from initial uniform temperature 27 °C (300 K) to a steady state of 1097 °C (1370 K) at an ambient temperature of $T_a=27$ °C (300

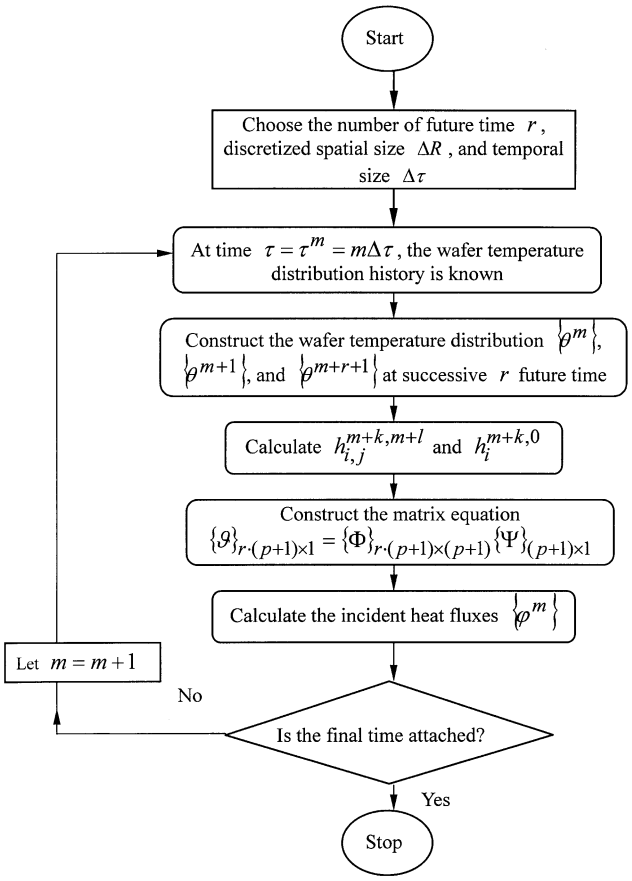

Fig. 3. Flow chart to the application of inverse problem algorithm.

K) were performed to examine the wafer temperature non-uniformity during RTP as a function of the ramp-up rate. Plots of the wafer center temperature for three linear ramp-up rates, 100, 200 and 300 °C/s are given in Fig. 4. While random errors of input data were added to the desired temperature trajectories, as described elsewhere [29]:

$$Y_1^n=\theta_1^n+\omega\sigma, \tag{27}$$

where the subscript 1 is the grid number of the spatial-coordinate at the wafer center, and the superscript n denotes the grid number of the temporal-coordinate. θ_1^n is the dimensionless 'exact' calculated temperature, Y_1^n is the dimensionless 'input-measured' temperature, σ is the standard deviation, and ω is a random number. The value of ω is calculated using the IMSL subroutine DRNNOR and chosen over the range $-2.576<\omega<2.576$, which represents the 99% confidence bound for the input temperature. In the present study, the respective dimensional input temperatures $T_1^n\pm0.7728$ °C and $T_1^n\pm3.864$ °C were simulated for the cases of $\sigma=0.001$ and 0.005.

We set

$$Y_2^n=Y_3^n=\ldots=Y_p^n=Y_{p+1}^n=Y_1^n, \tag{29}$$

as the desired uniform temperature tracking during processing, for the known temperature distribution used in the inverse problem algorithm to evaluate the unknown incident heat-flux profiles over the wafer. After that, the radial temperature distribution across the wafer could be computed. Finally, the wafer temperature non-uniformity both during transient and steady state was investigated.

The inverse wafer-center-temperature trajectory results for three linear ramp-up rates with various random errors, $\sigma=0.0$ (means 'exact'), $\sigma=0.001$ and $\sigma=0.005$, are shown in Fig. 4. The transient times from initial uniform 27 °C (300 K) to reach higher steady state 1097 °C (1370 K) are approximately 10.7, 5.35, and 3.57 s for 100, 200 and 300 °C/s, respectively. We can see that the differences between inverse-results and exact-result resulting from the random errors were reasonable. The greater the random errors, the less accurate the inverse-results.

Fig. 5a–c show the three-dimensional graph of the inverse incident-heat-flux profile results with error $\sigma=0.0$ for uniform temperature tracking of 100, 200 and 300 °C/s ramp-up rates, respectively. The axis 'Radial Position' shows the distance from the wafer center in centimeters. The axis 'Time' represents the time during this temperature transition. The vertical axis represents the calculated incident-heat-flux profile yielded by the inverse problem algorithm. The incident heat-flux energy is absorbed by top and bottom surfaces of the wafer, and the heat losses also occur at all wafer surfaces. Ramping of wafer temperature takes place when there is an excess of the absorbed energy over the energy of

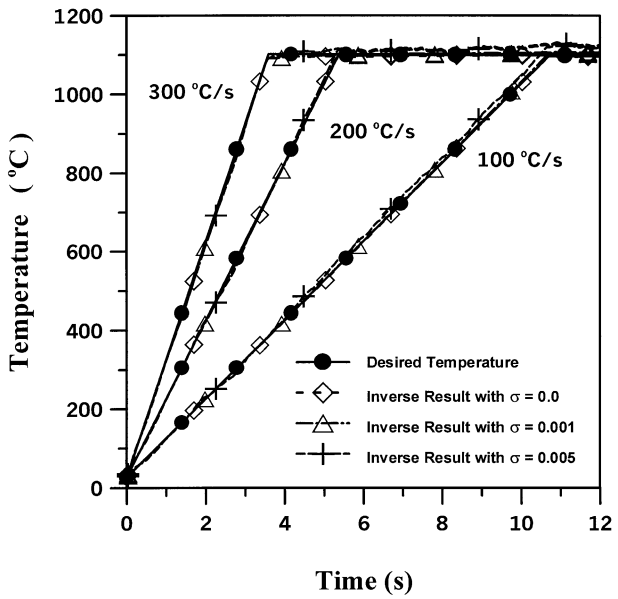

Fig. 4. Desired uniform temperature trajectories for 100, 200 and 300 °C/s linear ramp-up rates, and inverse results for random errors of $\sigma=0.0$, 0.001, and 0.005.

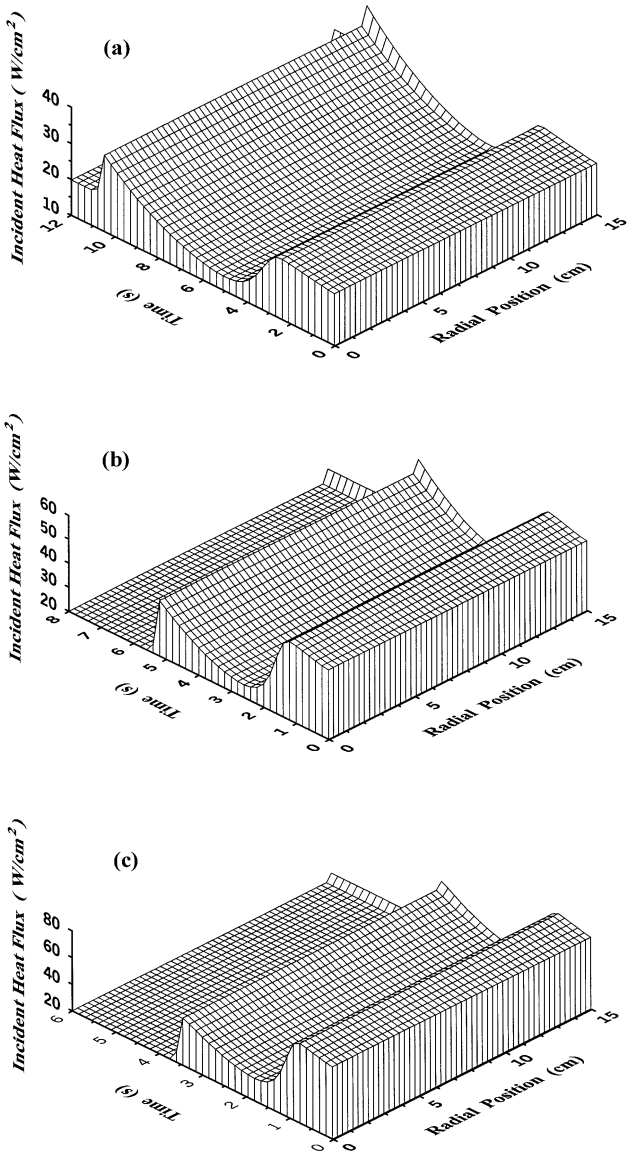

Fig. 5. Inverse results of incident-heat-flux profiles for a random error of $\sigma = 0.0$ at linear ramp-up rates of (a) 100 (b) 200 and (c) 300 °C/s.

heat losses. During the initial transient phase, the wafer temperature increases with the increasing absorbed energy, and heat losses also increase with the increasing wafer temperature. The initial absorbed energy, required for wafer uniform-temperature tracking, was larger than that during other periods for this temperature transient. This can be explained by the energy equation for the silicon wafer [Eq. (11)]. On the left-hand side of the equation, the term $\partial\theta/\partial\tau$ is constant during processing because of constant temperature ramp-up rate. Due to uniform temperature tracking, on the right-hand side, the radial temperature gradient terms, $\partial^2\theta/\partial R^2 + 1/R \times (\partial\theta/\partial R)$, are approximately zero. The temperature-dependent emissivity $\varepsilon(\theta)$ is 0.3 at initial lower temperatures, and 0.68 at higher temperatures from 800 to 1700 K (see Fig. 2). However, the temperature-dependent specific heat capacity $C(\theta)$ does not vary much with increasing wafer temperature. Thus the term, $\varepsilon(\theta)/C(\theta)$, on the right-hand side of the energy equation at the higher wafer temperature is greater than that at the initial lower wafer temperatures. To balance the energy equation during constant temperature ramp-up processing, the net of incident heat flux and heat losses, $Q(R,\tau) + A(1-\theta^4)$, is larger in the initial transient phase because of the lower $\varepsilon(\theta)/C(\theta)$. This results in larger incident heat fluxes being needed in the initial phase. When the wafer temperature reaches approximately 800 K, the necessary incident heat fluxes are reduced due to the larger $\varepsilon(\theta)/C(\theta)$. While, in the higher temperature periods, the heat losses occurring at all surfaces become much greater and greater incident heat fluxes are needed to counteract the heat losses. Thus, the necessary incident heat fluxes increase with the increasing wafer temperature until the wafer reaches the higher steady state (the temperature ramp-up rate becomes zero) and the necessary incident heat fluxes also become steady because the absorbed energy balances the heat losses. For these cases of three different ramp-up rates, they are all the same at the value of 20 W/cm^2 in this temperature transient during steady-state processing (after the time of 10.7, 5.35, and 3.57 s for 100, 200 and 300 °C/s, respectively). Due to the additional heat losses at the wafer edges, more heat compensation is needed at the wafer perimeters during processing. Since the wafer edge is slightly cooler than the center during the initial transient phase, the edge heating compensation was not significant. As the ramp-up proceeds, the temperature-dependent thermal conductivity $K(\theta)$ decreases with the increasing wafer temperature (see Fig. 2) and the temperature-dependent emissivity mentioned above; the edge heat losses increase with the increasing wafer temperature from the boundary condition described in Eq. (14). Edge-heating compensations are increasingly modulated to meet the requirement of uniform temperature tracking. Thus, the additional amounts of energy directed to the edge to offset the edge heat losses are apparent. Finally, as the wafer reaches the steady state, the edge heating compensation approaches the constant heat-flux scaling factor of 1.26 (25.2 W/cm^2) for a uniform temperature of 1097 °C (1370 K) for these cases of three different linear ramp-up rates. These figures show that the dynamically individual control of incident heat fluxes is needed for tracking the desired uniform-temperature trajectories during RTP. Fig. 6a,b show the inverse incident-heat-flux profiles results from tracking the desired uniform-temperature trajectory of 200 °C/s ramp-up rate with random errors of $\sigma = 0.001$ and $\sigma = 0.005$, respectively. These results are almost the same as those shown in Fig. 5b. However, the incident heat flux profiles have to be dynamically modulated according to the random-

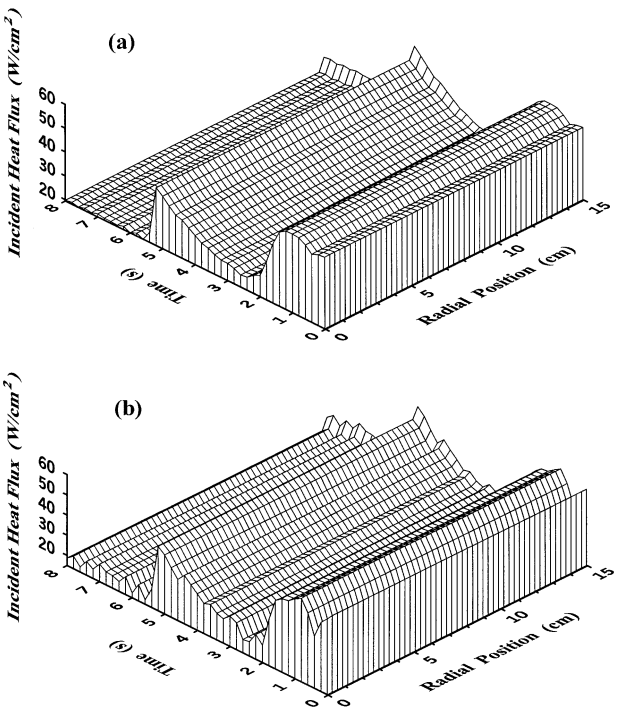

Fig. 6. Inverse results of incident-heat-flux profiles at a linear 200 °C/s ramp-up rate for random errors of (a) $\sigma = 0.001$ and (b) $\sigma = 0.005$.

error effects to maintain temperature uniformity during both transient and steady state.

Ideally, if it were not for the edge of a wafer, temperature uniformity could be achieved by applying uniform heat-flux profiles of varying strengths on the top and bottom surfaces of a wafer to achieve uniform temperature-trajectory tracking. However, temperature distortions develop near the wafer edges during processing. Many rapid thermal processes [20] direct additional amounts of energy toward the edges to counteract these temperature non-uniformities occurring at the wafer edges achieving results similar to our inverse incident-heat-flux results shown in Fig. 5. Inverse dynamic incident-heat-flux profile results on temperature non-uniformity for the three linear ramp-up rates with a random error of $\sigma = 0.0$ are shown in Fig. 7a–c, respectively. The vertical axes represent temperature non-uniformity graphed according to the temperature differences between points on the wafer and the wafer's center. These figures show that when incident heat-flux profiles are controlled as our inverse-results, temperature differences develop at the edge. Initially, the temperature difference is not significant, however, as the ramp-up proceeds, the temperature non-uniformity developed at the edge increases with increasing edge-heating compensation, as shown in Fig. 5. When temperature-dependent emissivity changes from 0.3 to 0.68, a sudden variation is seen in the temperature non-uniformity profiles (for instance, at 4 s in Fig. 7a, in the ramp-up rate of 100 °C/s). When the wafer reaches the higher steady state, the incident-heat-flux profile changes from the transient stage to the steady stage, the temperature non-uniformity drops gradually and approaches the steady-state. Thus, edge-heating compensation has an overheating effect on temperature uniformity during processing.

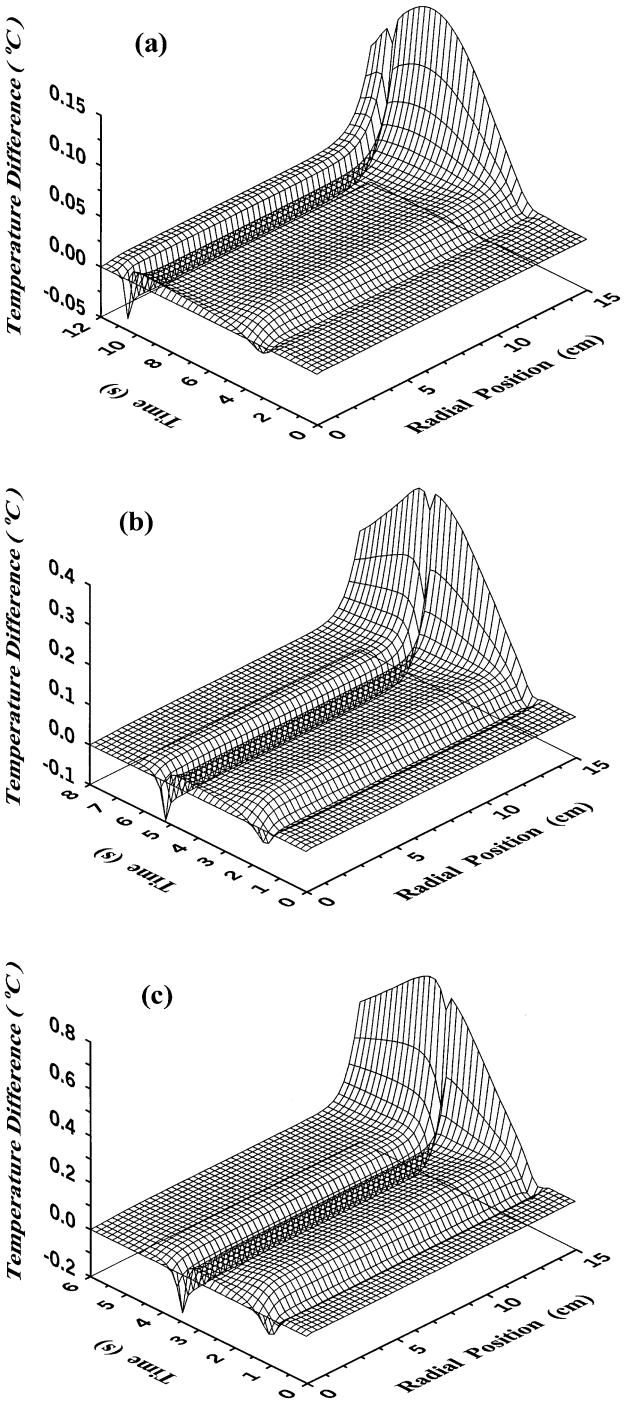

Fig. 7. Inverse results of temperature non-uniformity for a random error of $\sigma = 0.0$ at linear ramp-up rates of (a) 100 (b) 200 and (c) 300 °C/s.

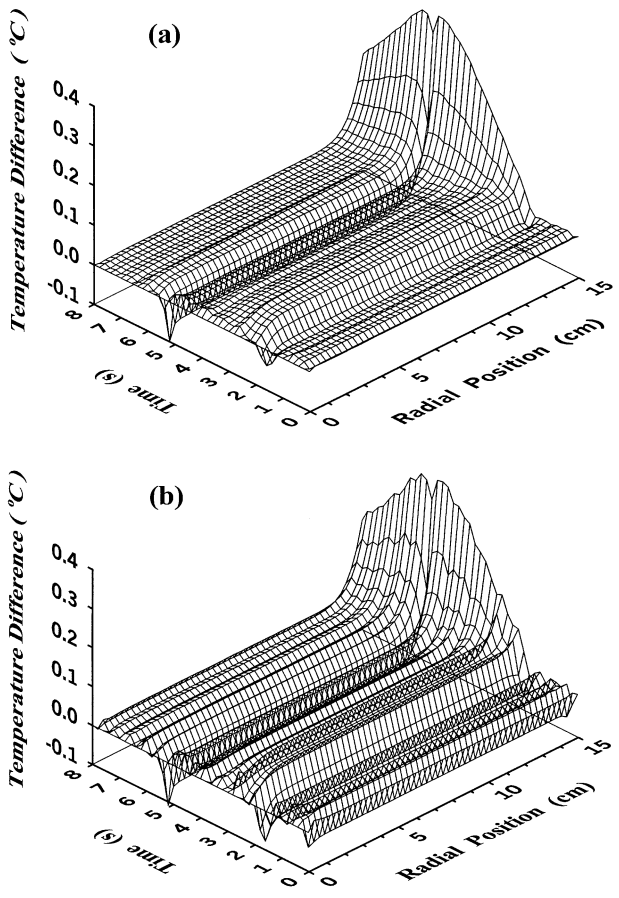

Fig. 8. Inverse results of temperature non-uniformity at a linear 200 °C/s ramp-up rate for random errors of (a) $\sigma = 0.001$ and (b) $\sigma = 0.005$.

Generally, the temperature over the wafer has to be maintained within 2 °C of the wafer center during rapid thermal processing [5]. Fig. 7a–c show that the temperature difference from the wafer center is not significant for our present inverse incident-heat-flux profiles. Even though during transient periods, the resulting maximum temperature differences were 0.152, 0.389 and 0.658 °C for the cases of 100, 200 and 300 °C/s ramp-up rates, respectively, it was found that temperature non-uniformity occurring during the ramp increased with the ramp-up rate, but within 1 °C during processing. Fig. 8a,b show the respective inverse-results on wafer temperature non-uniformity for the 200 °C/s ramp-up rate when the random errors of $\sigma = 0.001$ and 0.005 were introduced. The temperature non-uniformity increased as the random error was increased, as expected. However, the maximum temperature difference was less than 0.41 °C although the random error did reach 3.864 °C (in the case of $\sigma = 0.005$).

Fig. 9 illustrates the resulting maximum temperature differences ($|\Delta T|$, the absolute value of temperature difference between the wafer edge and the center) during transients as a function of the desired linear ramp-up rates for random errors of $\sigma = 0.0$, 0.001, 0.003 and 0.005, respectively. Our present results show that the maximum temperature differences occurring during the ramp increase with the ramp-up rate. Although a linear ramp-up rate of 300 °C/s was used and random errors did reach 3.864 °C (in the case of $\sigma = 0.005$), the temperature over the wafer was maintained within 0.665 °C of the wafer center if the incident-heat-flux profiles were dynamically controlled according to the inverse-results. Furthermore, random-error effects on temperature non-uniformity were not apparent when the desired linear ramp-up rate exceeded 200 °C/s. When ramp-up rate was lower than 150 °C/s, the random-error effect of $\sigma = 0.005$ on temperature non-uniformity was enlarged due to the dimensional error of 3.864 °C, but remained under 0.35 °C. These temperature non-uniformities could be acceptable in the advanced rapid thermal processing system. The extended model can then be combined with view factor information from more detailed modeling analysis to provide a realistic model for controller design for temperature uniformity problem.

5. Conclusion

This article presents a systematic method to the application of inverse problem algorithms for uniform temperature tracking of several different linear ramp-up rates in rapid thermal processing. Temperature-dependent thermal properties of the silicon wafer were considered in this study. Using a one-dimensional thermal model, temperature solutions for the inverse-method matrices can be constructed by applying the finite-

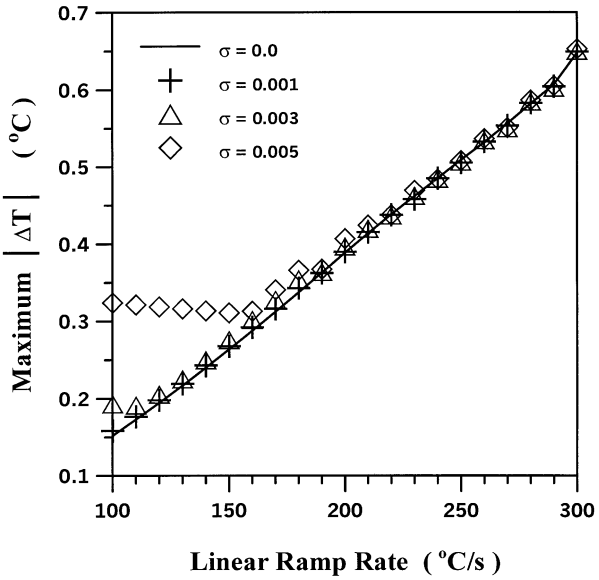

Fig. 9. Inverse results of maximum temperature difference for random errors of $\sigma = 0.0$, 0.001, 0.003 and 0.005 as a function of the linear ramp-up rate.

difference scheme to calculate the desired incident-heat-flux profiles required for uniform temperature tracking. In the present study, the wafer was ramped-up from initial uniform 27 °C temperature to a steady state of 1097 °C via simulation at several linear ramp-up rates. The resulting maximum temperature differences as a function of the desired linear ramp-up rates using several random errors were investigated. The maximum temperature differences in our present study were only 0.152, 0.388 and 0.658 °C, respectively, for the 100, 200 and 300 °C/s ramp-up rates when the incident heat fluxes on the wafer could be dynamically controlled according to the inverse-results. Temperature non-uniformity occurring during the ramp increased with the ramp-up rate. Furthermore, random-error effects on temperature non-uniformity were not apparent when the desired linear ramp-up rate exceeded 200 °C/s. When ramp-up rate was lower than 150 °C/s, the random-error effect of $\sigma = 0.005$ on temperature non-uniformity was enlarged due to the error of 3.864 °C, but remained under 0.35 °C. Although a linear 300 °C/s ramp-up rate was used and the error did reach 3.864 °C, the resulting maximum temperature differences were not significant and remained under 0.665 °C. These temperature non-uniformities could be acceptable in the advanced rapid thermal processing system.

References

[1] V.E. Borisenko, P.J. Hesketh, Rapid Thermal Processing of Semiconductors, Plenum Press, New York, 1997.
[2] K.S. Balakrishnan, T.F. Edgar, Thin Solid Films 365 (2000) 322.
[3] A. Kersch, T. Schafbauer, Thin Solid Films 365 (2000) 307.
[4] A. Theodoropoulou, E. Zafiriou, R.A. Adomaitis, IEEE Trans. Semicond. Manuf. 12 (1999) 87.
[5] F.Y. Sorrell, M.J. Fordham, M.C. Öztürk, J.J. Wortman, IEEE Trans. Electron Devices 39 (1992) 75.
[6] M. Glück, W. Lerch, D. Löffelmacher, M. Hauf, U. Kreiser, Microelectron. Eng. 45 (1999) 237.
[7] K.C. Lee, H.Y. Chang, H. Chang, J.G. Hwu, T.S. Wung, IEEE Trans. Semicond. Manuf. 12 (1999) 340.
[8] C.H. Huang, C.C. Yu, S.H. Shen, Automatica 36 (2000) 705.
[9] Y.M. Cho, P. Gyugyi, IEEE Trans. Contr. Sys. Tech. 5 (1997) 644.
[10] A. Theodoropoulou, R.A. Adomaitis, E. Zafiriou, IEEE Trans. Semicond. Manuf. 11 (1998) 85.
[11] S. Banerjee, J.V. Cole, K.F. Jensen, IEEE Trans. Semicond. Manuf. 11 (1998) 266.
[12] H.A. Lord, IEEE Trans. Semicond. Manuf. 1 (1988) 105.
[13] C. Hill, S. Jones, D. Boys, Rapid thermal annealing — theory and practice, in: R.A. Levy (Ed.), Reduced Thermal Processing for ULSI, NATO ASI Series B: Physics, Plenum Press, New York, 1988, pp. 143–180.
[14] R. Kakoschke, E. Bubmann, H. Föll, Appl. Phys. A 50 (1990) 141.
[15] R.S. Gyurcsik, T.J. Riley, F.Y. Sorrell, IEEE Trans. Semicond. Manuf. 4 (1991) 9.
[16] S.A. Norman, IEEE Trans. Electron Devices 39 (1992) 205.
[17] J.P. Zöllner, K. Ullrich, J. Pezoldt, G. Eichhorn, Appl. Surf. Sci. 69 (1993) 193.
[18] T.J. Riley, R.S. Gyurcsik, Mater. Res. Soc. Symp. Proc. 303 (1993) 223.
[19] Y.M. Cho, A. Paulraj, T. Kailath, G.A. Xu, IEEE Trans. Semicond. Manuf. 7 (1994) 34.
[20] R.H. Perkins, T.J. Riley, R.S. Gyurcsik, IEEE Trans. Semicond. Manuf. 8 (1995) 272.
[21] Y.K. Jan, C.A. Lin, IEEE Trans. Semicond. Manuf. 11 (1998) 75.
[22] M. Janicki, M. Zubert, A. Napieralski, Microelectron. J. 30 (1999) 1099.
[23] M. Kurz, G. Müller, J. Cryst. Growth 208 (2000) 341.
[24] J.V. Beck, B. Blackwell, C.R. St.Clair, Inverse Heat Conduction — Ill-posed Problem, Wiley, New York, 1985.
[25] K. Kurpisz, A.J. Nowak, Inverse Thermal Problems, Computational Mechanics Publications, Boston, 1995.
[26] S. Lin, H.S. Chu, IEEE Trans. Semicond. Manuf. 13 (2000) 448.
[27] S. Lin, H.S. Chu, Microscale Thermophys. Eng. 4 (2000) 245.
[28] A. Virzi, J. Cryst. Growth 112 (1991) 699.
[29] C.Y. Yang, Int. J. Heat Mass Transfer 42 (1999) 345.