Korean J. Chem. Eng., **26**(2), 307-312 (2009)

Iterative identification of temperature dynamics in single wafer rapid thermal processing

Wonhui Cho*, Thomas F. Edgar*, and Jietae Lee,†**

*Department of Chemical Engineering, University of Texas, Austin, TX78712, U.S.A.
**Department of Chemical Engineering, Kyungpook National University, Daegu 702-701, Korea
(*Received 19 November 2007 • accepted 3 September 2008*)

Abstract−As the standard size of silicon wafers grows and performance specifications of integrated circuits become more demanding, a better control system to improve the processing time, uniformity and repeatability in rapid thermal processing (RTP) is needed. Identification and control are complicated because of nonlinearity, drift and the time-varying nature of the wafer dynamics. Various physical models for RTP are available. For control system design they can be approximated by diagonal nonlinear first order dynamics with multivariable static gains. However, these model structures of RTP have not been exploited for identification and control. Here, an identification method that iteratively updates the multivariable static gains is proposed. It simplifies the identification procedure and improves the accuracy of the identified model, especially the static gains, whose accurate identification is very important for better control.

Key words: Temperature Control, Rapid Thermal Processing, Single Wafer, Iterative Identification, Nonlinear Model Identification

INTRODUCTION

Single wafer rapid thermal processing (RTP) consists of three steps of rapid heating up to a specified temperature, processing under a constant operating temperature and then cooling down. RTP is used for various processes in the fabrication of semiconductor devices, such as rapid thermal annealing, oxidation, nitration and chemical vapor deposition. A control system is required to follow the desired temperature trajectory while maintaining the desired temperature uniformity. RTP control becomes more stringent as the sizes of integrated circuits become smaller and wafer diameters grow. In control of RTP, there are several difficulties such as high nonlinearity of the heating process, drift, time-varying nature of wafer processing, and non-contact measurements of temperatures. A review of RTP control has been provided by Edgar et al. [1].

Various control methods have been applied to RTP, including decoupling control [1], iterative learning control [2,3], adaptive control [4], internal model control [1], model-based control [5] and nonlinear model predictive control [6]. Huang et al. [7] used a proportional-double integral-derivative (PI^{2}D) controller to eliminate offset error during the heating step. To design these control systems, accurate models for RTP are required. Models based on the first principles are available [6,8-10]. They have been shown to be approximated by diagonal nonlinear first order dynamics with multivariable static gains [5,7,11]. However, these model structures of RTP are not yet utilized for identification and control. When the multivariable static gains are identified accurately, the process controller can be decoupled and a linearizing control system having predictable responses can be designed.

Identification of the multivariable static gains is difficult due to drift and nonlinear time-varying dynamics. A poor model cannot ensure the given control performance. For the purpose of control system design, iterative identification methods are available [12]. Models are improved to guarantee a specified control performance, such as decoupling. An iterative identification method is applied to identify the multivariable static gains of RTP. At a given temperature level, locally linear models with accurate multivariable static gains are identified iteratively. Then, a nonlinear model is constructed by using analytic radiation terms [11] without model identifications at various temperature levels. The proposed method will simplify the identification procedure and improve models for the purpose of control system design.

SINGLE WAFER RAPID THERMAL PROCESSING AND TEMPERATURE CONTROL

This work was carried out with experimental RTP equipment developed by Texas Instruments to fabricate the 6-in wafers. The main chamber is shown in Fig. 1 [13]. Physical models for the wafer dynamics with different complexities have been proposed [6,7,14]. Here, a lumped model as in Huang et al. [7] and Kersch and Schafbauer [9] is used.

$$\rho V_i c_P \frac{dT_i}{dt} = q_i^{rad.waf} + q_i^{conv} + q_i^{cond.waf} + q_i^{surr.} + q_i^{lamps} \tag{1}$$

where i the index of the radial cell on the wafer with volume V_i, mass density ρ, heat capacity c_P, and temperature T_i.

The radiation term can be described as $q_i^{rad.waf} = A_i \sum_{j=1}^{N} a_i^{eff} \Gamma_{ij} e_j^{eff} \sigma T_j^4 - A_i e_i^{eff} \sigma T_i^4$, where A_i is the surface area of the i-th cell, a_i^{eff} is the effective absorptivity, and e_i^{eff} is the effective emissivity. The first term in Eq. (2) Γ_{ij} describes the fraction of radiation emitted from the surface element j. These radiative exchange factors depend on the geometry of the RTP chamber and the optical properties of the wafer surface. They need complicated computations such as Monte Carlo simulations [9]. Schaper et al. [8] showed that the off-diagonal terms in Γ_{ij} are small and can be ignored for the purpose of control system

†To whom correspondence should be addressed.
E-mail: jtlee@knu.ac.kr

design. The convective heat transfer term is given as $q_i^{conv}=-A_ih_i(T_i-T_{gas})$, where h_i is the heat transfer coefficient. The convective heat transfer coefficient depends on the gas properties and flow pattern [15]. The conduction term depends on the gradient of wafer temperature across the radial direction and is known to be negligible compared to radiation and convective terms. The heat exchange term q_i^{wall} is heat addition due to the heated wall. Its dynamics is about ten times slower than that of the wafer. Usually, the wall temperature is not measured but is considered as a slow drift. Heat addition from the lamp power is given as $q_i^{lamp}=A_ia_i^{eff}\sum_{j=1}^{N}L_{ij}P_j$.

In summary we have

$$\rho V_i c_P \frac{dT_i}{dt} = -A_i(h_iT_i + e_i^{eff}\sigma T_i^4) + A_ia_i^{eff}\sum_{j=1}^{N}L_{ij}P_j + d_i(t) \tag{2}$$

where $d_i(t)$ includes drift and effects of all ignored heat transfer effects. Eq. (2) has several parameters that are difficult to estimate theoretically. Cho et al. [11] conducted experiments for the RTP equipment in Fig. 1. They obtained local linear models for temperatures between 400 K and 700 K based on experimental data and constructed a nonlinear model by fitting those linear models. The model [11] is

$$\dot{T}(t) = -f(T(t)) + BP(t) + d(t) \tag{3}$$

where

$$T = (T_1, T_2, T_3)^T$$

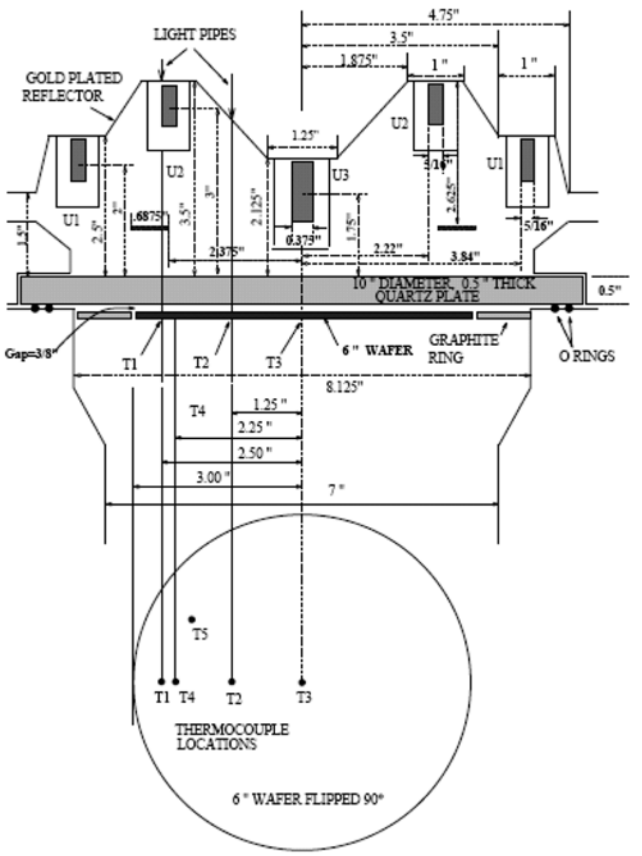

Fig. 1. Experimental single wafer rapid thermal processing equipment developed by Texas Instruments [13].

$$P = (P_1, P_2, P_3)^T$$

$$f(T) = \begin{bmatrix} 3.6170\times10^{-2}T_1 + 4.7203\times10^{-11}T_1^4 \\ 3.4455\times10^{-2}T_2 + 4.2441\times10^{-11}T_2^4 \\ 4.5144\times10^{-2}T_3 + 5.5126\times10^{-11}T_3^4 \end{bmatrix}$$

$$B = \begin{bmatrix} 4.9823\times10^{-2} & 3.6034\times10^{-2} & 0.6401\times10^{-2} \\ 3.6167\times10^{-2} & 4.1290\times10^{-2} & 0.8654\times10^{-2} \\ 4.6108\times10^{-2} & 4.8056\times10^{-2} & 1.8020\times10^{-2} \end{bmatrix}$$

d(t)=initial and drift term

Here this model is used as a real process to be identified.

Models of Eqs. (2) and (3) have uncertainties such as ignored conduction terms and quartz window dividing the chamber and lamp dynamics. They are valid for small perturbations of T_1, T_2 and T_3 that are maintained to be same, and are effective for control system design purposes because control systems will maintain wafer temperatures to be the same. However, to identify them, we are very cautious because models of Eqs. (2) and (3) have structural uncertainties. Large deviations of wafer temperatures will activate the conduction terms and quartz window dynamics, and can provide poor models. That is, identification of the full nonlinear models with large perturbations is not recommended. Here, an identification method that uses small perturbations is proposed. For this, we identify local linear models with perturbations whose sizes and times are minimized. Further, they are refined iteratively. Then nonlinear models are restored so that their linearized ones are the same as the identified linear models.

LOCAL LINEAR MODEL IDENTIFICATION

For a given operating condition, $T_{1s}=T_{2s}=T_{3s}=\overline{T}$, the system of Eq. (3) can be linearized as

$$\begin{bmatrix} \tilde{T}_1(s) \\ \tilde{T}_2(s) \\ \tilde{T}_3(s) \end{bmatrix} = G(s)\begin{bmatrix} \tilde{P}_1(s) \\ \tilde{P}_2(s) \\ \tilde{P}_3(s) \end{bmatrix} = \Lambda(s)K\begin{bmatrix} \tilde{P}_1(s) \\ \tilde{P}_2(s) \\ \tilde{P}_3(s) \end{bmatrix}$$
$$= \begin{bmatrix} 1/(\tau_1 s+1) & 0 & 0 \\ 0 & 1/(\tau_2 s+1) & 0 \\ 0 & 0 & 1/(\tau_3 s+1) \end{bmatrix}\begin{bmatrix} k_{11} & k_{12} & k_{13} \\ k_{21} & k_{22} & k_{23} \\ k_{31} & k_{32} & k_{33} \end{bmatrix}\begin{bmatrix} \tilde{P}_1(s) \\ \tilde{P}_2(s) \\ \tilde{P}_3(s) \end{bmatrix} \tag{4}$$

Here the tilde (~) means a deviation variable. The dynamic equation for T_1 is

$$\tau_1\dot{\tilde{T}}_1(t) = \tilde{T}_1(t) + k_{11}\tilde{P}_1(t) + k_{12}\tilde{P}_2(t) + k_{13}\tilde{P}_3(t) + \tilde{d}(t) \tag{5}$$

It can be discretized as

$$\tilde{T}_1(k+1) = \alpha\tilde{T}_1(k) + (1-\alpha)(k_{11}\tilde{P}_1(k) + k_{12}\tilde{P}_2(k) + k_{13}\tilde{P}_3(k)) + \tilde{d}(k+1) \tag{6}$$

where k means the k-th sampling and $\alpha=\exp(-h/\tau_1)$, h=sampling interval of 5 seconds.

At the start of the identification test, we set the deviation variables for temperature outputs and power inputs be zero. To reduce the effect of d(t), the experimental test times are minimized. For this, a sequential negative pulse with the size of s_1, a positive pulse

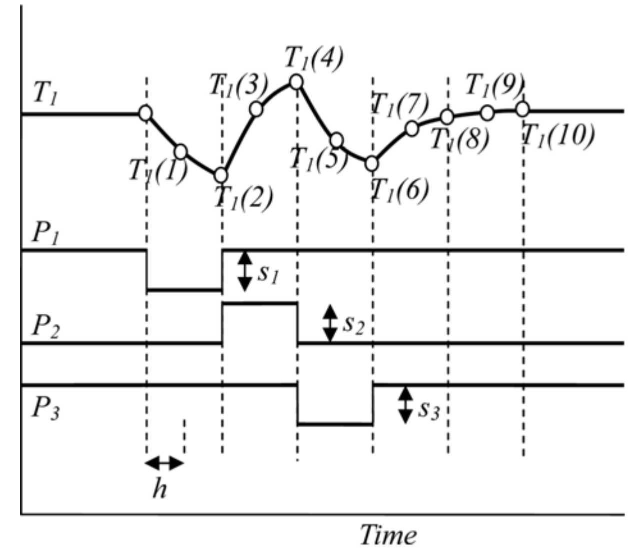

Fig. 2. Pulse input signals and their typical responses.

with the size of s_2 and a negative pulse with the size of s_3 are the forcing inputs as shown in Fig. 2. The pulse durations are all two sampling times. For the test inputs in Fig. 2, we have

$$\begin{bmatrix} \tilde{T}_1(1) \\ \tilde{T}_1(2) \\ \tilde{T}_1(3) \\ \tilde{T}_1(4) \\ \tilde{T}_1(5) \\ \tilde{T}_1(6) \\ \vdots \\ \tilde{T}_1(10) \end{bmatrix} = -\begin{bmatrix} 1 \\ \alpha+1 \\ \alpha^2+\alpha \\ \alpha^3+\alpha^2 \\ \alpha^4+\alpha^3 \\ \alpha^5+\alpha^4 \\ \vdots \\ \alpha^9+\alpha^8 \end{bmatrix} \beta k_{11} s_1 + \begin{bmatrix} 0 \\ 0 \\ 1 \\ \alpha+1 \\ \alpha^2+\alpha \\ \alpha^3+\alpha^2 \\ \vdots \\ \alpha^7+\alpha^6 \end{bmatrix} \beta k_{12} s_2$$

$$-\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ \alpha+1 \\ \vdots \\ \alpha^5+\alpha^4 \end{bmatrix} \beta k_{13} s_3 + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} d_0 + \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ \vdots \\ 9 \end{bmatrix} d_1 \tag{7}$$

Here, d(t) is assumed linear ($d(t)=d_0+d_1 t$) during the test.

By applying the nonlinear least squares method to Eq. (7), we obtain k_{11}, k_{12}, k_{13}, and τ_1. Specifically, the time constant τ_1 is optimized such that it minimizes the sum of squares for the linear least squares fit of k_{11}, k_{12}, k_{13}, d_0 and d_1.

Similarly, applying the above nonlinear least squares method to the responses of T_2 and T_3 for the same series of inputs, we can obtain all the other parameters of Eq. (4). Fig. 3 shows responses for this test. The wafer is heated to 500 °C with loosely tuned PI controllers [13,16]. At 100 seconds, controllers are put in manual and test signals are applied sequentially. In this simulation, small dynamic drifts with a time constant of 80 seconds are added. The identified linear model is given in Table 1.

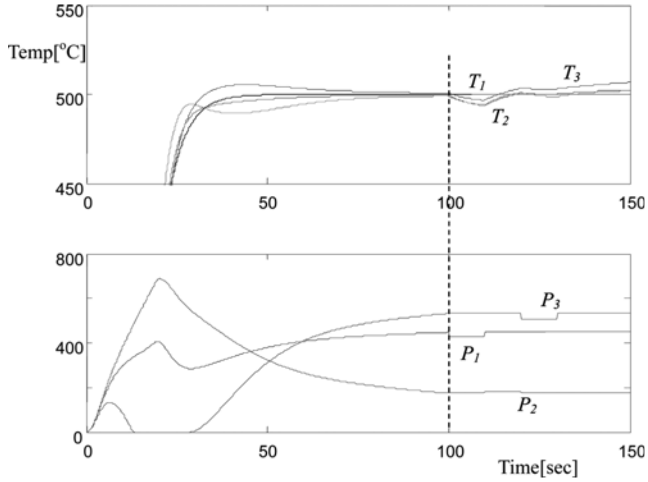

Fig. 3. Identification test inputs and temperature responses for RTP.

A longer pulse width improves the identification accuracy of the steady state gains, but it increases the identification times. Because a first order dynamic system is identified, a more complicated input sequence such as pseudo-random binary sequence is not needed.

ITERATIVE REFINEMENT

For controlling an RTP system described by Eq. (3), accurate identification of matrix B is very important. When B is found accurately, the system of Eq. (3) can be decoupled and the nonlinear first order system without time delay can be controlled robustly. Lee et al. [12] proposed an iterative identification method for ill-conditioned processes where accurate identification of models is very important, provided it converges. In this paper this iterative method is applied to RTP. First, using the identified matrix in the previous section, we decouple the RTP system. Finding the gain matrix of the decoupled system, we update the gain matrix of RTP system and repeat this decouple-update procedure.

Let the k-th identified gain matrix K_k and decouple RTP by $\tilde{P}(s)=K_k^{-1}\tilde{Q}(s)$ (Fig. 4). Then we have

$$\tilde{T}(s)=G(s)K_k^{-1}\tilde{Q}(s) \tag{8}$$

Applying the above test procedure, we obtain the multivariable steady state gain matrix, $\tilde{K}$, between $\tilde{T}$ and $\tilde{Q}$. A refined (k+1)-th gain matrix of RTP becomes

$$K_{k+1}=\tilde{K}K_k \tag{9}$$

When the process is not ill-conditioned, two or three iterations are sufficient. Fig. 5 shows responses of the second iteration test. Table 1 shows the refined gain matrix.

NONLINEAR MODEL AND CONTROL SYSTEM

In Eq. (2), coefficients for the term T_i^4 are

$$\kappa_i \equiv \frac{A_i e_i^{eff}\sigma}{\rho V_i c_P} = \frac{2e_i^{eff}\sigma}{\rho z c_P} \tag{10}$$

where z is the thickness of wafer. The coefficient κ_i is dependent

Table 1. Results of the iterative identification

Iteration no.	1	2
Decoupler	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0.4162 & 0.2806 & 0.0506 \\ 0.3484 & 0.3298 & 0.0761 \\ 0.3215 & 0.3157 & 0.1186 \end{bmatrix}^{-1}$
K	$\begin{bmatrix} 0.4162 & 0.2806 & 0.0506 \\ 0.3484 & 0.3298 & 0.0761 \\ 0.3215 & 0.3157 & 0.1186 \end{bmatrix}$	$\begin{bmatrix} 0.4255 & 0.3066 & 0.0541 \\ 0.3588 & 0.3925 & 0.0827 \\ 0.3150 & 0.3244 & 0.1202 \end{bmatrix}$
$100{*}B_m$	$\begin{bmatrix} 5.1427 & 3.4672 & 0.6252 \\ 3.9590 & 3.7476 & 0.8647 \\ 4.8076 & 4.7209 & 1.7735 \end{bmatrix}$	$\begin{bmatrix} 5.0500 & 3.6388 & 0.6421 \\ 3.9025 & 4.2690 & 0.8995 \\ 4.7744 & 4.9169 & 1.8218 \end{bmatrix}$
BB_m^{-1}	$\begin{bmatrix} 0.78 & 0.34 & -0.08 \\ -0.54 & 1.93 & -0.26 \\ -0.20 & 0.25 & 0.96 \end{bmatrix}$	$\begin{bmatrix} 0.98 & 0.01 & 0.00 \\ -0.09 & 1.07 & -0.02 \\ -0.04 & 0.01 & 1.00 \end{bmatrix}$
RGA (BB_m^{-1})	$\begin{bmatrix} 0.90 & 0.11 & -0.01 \\ 0.11 & 0.85 & 0.04 \\ -0.01 & 0.04 & 0.97 \end{bmatrix}$	$\begin{bmatrix} 0.999 & 0.001 & 0.000 \\ 0.001 & 0.999 & 0.000 \\ 0.000 & 0.000 & 1.000 \end{bmatrix}$
τ	8.0931 (8.1050*) 8.8003 (8.8600*) 6.6873 (6.8031*)	8.4258 9.1942 6.5977

*Exact values

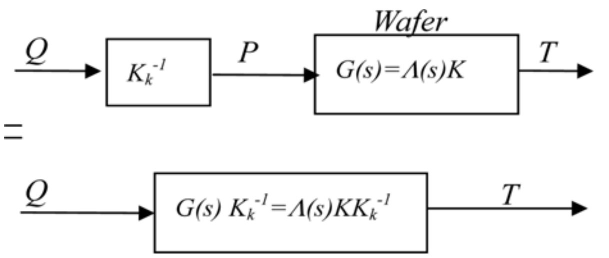

Fig. 4. Iterative identification procedure.

on physical properties of the wafer. Here, instead of identifying it from experimental responses, its theoretical estimate is used. Lee et al. [14] used ρ=2330[kg/m^3], z=6.75×10^{-4}[m], c_P=748.38+0.1678T [Ws/kgK], e_i^{eff}=0.6, and σ=5.67×10^{-8}[W/m^{2}K^4]. In this case, the value κ is between 4.75×10^{-11} and 5.03×10^{-11} for T∈[400 ºC, 700 ºC]. When data in Dassu et al. [6] are used, it is between 3.91×10^{-11} and 5.91×10^{-11}. Here we used a constant, κ=5.00×10^{-11}[1/sK^3], instead of letting unknown parameters for a simpler identification [11]. The RTP model becomes

$$\begin{bmatrix} \dot{T}_1 \\ \dot{T}_2 \\ \dot{T}_3 \end{bmatrix} = -\begin{bmatrix} a_{11}T_1+\kappa T_1^4 \\ a_{21}T_2+\kappa T_2^4 \\ a_{31}T_3+\kappa T_3^4 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}\begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} + d(t),\quad \kappa=5.00\times10^{-11} \tag{11}$$

The nonlinear model of Eq. (12) can be linearized as

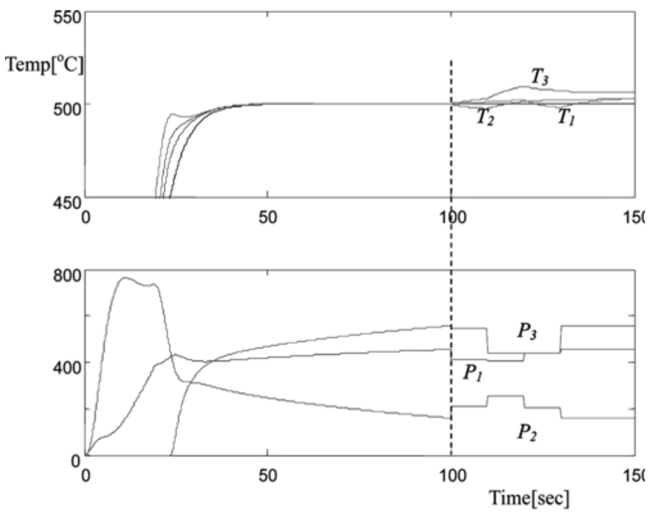

Fig. 5. Identification test inputs of the second iteration and temperature responses for RTP.

$$\begin{bmatrix} \dot{\tilde{T}}_1 \\ \dot{\tilde{T}}_2 \\ \dot{\tilde{T}}_3 \end{bmatrix} = -\begin{bmatrix} (a_{11}+4\kappa T_1^3)\tilde{T}_1 \\ (a_{21}+4\kappa T_2^3)\tilde{T}_2 \\ (a_{31}+4\kappa T_3^3)\tilde{T}_3 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}\begin{bmatrix} \tilde{P}_1 \\ \tilde{P}_2 \\ \tilde{P}_3 \end{bmatrix} \tag{12}$$

In the previous section we obtained a local linear model at 500 ºC as

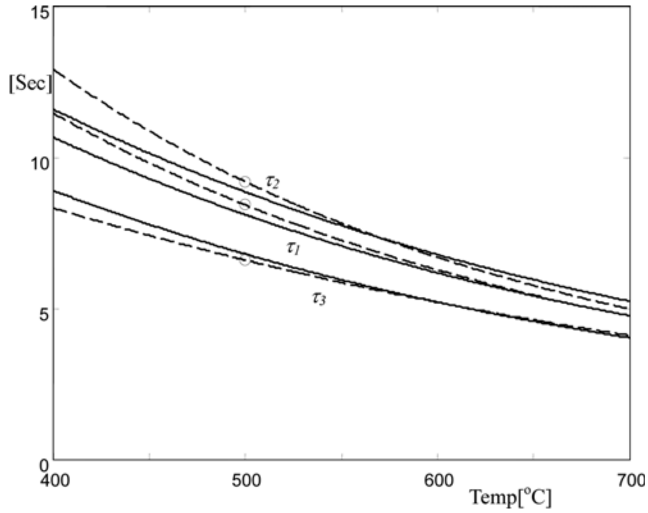

Fig. 6. Time constant estimates (Solid: exact, Dashed: estimate).

$$\begin{bmatrix} \tilde{T}_1(s) \\ \tilde{T}_2(s) \\ \tilde{T}_3(s) \end{bmatrix} = \begin{bmatrix} \frac{1}{8.426s+1} & 0 & 0 \\ 0 & \frac{1}{9.194s+1} & 0 \\ 0 & 0 & \frac{1}{6.598s+1} \end{bmatrix}$$

$$\begin{bmatrix} 0.4255 & 0.3066 & 0.0541 \\ 0.3588 & 0.3925 & 0.0827 \\ 0.3150 & 0.3244 & 0.1202 \end{bmatrix} \begin{bmatrix} \tilde{P}_1(s) \\ \tilde{P}_2(s) \\ \tilde{P}_3(s) \end{bmatrix} \tag{13}$$

Hence we can obtain model parameters by matching Eq. (12) and Eq. (13). They are

$$f(T) = \begin{bmatrix} 3.6170\times10^{-2}T_1 + 5.00\times10^{-11}T_1^4 \\ 3.4455\times10^{-2}T_2 + 5.00\times10^{-11}T_2^4 \\ 4.5144\times10^{-2}T_3 + 5.00\times10^{-11}T_3^4 \end{bmatrix}$$

$$B = \begin{bmatrix} 5.0500\times10^{-2} & 3.6388\times10^{-2} & 0.6421\times10^{-2} \\ 3.9025\times10^{-2} & 4.2690\times10^{-2} & 0.8995\times10^{-2} \\ 4.7744\times10^{-2} & 4.9169\times10^{-2} & 1.8218\times10^{-2} \end{bmatrix} \tag{14}$$

When a model is available, model-based control systems can be designed. A simple analytic control system designed by the external linearization method is tested here:

$$P(t) = \overline{P} + B^{-1}(f(T) - \alpha T + V(t)) \tag{15}$$

where $\overline{P}$ is the initial value of P(t). Then the system becomes linear and the transfer functions between $\tilde{T}(s)$ and $\tilde{V}(s)$ are all $1/(s+\alpha)$.

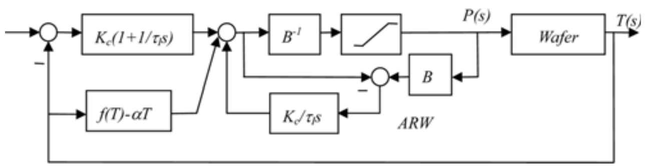

Fig. 7. Decoupling control system with an antireset windup feature.

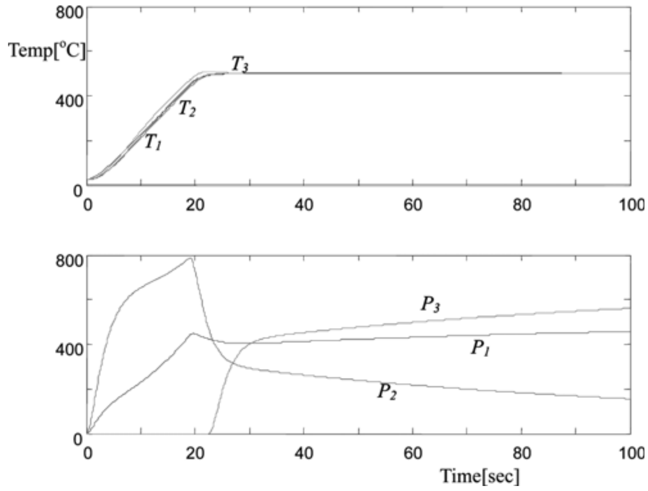

Fig. 8. Control system responses.

The PI controller can be used for this linearized system. Fig. 7 shows this linearizing control system, and Fig. 8 shows its closed-loop responses. We can see that the proposed control system is comparable with previous, more complicated control systems.

CONCLUSION

For the purpose of controller design, a very simple nonlinear model which consists of diagonal nonlinear first-order dynamics and multivariable static gains can be used for RTP. For this simple nonlinear model identification, it is very important to estimate the multivariable static gain matrix accurately. However, accurate estimation of multivariable static gains is difficult because of drift and the time-varying nonlinear dynamics. To reduce the effects of drift, test inputs are selected to be short pulses. From these pulse responses, linearized models are obtained. They are further improved by an iterative identification method.

After the term due to radiation on the wafer is fixed, a nonlinear model is constructed from the linearized models identified at a given temperature level. Simple nonlinear control law can be designed by using the nonlinear model.

ACKNOWLEDGMENT

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2006-D00099-100554).

REFERENCES

1. T. F. Edgar, S. W. Butler, W. J. Campbell, C. Pfeiffer, C. Bode and S. B. Hwang, *Automatica*, **36**, 1567 (2000).
2. J. Y. Choi and H. M. Do, *IEEE Trans. Semiconductor Manufacturing*, **14**, 1 (2001).
3. M. Cho, Y. Lee, S. Joo and K. S. Lee, *IEEE Trans. Semiconductor Manufacturing*, **18**, 430 (2005).
4. J. Y. Choi, H. M. Do and H. S. Choi, *IEEE Trans. Semiconductor Manufacturing*, **16**, 621 (2003).

5. K. S. Balakrishnan and T. F. Edgar, *Thin Solid Films*, **365**, 322 (2000).
6. E. Dasssu, B. Grosman and D. R. Lewin, *Comp. Chem. Eng.*, **30**, 686 (2006)
7. C. J. Huang, C. C. Yu and S. H. Shen, *Automatica*, **36**, 705 (2000).
8. C. D. Schaper, Y. M. Cho and T. Kailath, *Appl. Phys. A*, **54**, 317 (1992).
9. A. Kersch and T. Schafbauer, *Thin Solid Films*, **365**, 307 (2000).
10. S. J. Kim and Y. M. Cho, *Control Engineering Practice*, **10**, 1199 (2003).
11. W. Cho, T. F. Edgar and J. Lee, *Ind. Eng. Chem. Research*, **47**, 4791 (2008).
12. J. Lee, W. Cho and T. F. Edgar, *I&EC Research*, **37**, 1018 (1998).
13. W. Cho, *Temperature control and modeling of the rapid thermal processing chamber*, PhD Dissertation, University of Texas at Austin (2005).
14. K. S. Lee, J. Lee, I. Chin, J. Choi and J. H. Lee, *Ind. Eng. Chem. Research*, **40**, 1661 (2001).
15. S. A. Campbell, K. H. Ahn, K. L. Knutson, B. Y. H. Liu and J. D. Leighton, *IEEE Trans. Semiconductor Manufacturing*, **4**, 14 (1991).
16. W. Cho, T. F. Edgar and J. Lee, *Korean J. Chem. Eng.*, **23**, 171 (2006).