Technical Research Report

The Set-Valued Run-to-Run Controller in Semiconductor Manufacturing Processes

by Chang Zhang

T.R. 99-64

The Set-Valued Run-to-Run Controller in Semiconductor Manufacturing Processes

Chang Zhang / Advisor: John S. Baras*

August 13, 1999

Abstract

In semiconductor manufacturing, run-to-run (RtR) control is paid more and more attention. In this paper a set-valued RtR control scheme is introduced. Different from conventional RtR control methods, the set-valued method first calculates the feasible parameter set at the beginning of each run, then estimates the model parameters within this set. Compared to other RtR control schemes, it does not assume any statistical property of the noises. In simulation it was shown that it is robust to model and sensor errors, and it has the potential to be applied to highly nonlinear processes. Furthermore the set-valued method can be applied to other fields such as signal processing and chemical processes.

1 Introduction

In semiconductor manufacturing, run-to-run (RtR) control is paid more and more attention. The processes that could be applied the RtR control method are photolithography, chemical mechanical planarization (CMP), light pressure chemical vapor deposition (LPCVD), sputter deposition, ion implantation, photoresist and epi etch processes, etc.

A good RtR controller should be able to compensate for various disturbances, such as process drifts, process shifts due to maintenance or other factors, model or sensor errors, etc. Moreover, it should be able to deal with limitations, bounds, cost requirement, multiple targets, time delays that are often encountered in real processes.

*This work was supported by the Center for Satellite and Hybrid Communication Networks, under NASA cooperative agreement NCC3-528

At present there are the following RtR control methods available:

1. EWMA (Exponential Weighted Moving Average) method. The most common method in RtR control is the EWMA method [6],[15]. The EWMA method is applicable to some processes that can be approximated by linear models in a small range. Sometimes multiple linear models are used and a supervisory module is used to switch the models. For a detailed theoretical analysis of the EWMA method, please refer to [9]. There are lots of modifications to the EWMA method such as double exponential forecasting filter method [13]. It uses a predictor corrector controller (PCC) to eliminate the impact of machine and process drift. ANN EWMA [10] method makes the EWMA method applicable to some higher order linear-in-parameter processes. It will be discussed later. J. A. Mullins [26] uses a discrete process model for a linear process that has the ability of model-predictive control.

2. KIRC. It is the abbreviation of knowledge based interactive controller [29]. It is a machine learning algorithm for run-to-run control that uses leaves in a classification decision tree to suggest control actions. The algorithm generates a decision tree by using an information space with attribute tests. The starting operating point is chosen from the largest leaf in the decision tree where all outputs are inside the target range. It was shown that it can only be applicable to linear processes in a comparative simulation[29].

3. Least square recursive (LSR) method. It uses least square error to recursively approximate the process model. Typical examples are the optimized adaptive quality control (OAQC) [5] method and the Kalman Filter approach [18]. They are limited to polynomial processes that are not higher than 2nd order.

4. A probabilistic approach is proposed to RtR control too [11]. It uses probability theory to analyze the process. The concept of probability of stability and confidence of reliability are defined. Analytic formulas for the probability of stability were given in the particular case of an EWMA controller. However there is a key question of reliability of this methodology. Furthermore it assumes that the noises are Gaussian to derive the formulas, which limits the practical meaning of this method. At present it is limited to first order processes, though it has the promise to be used in higher order processes.

5.Artificial neural network method. It has the promise to be applied to severe nonlinear processes in principle. It was shown that ANN has great potential in modeling severe nonlinear semiconductor processes [34], [35], [36]. However a drawback to ANN method is that it does not supply an explicit model for the process. Thus it causes difficulties when one tries to apply optimal control method to adjust recipes. Wang[19] used Taylor expansion to find a first-order linear model to describe the ANN model. But it finally became a linear RtR control method. Its performance is only comparable to the EWMA method in the linear case. It may not be applied to high order linear-in-parameter processes. T. H. Smith [10] etc used ANN EWMA method to control a 2nd order process. It is successful for small disturbances or parameter variations in a limited run numbers(only 40). For large model error or disturbance, the process will become unstable. The reason is that it used an EWMA module to feed into the ANN model, which limit the ability of ANN to approximate a nonlinear process. D. Dong and Zafiriou used ANN approach to control batch-to-batch processes in chemical engineering, which may include reactors, crystallizers distillation towers and bio-chemical processes, etc, that can be modeled by a first principle model [27]. The processes are nonlinear chemical processes, but not the semiconductor process, which are usually much more complex and delicate. Therefore so far to the best of the author's knowledge, there is not a successful scheme to apply the ANN approach to the RtR control of semiconductor manufacturing nonlinear processes.

6. The set-valued approach. The set-valued based RtR controller can be used on both linear and nonlinear models. The main difficulty of the set-valued based RtR controller is the excessive computational time required to calculate the feasible sets and solve the optimization problem within this set.

Generally the feasible parameter set of linear or linear-in-parameter process can be estimated by set membership approach in the following ways [30]: 1. Ellipsoidal approach. It will be discussed in detail in the next part. 2. Orthotopic bounding. The feasible parameter set S is bounded with an orthotope aligned with the co-ordinate axes. S is defined by a set of 2N (N is the output parameter vector dimension) linear inequalities. Each bound can be obtained by solving a linear programming problem (e.g. simplex method). 3. Exact bounding. Some approaches are applicable to obtain the exact description of the set S in some specific situations [31], [32].

It is natural to use ellipsoids to approximate the region of indeterminacy. Because it has the following advantages: An ellipsoid is characterized by a vector center and a matrix; for convex regions ellipsoids can be used to obtain a satisfactory approximation; linear transformations map ellipsoids into ellipsoids.

In application of the ellipsoid algorithms, the minimum bounding ellipsoid which bounds the parameter set is desired. According to the difference of the search for the minimum bounding ellipsoid, there are mainly two different kinds of ellipsoid algorithms: Optimal Volume Ellipsoid Algorithm (OVE algorithm)[1] and Optimal Bounded Ellipsoid (OBE algorithm)[4]. The OVE algorithm was developed by M. F. Cheung, etc. It is based on the Khachiyan's ellipsoid algorithm[2] developed for solving the linear programming problem.

The OBE algorithm was developed by Fogel and Huang as a set-membership parameter estimation algorithm[3]. It is important because of its relationship to the LSE techniques. A main OBE algorithm is the Dasgupta and Huang OBE algorithm (DHOBE algorithm). It differs from the previous OBE method by introducing a forgetting factor which tries to shrink the ellipsoid each time the model is updated.

In the next we first introduce the OVE algorithm and the DHOBE algorithm. In section 3, the application of the set-valued RtR control with OVE approximation on LPCVD process is discussed. Then it follows the application of the set-valued OVE algorithm on the photoresist process. The OVE algorithm and the DHOBE algorithm are compared in simulation and theory in section 5. In section 6 conclusions are given.

2 The OVE algorithm and the DHOBE algorithm

For a linear-in-parameter system, it can be rewritten as the following form:

$$y_k = X_k^T \theta_k + \eta_k \tag{1}$$

where y_k is the output, X_k is the vector of inputs, θ_k is the vector to be estimated and η_k is the noise. The noise bound is assumed to be γ. It means that

$$|y_k - X^T \theta_k| \leq \gamma \tag{2}$$

The ellipsoid algorithm then produces, at each time step k, a set of estimates bounded by the ellipsoid:

$$E_k = \{\theta_k : (\theta_k - \hat{\theta}_k)^T P_k^{-1} (\theta_k - \hat{\theta}_k) \leq \sigma_k^2\} \tag{3}$$

Where P_k and σ_k determine the volume of the ellipsoid, and $\hat{\theta}_k$ is the center of the ellipsoid at run k.

1. The OVE algorithm.

In the OVE algorithm, $\sigma_k^2 \equiv 1$.

step 1. Calculate the following parameters

$$\alpha = \frac{y_k + \gamma - X_k^T \hat{\theta}_{k-1}}{\sqrt{X_k^T P_{k-1} X_k}} \tag{4}$$

$$\beta = \frac{\gamma}{\sqrt{X_k^T P_{k-1} X_k}} \tag{5}$$

If $\alpha > 1$, then reset β to $\beta - (\alpha - 1)/2$ and $\alpha = 1$. If $2\beta - \alpha > 1$, then reset β to $(1 + \alpha)/2$.

step 2. Calculate 3 intermediate variables.

If $\alpha \neq \beta$, then find the real solution τ of

$$(n+1)\tau^2 + \{\frac{(1+\alpha)(\alpha - 2\beta + 1}{\beta - \alpha} + 2[n(\beta - \alpha) + 1]\}\tau + n\alpha(\alpha - 2\beta) + 1 = 0 \tag{6}$$

such that $\alpha - 2\beta < \tau < \alpha$. Here n is the dimension of the estimated vector $\hat{\theta}$.

$$\delta = \frac{(\tau+1)^2(\beta-\alpha) - \tau(1+\alpha)(2\beta-\alpha-1)}{\tau+\beta-\alpha} \tag{7}$$

$$\sigma = \frac{-\tau}{\beta-\alpha} \tag{8}$$

If $\alpha = \beta$, then $\tau = 0$, and

$$\delta = \frac{n}{n-1}(1-\beta^2) \tag{9}$$

$$\sigma = \frac{1-n\beta^2}{1-\beta^2} \tag{10}$$

step 3. Update the ellipsoid.

$$\hat{\theta}_k = \hat{\theta}_{k-1} + \frac{\tau P_{k-1} X_k}{\sqrt{X_k^T P_{k-1} X_k}} \tag{11}$$

$$P_k = \delta(P_{k-1} - \sigma \frac{P_{k-1} X_k X_k^T P_{k-1}}{X_k^T P_{k-1} X_k}) \tag{12}$$

The algorithm can be initiated with a sufficiently large E_0 containing the feasible parameter set.

2. The OBE algorithm.

step 1.

$$\delta_k = y_k - X_k^T \hat{\theta}_{k-1} \tag{13}$$

step 2. If

$$\sigma_{k-1}^2 + \delta_k^2 \leq \gamma^2 \tag{14}$$

then it is thought as a noise disturbance and there is no update of the ellipsoid, otherwise go to step 3.

step 3.

step 3a. Compute two intermediate scalar variables:

$$G_k = X_k^T P_{k-1} X_k \tag{15}$$

step 3b.

$$\beta_k = (\gamma^2 - \sigma_{k-1}^2)/\delta_k^2 \tag{16}$$

step 4. Compute an update factor λ_k. Here $1-\lambda_k$ is the forgetting factor. It is confined to the range $0 \leq \lambda_k < 1$.

$$\lambda_k = min(\lambda_{max}, \upsilon_k) \tag{17}$$

where

$$\upsilon_k = \begin{cases} \lambda_{max} & \text{if } \delta_k^2 = 0 \\ (1-\beta_k)/2 & \text{if } G_k = 1 \\ (1-\sqrt{G_k/(1+\beta_k(G_k-1))})/(1-G_k) & \text{if } \beta_k(G_k-1)+1 > 0 \\ \lambda_{max} & \text{if } \beta_k(G_k-1)+1 \leq 0 \end{cases} \tag{18}$$

step 5. Update the parameter uncertainty factor

$$\sigma_k^2 = (1-\lambda_k)\sigma_{k-1}^2 + \lambda_k\gamma^2 - \lambda_k(1-\lambda_k)\delta_k^2/(1-\lambda_k+\lambda_k G_k) \tag{19}$$

step 6.This is the rescue procedure. When $\sigma_k^2 > 0$, proceed to step 7, otherwise:

Compute

$$\kappa = \begin{cases} \delta_k^2 + \gamma^2 - 2\gamma|\delta| & \text{if } \lambda_k \neq \lambda_{max} \\ \lambda_{max}[\frac{\delta_k^2}{1-\lambda_{max}+\lambda_{max}G_k} - \frac{\gamma^2}{1-\lambda_{max}}] & \text{if } \lambda_k = \lambda_{max} \end{cases} \tag{20}$$

Reset the uncertainty parameter for time k-1

$$\sigma_{k-1}^2 = \kappa + \zeta \tag{21}$$

then return to step 3b.

step 7. Update the ellipsoid parameters.

$$P_k^{-1} = (1-\lambda_k)P_{k-1}^{-1} + \lambda_k X_k X_k^T \tag{22}$$

$$P_k = \frac{1}{1-\lambda_k}[P_{k-1} - \frac{\lambda_k P_{k-1} X_k X_k^T P_{k-1}}{1-\lambda_k+\lambda_k G_k}] \tag{23}$$

$$\hat{\theta}_k = \hat{\theta}_{k-1} + \lambda_k P_k X_k \delta_k \tag{24}$$

The initial conditions of the OBE algorithm can be set as $\sigma_0^2 = 100$,$P_0^{-1} = I$ to include the optimal estimation point.

For more details about the OVE algorithm and DHOBE algorithm, please refer to [1], [3] and [4].

3 Application of the set-valued RtR controller with OVE approximation on LPCVD process

The idea of the controller here is to use the OVE algorithm to approximate the set of the feasible states. The set is updated each time when new measurements are available. Once the set is updated, the process model is updated within this region. Usually the center of the ellipsoid is an estimate of the process model. However the OVE algorithm may not propagate the center of the ellipsoid as the estimate. For example, the estimate may be based on worst case[7].

In the linear case, the set-valued RtR controller with OVE approximation works very well. Its performance is comparable to the EWMA method when the disturbance is a drift. Under a shift disturbance, its performance is better than the EWMA method. We can validate this performance from simulation results. The model we used is the following light pressure chemical vapor deposition (LPCVD) process. Here equation (1) can be simplified to a linear process by taking "logarithm" operation. But equation (2) is a nonlinear model. In this process we control two target rates R_1 and R_2. They are the deposition rates in $\overset{o}{A}/min$ on the first and last wafer respectively. The target rates are fixed at 169.75 $\overset{o}{A}/min$ and 141.7 $\overset{o}{A}/min$ respectively.

$$R_1 = exp(c_1 + c_2 lnP + c_3 T^{-1} + c_4 Q^{-1}) \tag{25}$$

$$R_2 = R_1 \frac{1 - S' C_{gs} R_1 Q^{-1}}{1 + S' C_{gs} R_1 Q^{-1}} \tag{26}$$

T stands for the temperature in K, P the pressure in mtorr, Q the silane flow rate in sccm. They are the inputs (recipes) to the process. We adjust them to maintain the process outputs on targets. The process parameters are $c_1 = 20.65$, $c_2 = 0.29$, $c_3 = -15189.21$, $c_4 = -47.97$, $S' = 4777.8$, $C_{gs} = 1.85 \times 10^{-5}$, where the units are omitted for convenience[7].

The simulation results are shown in Figure 1, Figure 2 and Figure 3. The target, the 3σ upper bound and lower bound are shown in these figures by three straight lines.

In Figure 1, there is large model error. However it can be seen that the controller returns the process output to the target value immediately.

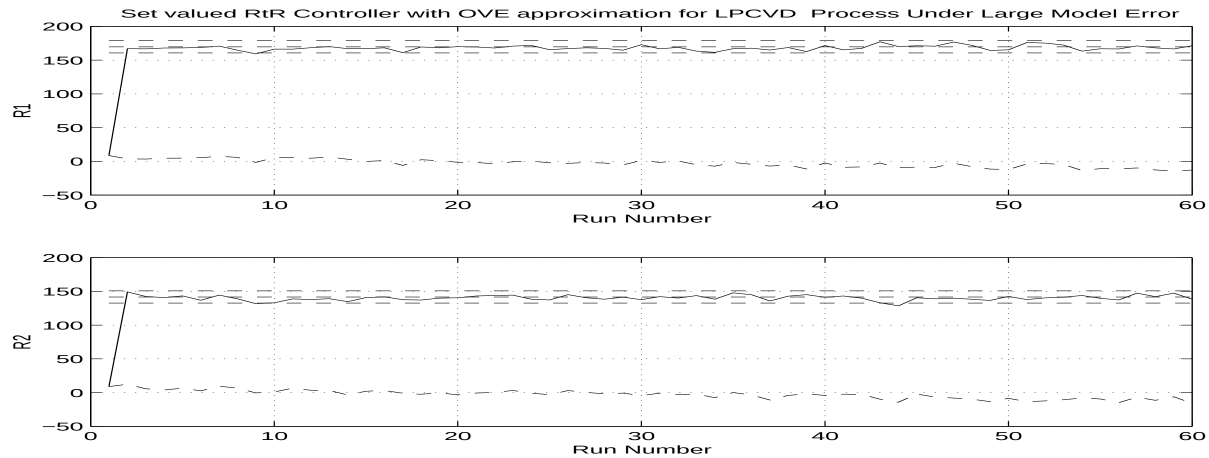

Figure 1: Set-valued RtR controller with OVE approximation for LPCVD process with large model error

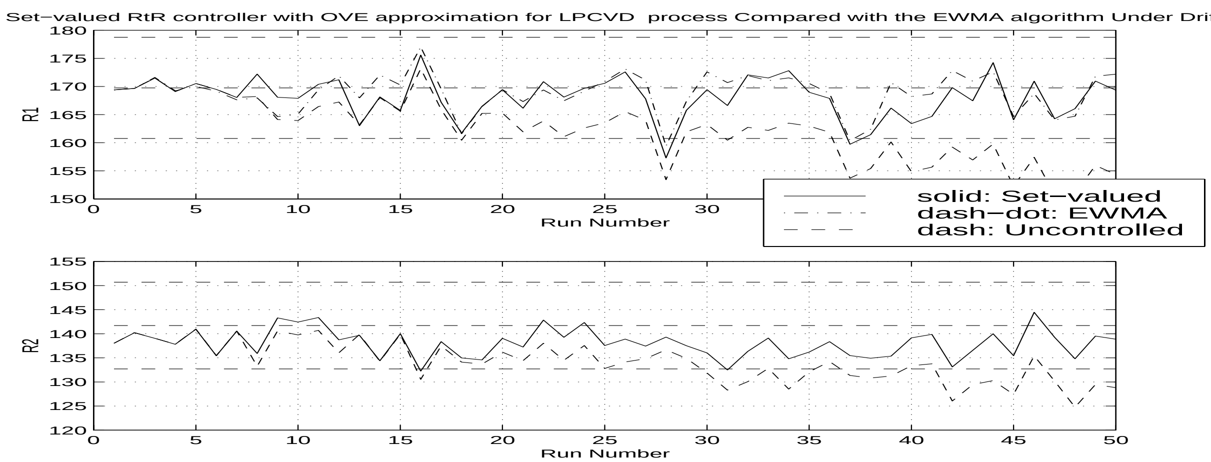

Figure 2: Set-valued RtR controller with OVE approximation for LPCVD process under drift

In Figure 2 and Figure 3, the set-valued RtR controller is compared with the EWMA controller under the disturbance of drift and shift separately. The weight parameter for the EWMA algorithm is 0.35. It was obtained by selecting the optimal performance among multiple weight parameters. The EWMA method is used to control only the single target R_1. The set-valued RtR controller is used to control the two targets R_1 and R_2. From Figure 2, it can be seen that both methods control R_1 well. Figure 3, it shows that the set-valued RtR controller returns the output to the target immediately after detecting the disturbance; The EWMA controller needs more steps to return the output within the 3σ bound.

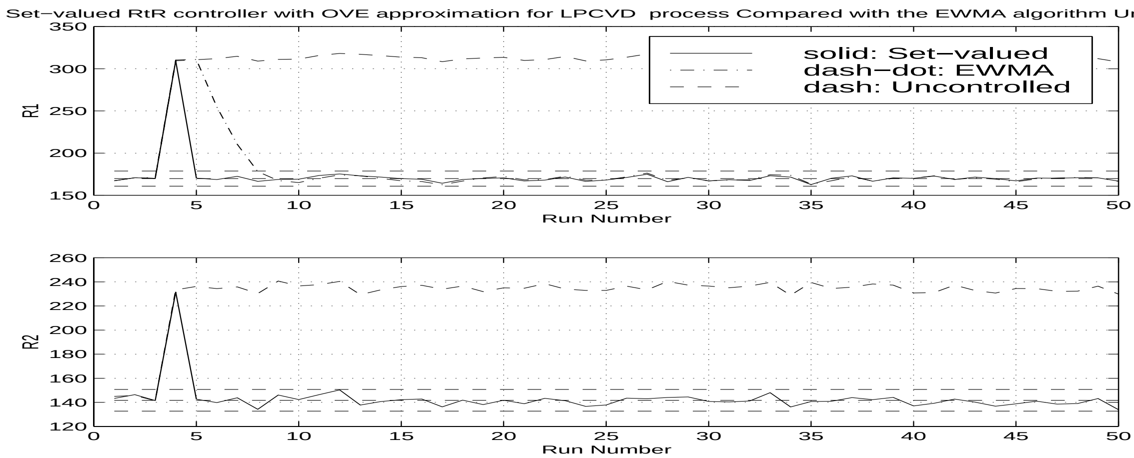

Figure 3: Set-valued RtR controller with OVE approximation for LPCVD process under shift

4 Application of the set-valued RtR controller with OVE approximation on photoresist process

In semiconductor manufacturing, many processes can be subjected to small shifts or drift changes. These perturbations can be compensated using the EWMA method. But, unfortunately, this is not always the case[5]. For example, many plasma processes have been shown to exhibit small to large nonlinearity in behavior. Further, the photoresist process and the chemical mechanical planarization (CMP) process require dynamic process models too. Thus it is necessary to develop non-linear algorithms to solve this type of problems.

The EWMA method updates only the constant term of the linear equation. Because many semiconductor processes are slow processes, suffering from noises such drifts, updating the constant term is usually enough to track the process. Since the ellipsoid algorithm originates from a linear model with quadratic bounds, the idea here is to only update the constant and the linear part of the nonlinear models for the photoresist process[17].

The following is the model used in the photoresist process. After changing process variables, it can be simplified to a second order nonlinear process.

$$T = -13814 + \frac{2.54 \cdot 10^6}{\sqrt{SPS}} + \frac{1.95 \cdot 10^7}{BTE\sqrt{SPS}} - 3.78BTI - 0.28SPT - \frac{6.16 \cdot 10^7}{SPS} \qquad (27)$$

Where T is the resist thickness in Angstroms, SPS the spin speed in RPM, SPT the spin time in seconds, BTI the baking time in seconds, and BTE the baking temperature in degrees Celsius. They are the inputs (recipes) to the process. They are confined to the following bounds:

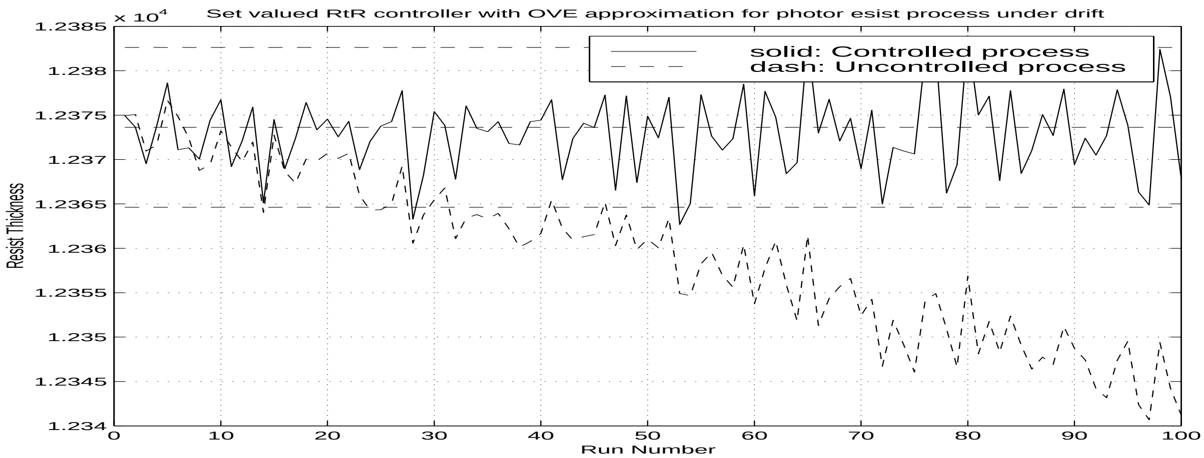

Figure 4: Set-valued RtR controller with OVE approximation for photoresist process under drift

$$
\begin{aligned}
4500 < SPS < 4700 \\
15 < SPT < 90 \\
105 < BTE < 135 \\
20 < BTI < 100
\end{aligned}
$$

The simplified model is shown in the following equation:

$$
T = -13814 + 2.54 \cdot 10^6 u_1 + 1.95 \cdot 10^7 u_1 \cdot u_2 - 3.78 \cdot u_3 - 0.28 \cdot u_4 - 6.16 \cdot 10^7 {u_1}^2 \tag{28}
$$

where:

$$
\begin{aligned}
u_1 &= \frac{1}{\sqrt{SPS}} \\
u_2 &= \frac{1}{BTE} \\
u_3 &= BTI \\
u_4 &= SPT
\end{aligned}
$$

The simulation results for the set-valued RtR controller with OVE approximation to control the resist thickness are shown in Figure 4 and Figure 5. It shows that the set-valued RtR controller can control the second order process very well under the disturbance of drift or shift.

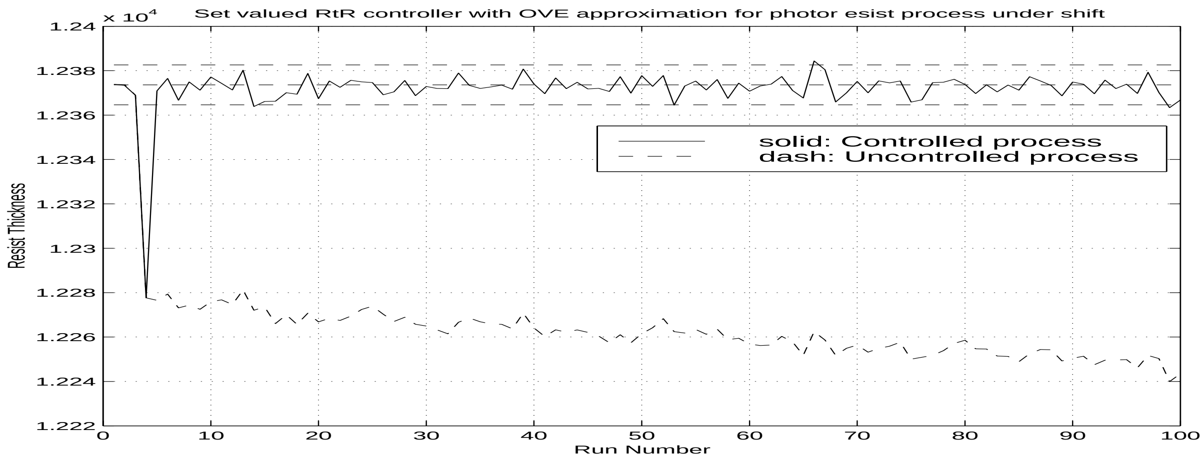

Figure 5: Set-valued RtR controller with OVE approximation for photoresist process under shift

5 Comparison of the set-valued RtR controller with OVE approximation and the set-valued RtR controller with OBE approximation

The derivation of the OVE algorithm is based on a geometrical point of view, but the OBE algorithm uses a RLS type scheme to update the center of the ellipsoid. The distinction between the OVE and the OBE algorithm lies in that the new ellipsoid from the updating is optimal under no constraints for the OVE, and the center of the ellipsoid is a “modified” recursive least square estimate (LSE) for the OBE[1]. Furthermore, the OVE algorithm always update the ellipsoid even when it is unnecessary, but the OBE algorithm only update the ellipsoid when it is necessary. However a smaller bound in the OBE algorithm usually leads to the update of the ellipsoid in each run.

The comparison was made by simulation for a different target in the photoresist process. The process is a second-order process with higher non-linearity. After some simplification, it has the following form:

$$\begin{aligned} R = \quad & 134.4 - 0.046u_1 + 0.32u_2 - 0.17u_3 + 0.023u_4 - 4.34 \times 10^{-5}u_1u_2 \\ + \quad & 5.19 \times 10^{-5}u_1u_3 - 1.07 \times 10^{-3}u_2u_3 \\ + \quad & 5.15 \times 10^{-6}u_1^2 - 4.11 \times 10^{-4}u_2u_4 \end{aligned} \tag{29}$$

Where R is the reflectance in % and the other variables are defined as in previous section.

More simulations have been made for different processes and it was found that both algorithms work well in the linear case or low order nonlinear case. However in the high order nonlinear case, the OVE algorithm performs better than the DHOBE algorithm. From the Figure 6 and Figure 7 it can be seen that the DHOBE algorithm has some overshooting.

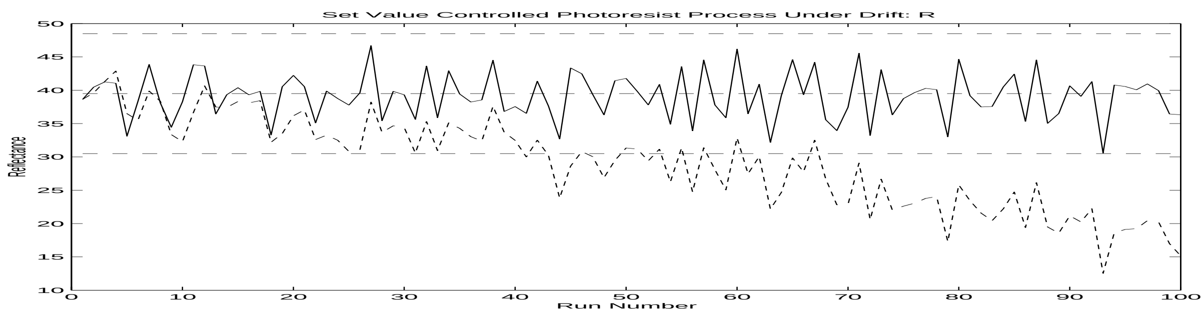

Figure 6: Set-valued RtR controller with OVE approximation for photoresist process II under drift

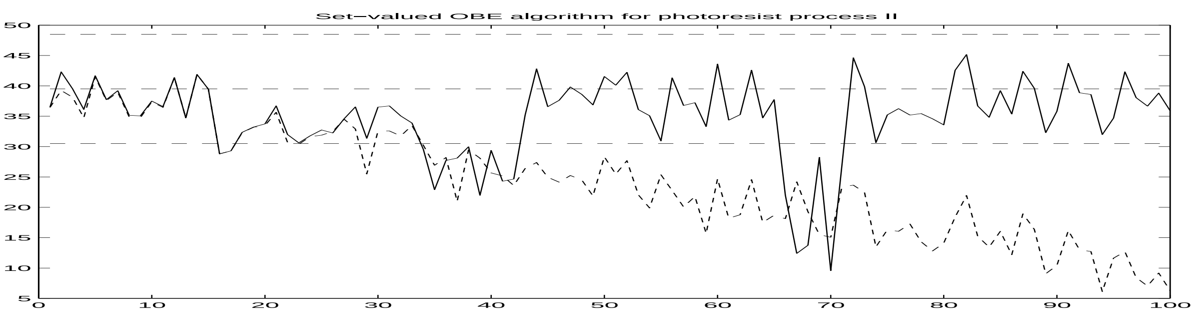

Figure 7: Set-valued RtR controller with OBE approximation for photoresist process II under drift

Actually it was found that even in the linear case, the over shooting may appear when the noises exist.

6 Summary

The set-valued RtR control scheme with ellipsoid approximation is applicable to various semiconductor manufacturing processes. The simulation results show that it is robust to model and sensor errors. The advantage of the set-valued method lies in the fact that it does not assume any statistical distribution for the noise. Knowing the bound of the noise is sufficient for applying this method. By using the ellipsoid approximation, the set-valued method is made easily applicable in practice.

The ellipsoid algorithm usually requires an estimate of the error bound, which makes it difficult to realize in some sense. It was shown that an OBE estimator is not theoretically consistent with conservative bounds[12]. Practical application of the OBE identification is made possible by some new algorithms such as automatic estimation of the model error bounds. For example, the OBE algorithm with automatic bound estimation (OBE-ABE) was developed by T. M. Lin, etc[12]. It is the first OBE algorithm to solve the difficult problem of blindly estimating error bounds. It may remove the major obstacle preventing the application of ellipsoid algorithms. However in simulation with LPCVD process it was shown that the OBE-ABE algorithm does not work well. It may only be applied to time-invariant processes or slow changing processes. In general the OVE algorithm works better than the OBE algorithm because of the existence of overshooting in the OBE algorithm.

Further research needs to be done on the generalization of the set-valued methods to the larger set of semiconductor processes. For those processes with dynamic nature that can not be expressed in a polynomial form, a much more general scheme is still needed. The immediate goal is to develop a general set-valued RtR controller for semiconductor manufacturing processes. A promising way is to introduce intelligent control and neural network into the framework of the set-valued method. It may lead to an exciting new field in the future.

References

[1] M. F. Cheung, S. Yurkovich and K. M. Passino, “An Optimal Volume Ellipsoid Algorithm for Parameter Set Estimation”, Proceedings of the 30th Conference on Decision and Control, 1991.

[2] B. Aspvall and R. E. Stone, “Khachiyan's linear programming algorithm”, Journal of Algorithms, vol. 1, pp. 1-13, 1980.

[3] E. Fogel and Y. F. Huang, “On the value of information in system identification-bounded noise case”,Automatica, vol. 18, no. 2, pp. 229-238, Mar. 1982.

[4] S. Dasgupta and Y.F.Huang,“Asymptotically Convergent Modified Recursive Least-Squares with Data-dependent Updating and Forgetting Factor for Systems with Bounded Noise”, IEEE Transactions on Information Theory, vol IT-33, No. 3, p383-392, May 1987.

[5] E. D. Castillo and J. Y. Yeh, “An Adaptive Run-to-Run Optimizing Controller for Linear and Nonlinear Semiconductor Processes”, IEEE Transactions on Semiconductor Manufacturing, Vol. 11, No. 2,pp. 285-295, May 1998.

[6] E. Sachs, A. Hu, and A. Ingolfsson, ”Run by run process control: combining SPC and feedback control”, IEEE Trans. Semiconduct. Manufact., vol. 8, pp. 26-43, Feb. 1995.

[7] J. S. Baras, N. S. Patel, “Designing Response Surface Mdoel-Based Run-by-Run Controllers: A Worst Case Approach”, IEEE Transactions on Components, Packaging and Manufacturing Technology, vol. 19, April 1996.

[8] J. S. Baras, N. S. Patel, ”A framework for robust run by run control with lot delayed measurements,” IEEE Trans. Semiconduct. Manufact., vol. 10, no. 1, pp. 75-83, Feb. 1997.

[9] A. Ingolfsson and E. Sachs, "Stability and sensitivity of an EWMA controller", Journal of Quality Technology, vol. 25, pp. 271-287, 1993.

[10] T. H. Smith and D. S. Boning, "Artificial neural network exponentially weighted moving average controller for semiconductor processes," J. Vacuum Science Tech. A, vol. 15, no. 3, pp. 236-239, May 1997.

[11] E. S. Hamby, P. T. Kabamba, and P. Khargonekar, "A probabilistic approach to run-to-run control," IEEE Trans. Semiconduct. Manufact., vol. 11, no. 4, pp. 654-669, 1998.

[12] T. M. Lin, M. Nayeri, and J. R. Deller, "A practical development in optimal bounding ellipsoid processing", IEEE Signal Processing Letters, vol. 4, no. 8, pp. 236-239, August 1997.

[13] S. W. Butler and J. A. Stefani, "Supervisory run-to-run control of polysilicon gate etch using in situ ellipsometry," IEEE Trans. Semiconduct. Manufact., vol. 7, pp. 193-201, May 1994.

[14] M. Hankinson, T. Vincent, K. Irani, and P. Khargonekar, "Integrated real-time and run-to-run control of etch depth in reactive etching," IEEE Trans. Semiconduct. Manufact., vol. 10, pp. 121-130, Feb. 1997.

[15] D. Boning, W. Moyne, T. Smith, etc., "Run by run control of chemical-mechanical polishing," 1995 IEEE/CPMT Int'l Electronics Manufacturing Technology Symposium, pp. 81-87, 1995.

[16] E. Del Castillo and A. Hurwitz, "Run to run process control: Literature review and extensions," J. Qual. Technol., vol. 29, no. 2, pp. 184-196, 1997.

[17] S. Leang and C. J. Spanos, "Statistically Based Feedback Control of Photoresist Application", IEEE/SEMI Advanced Semiconductor Manufacturing Conference, pp.185-190, 1991

[18] E. Palmer, W. Ren, C. J. Spanos, "Control of photoresist properties: A Kalman filter based approach," IEEE Trans. Semiconduct. Manufact., vol. 9, no. 2, pp. 208-214, May 1996.

[19] X. A. Wang and R. L. Mahajan, "Artificial neural network model-based run-to-run process controller", IEEE Trans. Components, Packaging and Manu. Tech., Part C, vol. 19, no. 1, Jan. 1996.

[20] S. Dasgupta and Y.F.Huang, "Asymptotically convergent modified recursive least-squares with datadependent updating and forgetting factor for systems with bounded noise," IEEE Trans. Information Theory, vol. IT33, no. 3, pp. 383392, May 1987.

[21] T. M. Lin, M. Nayeri, and J. R. Deller, Jr., "Automatic bound estimation: A practical development in optimal bounding ellipsoid processing," IEEE Signal Processing Letters, vol. 4, no. 8, pp. 236-239, 1997.

[22] J. R. Deller, Jr., S. F. Odeh, "Adaptive set-membership identification in q(m) time for linear-in-parameters models," IEEE Trans. Signal Processing, vol. 41, no. 5, pp. 1906-1924, May 1993.

[23] A. Sabharwal and L. Potter, "Set estimation via ellipsoidal approximations," IEEE Trans. Signal Processing, vol. 45, no. 12, pp. 3107-3111, Dec. 1997.

[24] A. K. Rao, Y. Huang, "Tracking characteristics of an OBE parameter-estimation algorithm," IEEE Trans. Signal Processing, vol. 41, no. 3, pp. 1140-1148, March 1993.

[25] T. Chen and M. J. Wang, "A fuzzy set approach for yield learning modeling in wafer manufacturing", IEEE Trans. Semiconduct. Manufact., vol. 12, no. 2, pp. 252-258, May 1999.

[26] J. A. Mullins, etc, "An evaluation of model predictive control in run to run processing in semiconductor manufacturing", SPIE, vol. 3213, pp. 182-189, 1997.

[27] D. Dong, T. J. McAvoy, E. Zafiriou, "Batch-to-batch optimization using neural network models", Int. Eng. Chem. Res. Vol. 35, pp. 2269-2276, 1996.

[28] A. Wu, etc, "Parallel optimal statistical design method with response surface modeling using genetic algorithms", IEE Proc.-Circuits Devices, Syst., Vol. 145, no. 1, pp. 7-12, 1998.

[29] Z. Ning, etc, "A comparative analysis of run-to-run control algorithms in the semiconductor manufacturing industry", 1996 IEEE/SEMI Advanced Semiconductor Manufacturing Conference, pp. 375-381, 1996.

[30] E. Walter, etc, "Estimation of parameter bounds from bounded-error data: A survey", Mathematics and Computers in Simulation 32, pp. 449-468, 1990.

[31] V. Broman and M. J. Shensa, "Polytopes, a novel approach to tracking", Proc. 25 th IEEE Conference on Decision and Control, Athens, pp. 1749-1752, 1986.

[32] S. H. Mo and J. P. Norton, "Fast and robust algorithm to compute exact polytope parameter bounds", Math. Comput. Simulation 32, pp. 481-493, 1990.

[33] P. L. Combettes, "The foundations of set theoretic estimation", Proc. IEEE, vol. 81, no. 2, Feb. 1993.

[34] C. D. Himmel and G. S. May, "Advantages of plasma etch modeling using neural networks over statistical techniques", IEEE Trans. Semiconduct. Manufact., vol. 6, no. 2, May 1993.

[35] C. D. Natale, etc, "Modeling of APCVD-doped silicon dioxide deposition process by a modular neural network", IEEE Trans. Semiconduct. Manufact., vol. 12, no. 1, Feb. 1999.

[36] Y. L. Huang, etc, "Constructing a reliable neural network model for a plasma etching process using limited experimental data", IEEE Trans. Semiconduct. Manufact., vol. 7. No. 3, Aug. 1994.