MORE PROCESS SYSTEM ENGINEERING (PSE) APPLICATIONS IN IC MANUFACTURING

Daniel R. Lewin, Sivan Lachman-Shalem, Benyamin Grosman

*PSE Research Group, Wolfson Department of Chemical Engineering,
Technion–IIT, Haifa 32000, Israel*

Abstract: The manufacture of integrated circuits is driven by a demand for faster calculation capabilities and lower costs, which will require the development of a new generation of manufacturing tools to increase yield productivity, spearheaded by improved measurement devices and advanced process control. The objectives of this paper are review of the challenges in two main PSE areas: process monitoring and process control. PSE solutions appropriate for these challenges involve harnessing multivariate statistics, automated modeling approaches like genetic programming, and multivariable model-based control. The paper is illustrated with several example applications, all tested in fabrication facilities in Israel.

Keywords: Integrated circuit manufacturing; Process systems engineering; Model-based control; Process monitoring; Yield enhancement.

1. INTRODUCTION

Semiconductor devices, usually referred to as integrated circuits (ICs), are created by subjecting wafers of doped silicon to a cycle of processes designed to form layers of conducting, semi-conducting and resistive layers with a prespecified topology, to impart the desired electronic functions on the final device. The width of polysilicon lines that can be repeatedly attained, defining transistor gates in IC circuits, is the key performance factor, usually referred to as the *design rule*. Technology commonly in use today has the capability of generating topology with design rules of 130 nm, based on 200 mm diameter silicon wafers. As pointed out by the international technology roadmap for semiconductors (ITRS) 2003, the IC industry is driving towards the manufacture of more compact devices, based on 300 mm diameter wafers, and even smaller design rules, with 70 nm in ramp at several locations worldwide at this time.

The current practice in the IC industry is to implement open loop, recipe-driven, feedforward control strategies. Often, the desired operating point is determined following a statistical design-of-experiment (DOE) study to define the “stable” process window. Subsequently, the degrees-of-freedom of the process (the manipulated variables) are fixed according to the DOE results. Feedback control, if implemented at all, is usually limited to single loop PID control, and usually only for the lower level loops (e.g., temperature control). The main disadvantages arising from the widely accepted feedforward control strategy are obvious to anyone with modest control experience: (a) the approach cannot deal with unmeasured and/or unknown disturbances (which, of course, always occur in practice), and (b) since the feedforward correction is based of imperfect process knowledge, it will generally not allow product to be produced consistently on target (even in the absence of disturbances). Modeling activity is seldom seen, and if at all, usually limited to empirical, data-driven approaches. Typically, when loss-of-control (LOC) incidents are encountered, the process is stopped, with a resulting loss in yield, and a new DOE is initiated to assess the problem and suggest corrections. Surprisingly, process monitoring is usually carried out using univariate methods, despite the fact that hundreds of variables need to be monitored, often strongly correlated.

Modern PSE methodologies can significantly improve the management of IC manufacturing. These call for model-based approaches, where models are generated with accuracy at a level appropriate to the application. Multivariable feedback control is favored as a means to deal with unmodeled/unmeasured disturbances and to accommodate constraints. Process monitoring is carried out relying on multivariate methods, such as principal component analysis (PCA), coupled with first principles modeling to generate "smart" alarming. Ideally, the monitoring system should be integrated with the regulatory control system, to provide the framework of truly advanced process control (APC).

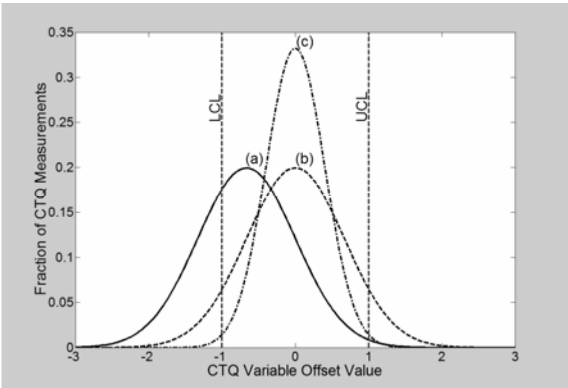

Fig. 1. Probability functions for CTQ variable: (a) solid line – without regulatory control; (b) dashed line – with regulatory control but subject to large variance; (c) dashed-doted line – with regulatory control with lower variance.

Figure 1 compares typical performances that can be expected depending on the control strategy implemented. The solid line (a) shows the expected distribution of the critical to quality (CTQ) variable when using only feedforward control, which cannot deal with the effect of unmeasured disturbances to the process, leading to a large fraction of the expected production under the LCL. The dashed line (b) indicates the improvement that can be expected by implementing a feedback control strategy designed to maintain the average CTQ measured on target. To significantly impact on product yield, however, it is necessary to apply not only feedback control, but also to reduce the CTQ variance as indicated by the dashed-dotted line (c) in Figure 1.

By the end of 2004, worldwide revenues for semiconductor device manufacturing are expected to exceed \$220 billion, a 27% increase from 2003 (according to the Gartner press release, 24th August 2004). Because of the huge market, there is room for significant improvements in profits, with each percent of yield improvement in semiconductor device manufacturing being worth \$2 billion on a global scale. Even on the scale of a modern manufacturing facility producing, say, 25,000 300-mm wafers per month, at a production cost of \$3,000 per wafer, a 1% increase in yield is worth \$30/wafer, or \$9M/year.

The objectives of this paper are to present the challenges and the state-of-the-art in two main PSE areas that have application in microelectronics manufacturing: process control and process monitoring, illustrated by example applications developed in cooperation between our research group and the IC industry in Israel.

2. PROCESS MONITORING

In the microelectronics industry, attempts are being made to enhance performance and yield via fault detection. Most of the attention is being directed at reducing process variation by various means: feedforward control for reducing run-to-run variation (Leang et al., 1996; Ruegsegger et al, 1999), as well as model predictive control (Edgar et al, 1999). Multivariate statistical methods have also been applied with varying degrees of success (e.g., Chen et al, 2000). Our contributions have involved the usage of model-based principal component analysis (MBPCA). In recent work (Lachman-Shalem et al, 2002a), a physical model describing an oxidation tube is used to simulate faulty oxidation ovens in the fabrication of a CMOS transistor, and combined with PCA to monitor simulated CMOS manufacturing. The results indicates that MBPCA is sensitive enough to identify abnormal operating conditions resulting from irregular correlations among monitored variables, even though the individual measured variables can be safely within their specific normal bounds.

An additional study on the feasibility of applying MBPCA to monitoring a complex IC manufacturing process was recently commissioned. In the process, each wafer in a lot is processed individually, and 15 transient variables are monitored for each processed wafer. Figure 2 shows typical transient profiles for each of the 15 monitored variables, for a lot of 17 wafers. As outlined in our publications on MBPCA (Rotem et al., 2000; Lachman-Shalem et al., 2002a), the steps required are:

a. Standardization and time-scale normalization of transient data for normal operating conditions (NOC). In this implementation, each wafer transient appears as 100 consecutive data points on the ordinate scale.
b. Generation of residual data by subtracting NOC model prediction from NOC data. In this implementation, average NOC transients are used as model predictions.
c. Principal component analysis is used to generate a NOC model. For this application, the first eight principal components, capturing 70% of the NOC variance, were used to construct a model that adequately limits the approximation error.

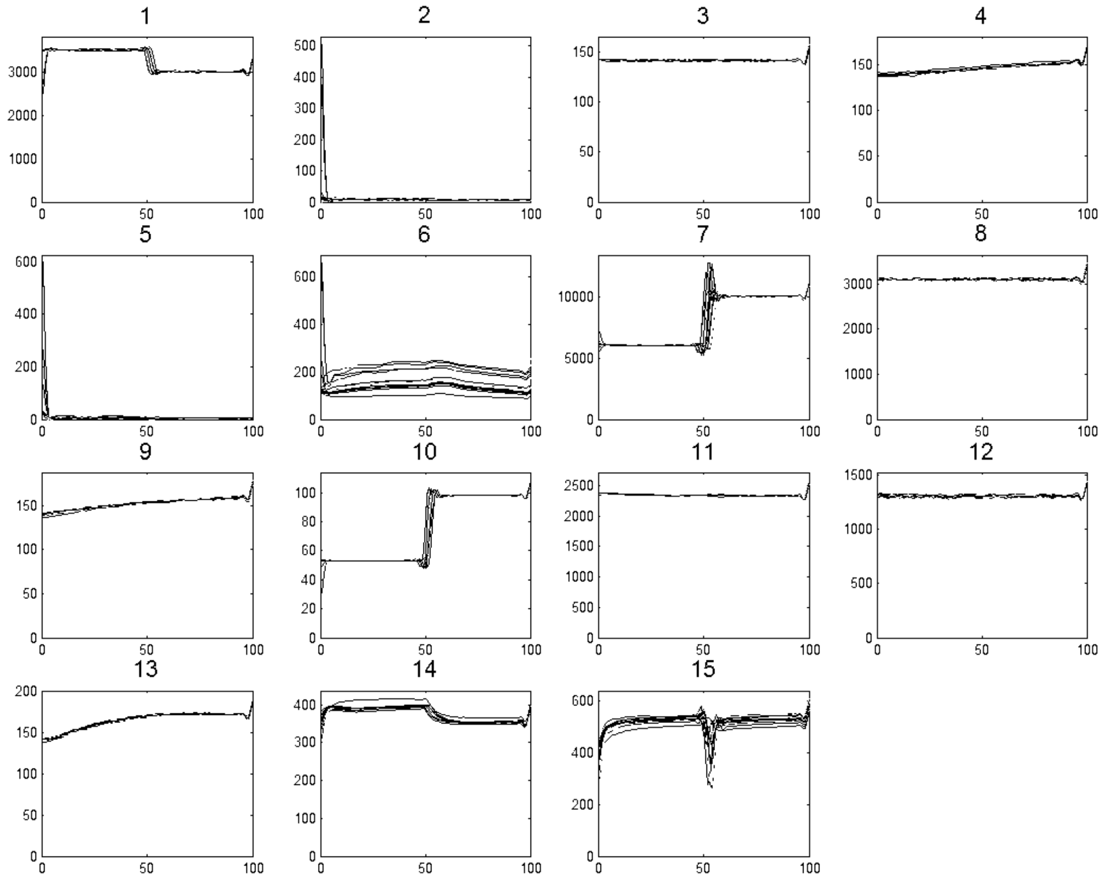

Fig. 2. Raw data for the 15 monitored variables, for all 17 wafers in the lot, after time-scale normalization.

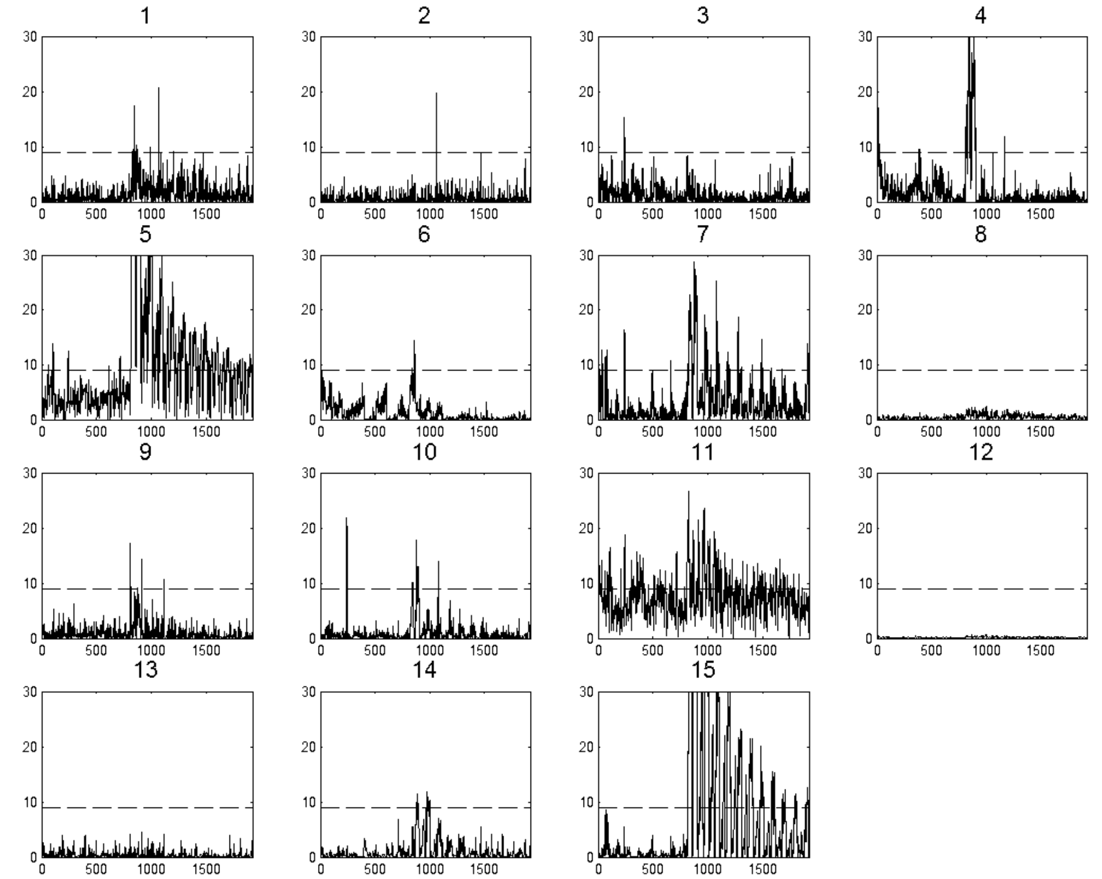

Fig. 3. MBPCA analysis of data in Fig. 2. Note the upper NOC limit at SPE = 9, indicating a root cause of the fault in wafer 8, identified in variable 4, propagating to variables 5, 7, 11, and 15.

d. The squared prediction error (SPE) is computed from the difference between the actual data points and the PCA model prediction. As the data has been standardized, an expected maximum SPE limit of 9 implies greater than 99.9% confidence. The SPE upper limit of 9 is plotted as a reference level in Figure 3.

In this way, an on-line monitoring of the measured variables can be tracked as the SPE as a function of sample time, as shown in the example in Figure 3. It is noted that the lot is normal until the eighth wafer (from sample 800 on the ordinate scale), when the SPE for measurement 4 registers a severe abnormality. This failure propagates to faults indicated in additional sensors (5, 7, 11 and 15), which continue to the end of the lot. This information is helpful in setting up preventative maintenance measures in the FAB. It is interesting to note that conventional univariate statistics is unable to identify failures of the kind demonstrated in these plots.

3. PROCESS CONTROL

The microelectronics industry today owes much of its success to remarkable advances in the lithographic process used to fabricate integrated circuits. The economic inducement of cheaper yet more sophisticated integrated circuits continues to motivate the industry to produce ever-smaller features. The miniaturization of the components of the integrated circuit has been achieved through improvements in projection printing, in the photoresists that are used to generate the structures, and better control of the photolithography cluster. The main production problem is the control of the precision of the printed line width (critical dimension, CD). The tolerance on the CD is delineated by upper and lower control limits, with the variability in CDs being one of the limitations to increased wafer yield.

Much work has been reported on the application of process control to lithography. Lithography is comprised of a number of basic, but interacting, operations. Despite this, to the best of our knowledge, all of the reported work on photolithography control relates to single-loop control, in which a key manipulated variable in the process track is selected for regulation of the specified CD. In theoretical work carried out in our group, a nonlinear model predictive control (NMPC) has been developed and tested on a full-track simulation using the commercial simulator PROLITH. The performance of the NMPC has been demonstrated to be significantly better than the best-possible single-loop controller (Lachman-Shalem et al, 2002b).

The success of the theoretical study provided the motivation to test the NMPC approach on a problematic layer at Tower Semiconductors Ltd. (Tower), an IC foundry in Migdal HaEmek in Israel. The methodology proposed in Lachman-Shalem et al (2002b) was followed, modified to account for FAB limitations, namely: (a) the manipulated variables used for the controller were limited to the stepper variables (dose and focus); (b) Both dense and isolated CDs were regulated; (c) The controller was implemented lot-by-lot, rather than wafer-to-wafer.

Realization of the NMPC approach to the Tower FAB involved the following steps (Grosman et al, 2005):

a. Calibration of the PROLITH model using focus exposure matrices (FEMs) supplied by Tower.

b. Generation of empirical nonlinear models relating the focus and dose to the CD outputs. This was carried out using our in-house genetic program (GP) – see Grosman and Lewin (2002 and 2004).

c. Development of NMPC using the nonlinear GP models to predict the process behavior. The controller uses filtered CD measurements, weights to favor the tracking of the more critical dense CD values, and to maintain the manipulated variables in mid-range, whenever possible. Figure 4 shows the simulated performance of the NMPC in rejecting a disturbance in the development time, imposed at sample 0 in the plot. Note that whereas this disturbance causes the uncontrolled dense CD to violate its lower acceptance limit bound, the NMPC succeeds in rejecting the disturbance almost immediately. The simulated noise levels correspond to those found in the FAB.

d. Implementation at Tower. The pre-tuned controller was implemented successfully with no further adjustments to any of the tuning parameters, giving the performance as shown in Figure 5. The same disturbance simulated in Figure 4 was imposed at wafer 2, causing the dense CD to violate its lower acceptance limit at wafer 6. At this point, the NMPC was activated, returning both dense and isolated CDs to their set points.

e. Note that in both the simulation and in the FAB test, the regulation of the CDs is attained at the expense of having both of the manipulated variables at their constrained limits (the dose at its minimum value and the focus at its maximum value). This is an indication of a process problem (indeed, the development time is excessive), and should be identified by the monitoring system. It is recommended that multivariate statistics be employed in setting up such a system, as discussed above.

The control strategy described above improves the yield by shifting the operating mode illustrated in Figure 1 from curve (a) to curve (b) using feedback control. Tower estimate that the implementation of the proposed control strategy on the recipe it was designed for is worth \$250,000/year in increased yields.

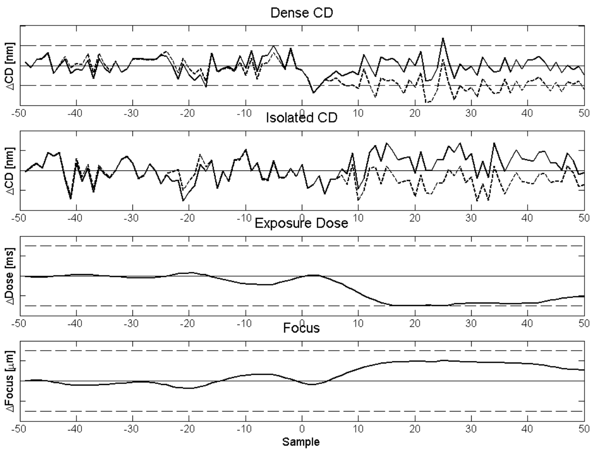

Fig. 4. Open-loop (dashed) and closed-loop (solid) response of CD dense and isolated lines to a simulated disturbance in development time.

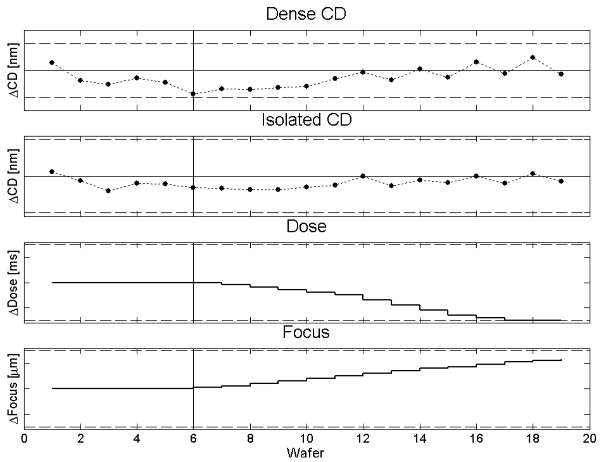

Fig. 5. True closed-loop response of CD dense and isolated lines to a disturbance in development time as implemented at Tower.

To make an even more significant impact, the variance in the CDs needs to be reduced. The variance in CDs is noticed in practice, since CDs are typically measured at several locations on a wafer (at Tower, five measurements are made). In the so-called *average mode* control described above, the feedback signal used for control is the average of these five measurements. To use the power of the stepper to reduce the variability, it is proposed to control CDs at each of the measured locations *independently*, by manipulating the local focus and dose settings. The main difficulty associated with this approach is the limited capability of existing stepper software and hardware to handle it.

The so-called field-to-field (F2F) control strategy involves:

a. Generating a PROLITH model relating local focus and dose settings with the CDs measured in each location. The five measurement points were selected to be close to the center, and east, north, west and south of center, at distinct radial locations.
b. Developing GP, reduced-order models relating the same CDs with the focus and dose settings. Each model provides the predictive capability required to control each of the five measured CDs using NMPC.
c. Tuning the NMPC involves the resolution of the following issues: (i) Defining the weights on tracking errors for dense and isolated CDs; (ii) Defining weights on moves using dose and focus; (iii) Defining the filter on CD measurements, since these are known to be noisy; (iv) Defining weights on offsets of the focus and dose settings from their mid-range value. This so-called "goalkeeper" feature was introduced to force the manipulated variables to mid-range values if there is no need for them to be elsewhere, to increase the capability of the controller to be resilient to unknown future disturbances.

The typical simulated performance of the F2F control strategy is presented in Figure 6. The plots show the response to a step disturbance in exposure time, imposed on the 11^{th} wafer and corrected in the 39^{th} wafer, with no simulated noise. In the first 11 wafers, the control succeeds in effectively reducing the F2F variance, as indicated by the sharp reduction in standard deviations (from 4 to 1 nm for isolated CDs and from 5 to 1 nm for sense CDs). The controller achieves this result by adjusting the five sets of focus and dose values (one corresponding to each location where CDs are measured) in a fan-like fashion. From the 11^{th} wafer, when the disturbance is applied, the manipulated variables are adjusted to try to maintain the CDs as close as possible to their setpoints. It is seen that this comes about at the cost of having to give up the minimization of the standard deviations of the CDs. However, from the 39^{th} wafer, where the disturbed exposure time is returned to its nominal value, the *goalkeeper* weights cause the controller to gradually readjust the manipulated variables to the minimum offsets from the mid-range values while satisfying the CD setpoints.

The simulations indicate that the *F2F mode* controller reduces the standard deviations of CDs to between 20-25% of the values obtained using *average mode* control. Returning to Figure 1, this in equivalent of going from curve (b) to curve (c), with the resulting expected product yields.

4. CONCLUSIONS

The drive for increasingly smaller critical dimensions in IC manufacture, together with increased yields is leading the industry to turn to APC solutions for the monitoring and regulation of production. This paper has focused on the possible impact of PSE tools in IC manufacturing. The applications demonstrated indicate that the implementation of model-based, multivariable control, backed up by a smart monitoring system using multivariate statistics, appears to have great promise in many applications.

From our experience, the entry of system engineers into the IC manufacturing business involves a steep learning curve. It is important to have a good working knowledge of the business, its vocabulary and technology, and this requires serious effort. Without this knowledge, communication with partners in the IC industry is difficult, but with it, and the necessary enthusiasm at the receiving end, the expertise and experience brought by the control systems engineer provide vehicles for significant contributions.

ACKNOWLEDGEMENTS

The photolithography control work described in this publication was conducted as part of the Wafer Fab Cluster Management (WFCM) Consortium supported by the "MAGNET" program of the Chief Scientist's Office at the Israeli Ministry of Industry and Trade. The authors also wish to thank Tower personnel for their enthusiastic support and collaboration, and especially Mrs. Raaya Swissa.

REFERENCES

Butler, S. W. and T. F. Edgar (1997). Case studies in equipment modeling and control in the microelectronics industry", *in Proc. of the 5th Int. Conf. on Chem. Proc. Control*, Eds. Kantor, J. C., Garcia, C. E. and Carnahan, B., *AIChE Symp. Series No.* 316, **93**, 133.

Chen, Q., R. J. Wynne, P. Goulding, and D. Sandoz (2000). The application of principal components analysis and kernel density estimation to enhance process monitoring, *Control Eng. Prac.* **8**, 531

Edgar, T. F., W. J. Campbell, and C. Bode (1999). Model-based control in microelectronics manufacturing, *Proc. of the IEEE 38th Conf. on Decision and Control*, **4**, 4185.

Gartner press release, 24th August 2004: http://www.gartner.com/5_about/press_releases/pr2004.jsp

Grosman, B. and D. R. Lewin (2002). Automated Nonlinear Model Predictive Control using Genetic Programming, *Comput. Chem. Eng*, **26**(4-5), 631.

Grosman, B. and D. R. Lewin (2004). Adaptive Genetic Programming for Steady-state Process Modeling, *Comput. Chem. Eng*, **28**(12), 2779.

Grosman, B. S. Lachman-Shalem, R. Swissa and D. R. Lewin (2005). Yield Enhancement in Photolithography through Model-based Process Control: Average Mode Control, *IEEE Trans. of Semiconductor Manuf.*, **18**(1) 86.

Lachman-Shalem, S., N. Haimovitch, E. N. Shauly, and D. R. Lewin (2002a). MBPCA Application for Fault Detection in NMOS Fabrication, *IEEE Trans. of Semiconductor Manuf.*, **15**(1), 60.

Lachman-Shalem, S., B. Grosman, and D. R. Lewin (2002b). Nonlinear modeling and multivariable control of photolithography, *IEEE Trans. of Semiconductor Manuf.*, **15**(3), 310.

Lai, J. H. and C. T. Lin (1999). Hybrid feedforward and feedback control of wafer temperature in RTP using genetic algorithm and fuzzy logic, *IEEE Trans. on Fuzzy Systems*, **7**(2), 160.

Leang, S. M. Shang-Yi, J. Thomson, B. J. Bombay, and C. J. Spanos. (1996). A Control System for Photolithographic Sequences, *IEEE Trans. of Semiconductor Manuf.*, **9**(2), 191.

Rotem, Y. A. Wachs and D. R. Lewin (2000). Ethylene Compressor Monitoring using Model-based PCA, *AIChE Jnl*., **46**(9), 1825.

Ruegsegger, S. A. Wagner, J. S. Freedenberg, and D. S. Grimard (1999). Feedforward control for reduced run-to-run variation in microelectronics manufacturing, *IEEE Trans. on Semiconductor Manuf.*, **12**(4), 493.

Scaper, C. D., T. Kailath, and Y. J. Lee (1999). Decentralized control of wafer temperature for multizone thermal processing systems, *IEEE Trans. on Semiconductor Manuf.*, **12**(2), 193.

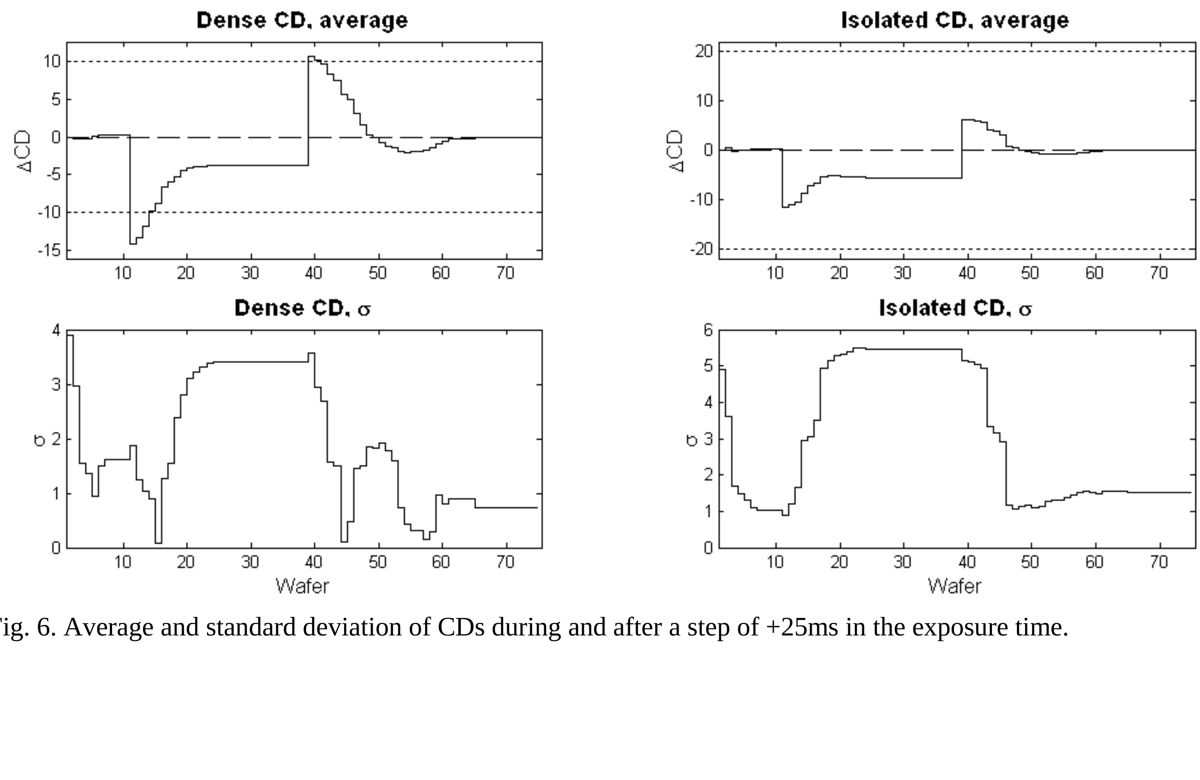

Fig. 6. Average and standard deviation of CDs during and after a step of +25ms in the exposure time.