Available online at www.sciencedirect.com
ScienceDirect
Computers & Industrial Engineering 50 (2006) 233–247
computers & industrial engineering
www.elsevier.com/locate/dsw

A dispatching rule for photolithography scheduling with an on-line rework strategy ☆

D.Y. Sha [a,b], S.Y. Hsu [c], Z.H. Che [d,*], C.H. Chen [e]

[a] *Department of Industrial Engineering and Management, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 30050, Taiwan, ROC*
[b] *Institute of Business Administration, Asia University, 500 Liufeng Road, Wufeng, Taichung 41354, Taiwan, ROC*
[c] *Department of Industrial Engineering and Management, Chien Kuo University of Technology, 1 Chieh-Sou N. Road, Changhua 50094, Taiwan, ROC*
[d] *Department of Industrial Engineering and Management, National Taipei University of Technology, 1, Sec. 3, Chung-Hsiao E. Rd., Taipei 106, Taiwan, ROC*
[e] *SiSoft Research Center, National Chiao Tung University, MIRC R316, 1001 Ta Hsueh Road, Hsinchu 30050, Taiwan, ROC*

Received 26 May 2005; received in revised form 27 January 2006; accepted 7 April 2006
Available online 5 July 2006

Abstract

Generally speaking, wafer fabrication factories define the photolithography area as the dispatching center of the entire factory. To establish a set of operative dispatching rules in the photolithography area while taking into consideration the rework of defective products would assist in coordinating and balancing the workload of the entire production line. Furthermore, it would help to enhance both the productivity and efficiency of the wafer fabrication, reduce the on-line WIP stock, shorten the production cycle time, and satisfy the requirements of customers regarding production due time and product quality. This research uses on-line rework as the basis for bringing the factor of reworking of a batch process into the dispatching rule for measurement. It then develops the dispatching rule (Rw-DR) which includes the rework strategy. In addition, this research focuses on the batch with high finished proportion in the photolithography area for finding a way to complete the manufacturing procedure faster, lighten the machine workload of the waiting line, and at the same time increase the output quantity.

Keywords: Photolithography area; Dispatching rule; On-line rework; Finished proportion

1. Introduction

Semiconductor manufacturing belongs to the high capital-technology-intensive industries. The equipment for wafer fabrication requires high capital investment for both initial purchasing as well as maintenance costs, resulting in a very high total production cost. Wafer fabrication plants expect to enhance the throughput by

☆ This manuscript was processed by Area Editor Maged Dessouky.
* Corresponding author. Tel.: +886 2 2771 2171x2346; fax: +886 2 2731 7168.
E-mail address: zhche@ntut.edu.tw (Z.H. Che).

0360-8352/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cie.2006.04.002

raising the machine utility and thereby decreasing production cost. However, the idea of raising the machine utility is contrary to cutting down the setup time and the work in process, and introduces an increased level of complexity in the production control of wafer fabrication.

The equipment investment for the photolithography area, is the highest among all areas of the wafer fabrication plant, and as such is the financial bottleneck for processing wafers. Goldratt (1990) referred to resource constraints in the bottleneck as being the major factor to affect the production performance. The key objective of production control is to decide how to use the bottleneck machines' utility, how to decrease the cost of the work performed in the process, and how to at the same time raise the throughput in order to improve the system's performance. Hence, up to now, scheduling has been the main area of concern for many researchers.

Because of economic considerations, a defective wafer must be reworked even though the wafer fabrication technique is improved continuously. Using the rework process to enhance the yield of wafer production is common in the plants. Consequently, to achieve the objectives of enhanced yield, reduced cost of the work in the process, reduced cost of material wasted, and on time delivery, a strategy of quality rework is necessary to process the rework of the wafers. In addition, since the photolithography area is the bottleneck of the entire plant, an effective dispatching rule for arranging jobs in this area becomes indispensable. Therefore, the basic principle of this paper is to develop a superior dispatching rule (Rw-DR) which considers reworking the defective wafer, the need of the production line to increase its production capability, the need to reduce the work in the process, and the need to shorten the cycle time for responding to the customer's request for quality, quick delivery, etc.

The remainder of this paper is organized as follows. Section 2 reviews the literatures hat relate to the rework strategies at the photolithography stage, dispatching rules, and performance indicators for wafer fabrication. Section 3 provides the definitive description of the problem. In Section 4, the methodology of the dispatching rule is developed for the scheduling of photolithography with the consideration of on-line rework. Section 5 develops the simulation model for this scheduling problem and shows the analytic results. Finally, a brief summary and conclusions are presented in Section 6.

2. Literature review

2.1. Rework strategies at the photolithography stage

In wafer fabrication, the photolithography process is used to define the graph of each layer. In each layer, the circuit patterns and components are constructed, which need a re-entry process in order to be completed. In order to complete this difficult and precise task, the shortest illuminant wavelength and the most precise image lens must be used. The precision of these high precision instruments which are so costly, is easily influenced by pressure, dust in the air, temperature, humidity, etc. As a result, the photolithography area becomes the bottleneck of the wafer fabrication plant (Akcal, Nemoto, & Uzsoy, 2001; Hung, 2003; Sha, Hsieh, & Chen, 2001).

When the wafer has undergone the photolithography process, it shall be subject to quality control "after development inspection" (ADI) before continuing the follow-up processes. The wafer shall be defined "mother lot" if it is accepted in ADI, and shall continue the follow-up processes. Inversely, it shall be defined "child lot" if it does not has good quality, and then it shall be reworked (Zargar, 1995). The rework will ensure the quality of the child lot, but also enhances the queuing time of the mother lot for completing the whole process. The whole photolithography process is shown in Fig. 1. The main mission of the rework strategy is to arrange child lots and mother lots for future processing so as to maximize wafer production efficiency.

Four rework strategies were proposed by Zargar (1995): (1) Lock-step strategy (LSte): the mother lot stops and waits to remerge with the child lot until the rework is completed before the next processing stage. In this strategy, the child lot shall be defined as a hot lot with a high processing priority to be directly reworked, so as to avoid that the mother lot spends a long waiting time. (2) Lot-split strategy (LSpl): the mother lot does not wait for the child lot and continues on to the next processing stage; the child lot is defined as an independent lot to be reworked, and then goes on the next processing stage. (3) Lot-staging strategy (LSta): the mother lot does not wait for the child lot and continues on to the next processing stage; the child lot must accumulate

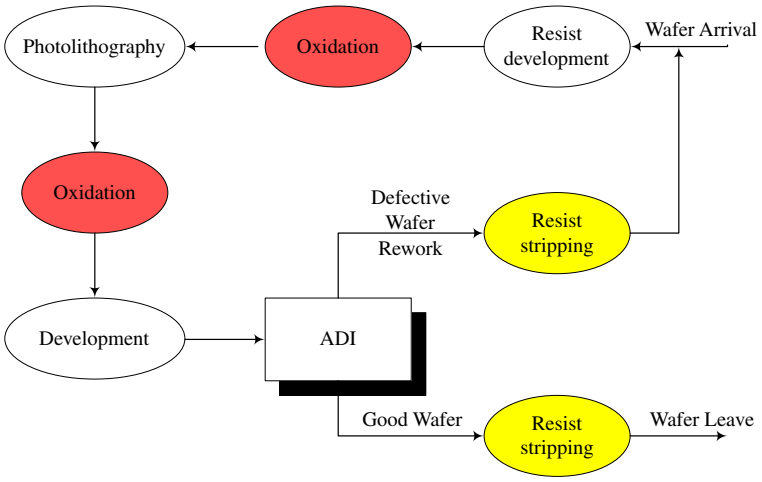

Fig. 1. The rework process in the photolithography area.

wafers up to a specific quantity to be a new lot, and then proceeds with the follow-up stage. (4) Step-mew strategy (SMew): the mother lot continues to be processed in the next step and the reworked child lot joins the next mother lot to be processed in the next step.

Of the four strategies, the second and third do not give priority to the reworking of the child lot. Instead, the child lot has to wait in line to be processed. Zarger drew the conclusion that the fourth strategy is the best by comparing the wafer cycle time of each of the four strategies. But as Zarger also mentioned in his research, the third and fourth strategies have not been implemented in industry practice owing to the difficulties in tracing.

Akcal et al. (2001) consider the test run policies in their research. In a test run, a single wafer from the lot is processed and then inspected. If the test wafer meets specifications, the stepper is qualified for the process. Otherwise, parameters of the stepper should be adjusted and test runs are repeated until the stepper is qualified. The rework process will be done for the defective layer on the test wafer. Three different test policies are considered in their research.

1. Test wafer joins its lot after rework.
2. Test wafer form a new lot for rework.
3. Test wafer is reworked with its lot.

The methodologies of policy 1 and 2 are similar to LSte and LSta separately, Zargar (1995) proposed. Policy 3, if the test wafer fails the inspection, processing is not interrupted and the entire lot is sent to stripping, while the stepper is adjusted and another test run is made. In this policy, test wafer is reworked with its lot, will cause more defective wafers and waste much more time and resource to repair them. In practice, it is very difficult to be adopted.

Sha et al. (2001) tried to find a better strategy for reworking child lots at the photolithography stage, in order to minimize their influence on other normal lots and at the same time to manage mother lots and child lots efficiently. They used simulation to compare rework strategies including the four strategies mentioned by Zarger, as well as the one proposed in their own research (Rendezvous strategy: Rend). In the Rendezvous strategy, the mother lot does not wait for the child lot, but instead goes ahead to be processed in the next step while the child lot is being reworked. The child lot continues on to be processed on its own in the following stage. After this is done, the mother lot and the child lot are rejoined for further processing. In the strategy, the child lot is designed as a hot lot, given the priority to be reworked first.

Based on the above descriptions, the process becomes more complex for the photolithography stage that involves a rework task. Therefore, for enhancing the performance of the photolithography process considering defective wafer rework, and so as to decrease waste, a quality dispatching rule will be developed in this research.

2.2. Dispatching rules

Numerous researches focus on developing suitable dispatching strategies for production control in wafer fab, except for the typical dispatching strategies, such as FIFO, EDD, CR, SPT, and others. Dabbas (1999) divided the dispatching rules into two types by considering the scopes of time and space of decision information: (1) local rule: this type includes several rules, as FIFO, SPR, EDD, CR, ATC, etc. (2) global rule: the DBH, NACH, WR, SA, etc., are contained in this type.

Glassey and Resende (1988) developed a dispatching strategy, SA+, to complement the Starvation Avoidance input regulation policy. In Wein's simulation model (Wein, 1988), 14 dispatching strategies were included and altered based on wafer fab's specific characteristics. In the research done by Kim, Lee, Kim, and Roh (1998), workstations in the wafer fab were divided into two groups according to the utilization level, which were photo workstations and non-photo workstations such as workstations for deposition, etching and ion implant. Different dispatching strategies were applied to different groups. Yoon and Lee (2000) employed stochastic petri nets to model and analyze the machine module, and to define operation due dates using a novel utilization index metric. An operation due date strategy for lot dispatching was proposed and compared with other dispatching policies. The methodologies of dispatching rules mentioned above can be found at Appendix A.

During the process of wafer manufacturing, defects often occur due to the conditions of physical environment, human errors, differences between machines, etc. In most production procedures, defective wafers are abandoned because they cannot be reprocessed. However, in the photolithography stage, wafer defects can be redressed. Although redressing defects by reprocessing the wafer increases cycle time and manufacturing costs, it can also reduce the costs that are associated with defects. Thus, how to arrange the job while considering the rework of wafers during the photolithography stage is an important topic to be studied (Sha et al., 2001).

2.3. Performance indicators for wafer fabrication

In the wafer fabrication system, when deciding which production model is best for a group of waiting jobs, several performance indicators are commonly used: utilization, throughput, quantity of WIP, cycle time, and due date, etc. Table 1 shows several related papers on system assessing by using specific performance indicators.

Table 1
Several related literatures of performance indicators for measuring wafer fabrication system

Author	Cycle time	Throughput	Delay	Queuing time	Tardiness	Variance of cycle time	Variance of delay	On time delivery	Utilization of bottleneck machine	Quantity of WIP in system
Apolloni, S., Lando, M., and Savino, M. M. (2004)	x	x								x
Chao and Sivakumar (2004)	x	x				x			x	x
Rooda and Vervoort (2003)	x	x				x			x	x
Dabbas and Fowler (2003)	x					x	x	x		
Dai and Neuroth (2002)	x	x								x
Andradottir et al. (2001)		x								x
Kidambi (2001)	x	x					x		x	x
Kim et al. (1998)		x			x					
Lu, S. C. H., Ramaswamy, D., and Kumar, P. R. (1994)	x				x					
Weng and Leachman (1993)				x						
Yan, H., Lou, S. X. C., and Sethi, S. P. (2000)										x
Lee and Park (1991)	x	x						x	x	
Glassey and Resende (1988)	x	x								
Wein (1988)	x									

According to these literatures, in the past, studies aimed at using single or a small number of indicators for measuring system performance. In recent years, multiple performance indicators have constantly been used in system assessing, so as to achieve the best production situation. Therefore, in this paper, multi-indicator analyzing is employed for evaluating the performance of the proposed methodology which is used to plan the photolithography process for wafer fabrication.

3. Problem description

The production process (Fig. 2) of wafer fabrication that includes reentry is very complex, it has a long cycle time, rework, and the forward process affects the backward process, making the production control become more difficult. In addition, the character of the wafer fabrication process becomes more uncertain, because several situations are frequently faced: equipment-down, batch/non-batch production, stable processing quality, etc. Therefore, to ensure that the process is maintained with a quality yield and throughput, the firms usually uses rework to reduce the influence of uncertainty.

Generally there are two types of machines in the fab. Serial machines which process one wafer at a time and batch machines which work on multiple wafer/lots. Most of the machines are belonging to series machine including the bottleneck machines, photolithography. Batching machines always are critical machines, like heat-treating ovens (oxide film), plating baths, kilns for drying lumber, and in semiconductor wafer fabrication, diffusion and oxidation ovens. Often the maximum batch size (machine capacity) of such operations is greater than the size of the arriving lots.

The photolithography area mainly includes three processes: coating, exposure, and development. To ensure the photolithography process accuracy, a quality inspection will be carried out after the development process is accomplished, referred to as 'after development inspection' (ADI). If defective wafers are detected in the ADI, they will be reworked to redeem their quality. In addition, owing to the fact that the amount of invested capital in wafer production is enormous, throwing wafers with only slight defects away is excessively wasteful. Hence, enhancing the yield and throughput of wafer by means of a rework strategy is the focus of this study.

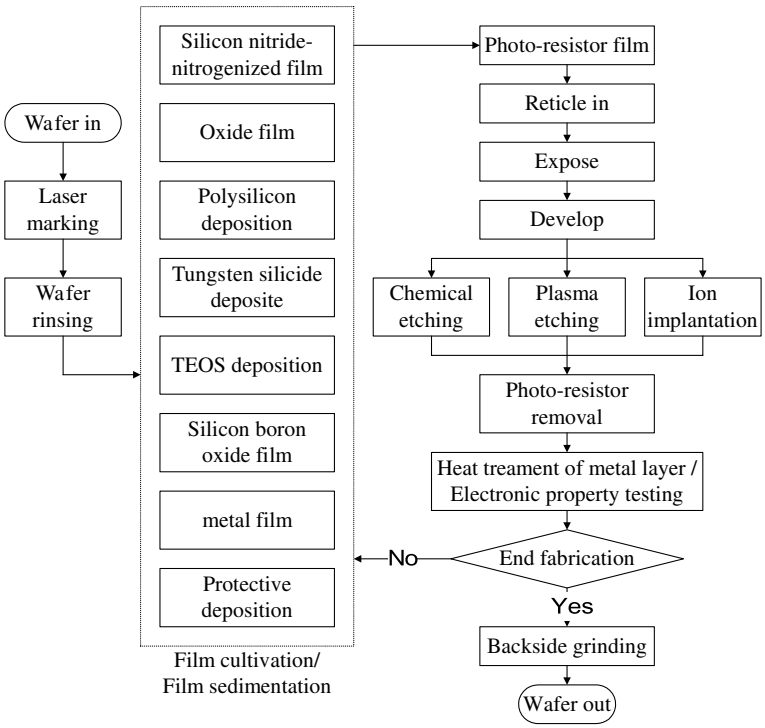

Fig. 2. Basic operations sequence of wafer fabrication (Sha & Hsu, 2004).

Past researches on the process in the photolithography area, only discussed improving machine utilization, on time delivery, and throughput, etc., and very few studies brought the rework strategies into the dispatching rules in order to measure system performance. Therefore, this research emphasizes the development of a quality dispatching rule which is in accordance with the system variations, in order to perform the different rework strategies for defective wafers. This proposed rule will enhance the production performance for a wafer fabrication plant by balancing system load and improving throughput.

4. Proposed methodology

In this study the following notations were used to develop the integrated rework-dispatching model.

k: quantity of a lot in the queuing line.
PT_{iz}: processing time of lot i in workstation z.
PT^c_{iz}: processing time of the child lot of lot i in workstation z.
PT^c_i: processing time of the child lot of lot i in the photolithography area.
PT^m_{iz}: processing time of the mother lot of lot i in workstation z.
PTR_z: remaining processing time of the working lot in workstation z.
PTR: remaining processing time of the working lot in the photolithography area.
RTR^c: processing time of the child lot in resist stripping.
IT^c: inspecting time of the child lot in ADI.
RT^c: rework processing time of the child lot.
WAT^m: queuing time of the mother lot for entering the next machine of the photolithography area.
N_{iz} : quantity of wafers of lot i in the queuing line of workstation z.
FT^m_i : flow time of the mother lot of lot i for achieving the next capacity-constricted resources machine.
FT^c_i : flow time of the child lot of lot i for achieving the next capacity-constricted resources machine.
MWT_z: average expecting queuing time of a lot in the queue line of workstation z.
TPT_i: total processing time of lot i.
RPT_i: remaining processing of lot i.
FP_i: finished proportion of lot i.
FWR_i: finished weighted ratio of lot i.
$Weight_i$: weight of lot i.
CR_i: critical ratio of lot i.

There are several rework strategies presented in the literature. In this paper, we adopt the rule which considers the length of queuing line of the follow-up process after photolithography area to decide whether the mother lot should wait for the child lot in photolithography. The rework procedure is as follows:

Step 1: The detective wafer is detected at the ADI workstation.
Step 2: Calculate the rework processing time of child lot (RT^c) and the queuing time of the mother lot for entering the next machine of the photolithography area (WAT^m).

$$RT^c = RTR^c + PTR + PT^c_i + IT^c$$

$$WAT^m = PTR_z + \sum_{i=1}^{k} N_{iz}(PT_{iz})$$

Figs. 2 and 3 express the meanings of these two arithmetic patterns separately.
Step 3: Compare the RT^c and WAT^m.
If $RT^c > WAT^m$, go to Step 4.
Else, the child lot is defaulted as a hot lot that can directly be reworked in the photolithography process without having to enter the queuing line. The mother lot shall wait to remerge with the child lot to become the original lot when rework is completed, before continuing with the follow-up processes.

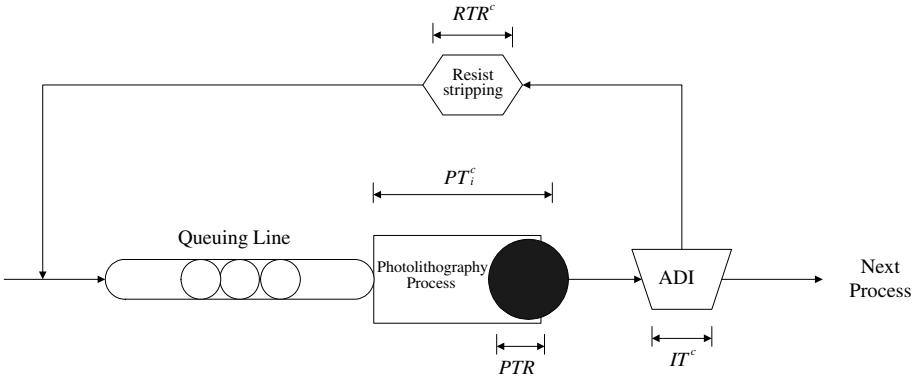

Fig. 3. The rework process of the child lot in the photolithography area.

Step 4: The mother lot continues with the follow-up processes without waiting for the child lot.

If the next process is the capacity-constricted resources station, the child lot is not defaulted as a hot lot to be reworked in the photolithography process. The mother lot will remerge with the child lot to become the original lot at the next machine of the capacity-constricted resources station.

Otherwise, the child lot is defaulted as a hot lot that can directly be reworked in the photolithography process without having to enter the queuing line. The mother lot shall wait to remerge with the child lot to become the original lot at the capacity-constricted resources station in the follow-up processes when rework is completed.

On the basis of the above rework strategy, there are three types of lots in the queuing line at the photolithography area. First, the child lot is defaulted as a hot lot that can directly be reworked. Second, the child lot is not defaulted as a hot lot to be reworked. Third, the original lot is in the queuing line. The following proposed dispatching procedure is used to decide what these three-type lots are, in order to enter the photolithography process.

The CR rule is employed as the base of the proposed dispatching procedure. In addition, in accordance with the lot with the higher finished ratio in the photolithography area being given a higher weight, ensures that it shall be finished as early as possible. This dispatching rule can initiate a reduction in the load of the queuing line, and enhance the throughput of the production system. Hence, this proposed dispatching rule not only considers the higher finished ratio lot but also allows for the flow time of the mother and child lots. The complete procedure is shown as follows:

Step 1: When the process is completed in the photolithography area, check the lot types in the queuing line. If there is a hot lot that requires reworking, let this child lot enter the photolithography process first. Else, go to Step 2.

Step 2: If there is a child lot which is not a hot lot that requires reworking, calculate the flow time of the mother lot for achieving the next capacity-constricted resources machine (FT_i^m) and the flow time of the child lot for achieving the next capacity-constricted resources machine (FT_i^c), then go to Step 3. If there are not rework lots, go to Step 3.

$$FT_i^c = RT^c + \sum_{j=1}^{n+1} PT_{i(z+j)}^c + \sum_{j=1}^{n+1} MWT_{z+j}$$

$$FT_i^m = WAT^m + \sum_{j=1}^{n+1} PT_{i(z+j)}^m + \sum_{j=1}^{n} MWT_{z+j}$$

Individually, Figs. 4 and 5 indicate the meanings of these two arithmetic patterns.

Step 3: Calculate the $CWCR_i$ for each lot in the queuing line of the photolithography area. The lot with the smallest $CWCR$, shall be dispatched first.

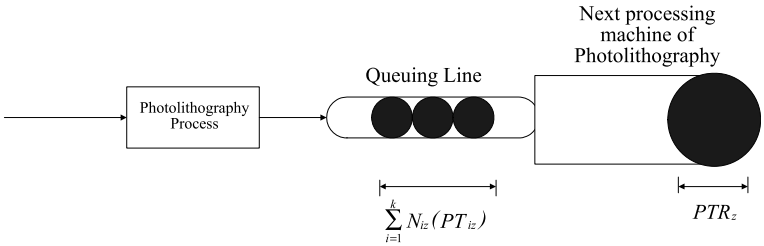

Fig. 4. The queuing time of the mother lot before entering the next photolithography machine.

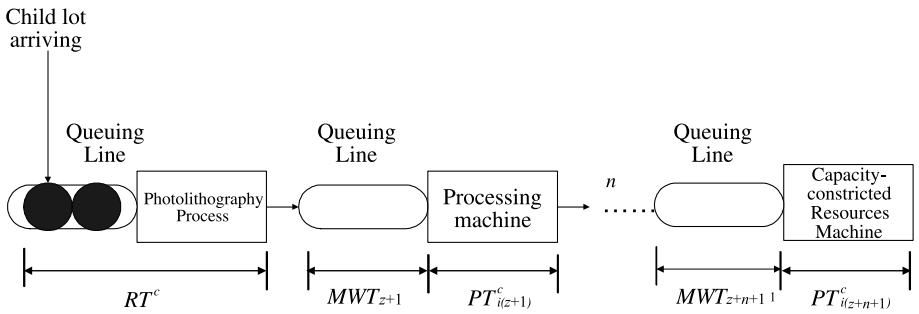

Fig. 5. The child lot arrives at the capacity-constricted resources machine.

$$CWCR_i = CR_i \times Weight_i \times FWR_i$$

$$CR_i = \frac{\text{time to due date}}{\text{total remaining production time}}$$

$$Weight_i = \begin{cases} 1 & \text{weight of the original lot} \\ \frac{FT_i^m}{FT_i^c} & \text{weight of the rework lot} \end{cases}$$

$$FWR_i = \begin{cases} 1 - \left(\frac{0.3 \times TPT_i - RPT_i}{0.3 \times TPT_i}\right)^2 & \text{if } FP_i > 0.7 \\ 1 & \text{if } FP_i \leqslant 0.7 \end{cases}$$

$$FP_i = \frac{TPT_i - RPT_i}{TPT_i}$$

The meaning of *FWR* and *FP* is clarified in Fig. 6.

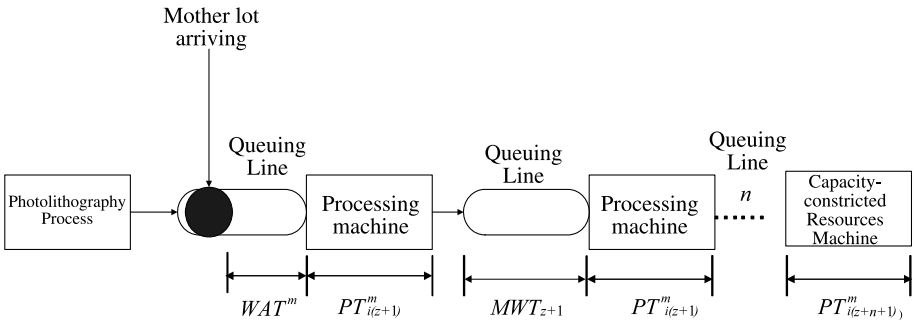

Fig. 6. The mother lot arrives at the capacity-constricted resources machine.

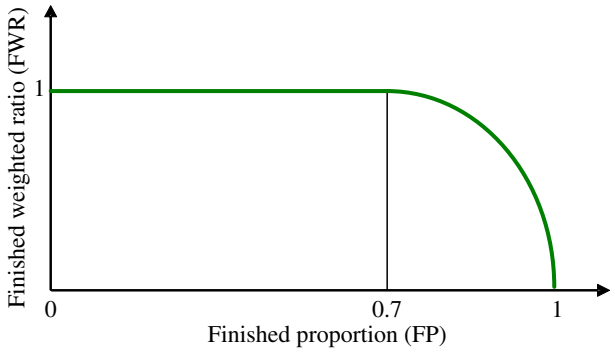

Fig. 7. The relationship between the finished proportion and the finished weighted ratio.

The main concept in developing the dispatching rule is the fact that the state of the next machine of the photolithography area is the critical factor to influence the performance of the wafer fabrication. In this rule, for reducing the queuing time for the mother lots, the weight is adopted to do this task. If the weight of the lot is 1, the flow times of the mother lot and the child lot are equal. That means that there is no difference in which lot is dispatched first. If the weight of the lot is less than 1, then the child lot's flow time of achieving the next capacity-constrict station is greater than that of the mother lot. In that case the child lot shall be processed first so as to reduce the queuing time of the mother lot. Based on this notion of weight, the *CR* and weight are integrated in this dispatching rule. The child lot has a high priority to be dispatched when its *CR* is multiplied by a smaller weight. If the weight is 1, the *CR* of the child's lot shall remain unchanged, and its processing order remains unmodified.

To briefly and clearly describe the developed dispatching methodology for the scheduling of photolitho graphy with the consideration of on-line rework, the overall procedure is shown in Fig. 7.

5. Simulation model and result analysis

In this paper, the configuration of wafer fabrication considered is the *H* wafer fabrication plant in Taiwan. The eM-Plant software is employed to construct the virtual wafer fabrication plant. The *H* plant provides all simulated data for developing the simulation model, including: process of the wafer fabrication, product data, product quantity, processing time, repair time and down time of machine, etc. The simulation results are analyzed by adopting the Statistica 6.0 statistic software. The simulation tests and statistic analyses are done on personal computers with Pentium III 800 processors.

In the virtual plant, 53 workstations and 300 machines are included for performing the production tasks. There are three product types – DRAM, SRAM, and LOGIC produced in this plant. For these three product types, the product mix is 4:7:9 and the loops of processing are 15, 18, and 17 for each type. The reworking rate is 3% after ADI, and the order review and release strategy utilized is CONWIP. CONWIP is the abbreviation of constant work in process. CONWIP regulating new wafer released to maintain a constant number of lots in the production system. CONWIP starts a new lot whenever a lot is completed. The level of WIP in our simulation model is 750 lots for maintaining the system utilization to 90%.

The simulation day is set at 515, the warming up period takes up 150 days. Data collection is performed one year from the 151th through the 515th simulated day. The due day of each lot is calculated as the total processing time multiplied by 4.4.

The following assumptions are to build into the virtual wafer fabrication plant model:

1. Product mix and product demand are stable.
2. Lot sizes are fixed (a cassette contains 24 wafers).
3. The quantities of machines, mean time between failure (MTBF), mean time to repair (MTTR), and processing time are constants.
4. The single bottleneck is in the photolithography area.

Table 2
Whole combinations of compared forms

Rework strategies	Dispatching rules		
	FIFO	EDD	CR
LSte	Le-FIFO	Le-EDD	Le-CR
LSta	La-FIFO	La-EDD	La-CR
Rend	Re-FIFO	Re-EDD	Re-CR

5. The reworking rate is stable.
6. Defective lots shall become intact lots after being reworked once.
7. Transfer time between workstations is ignored.
8. If two or more hot lots need to be reworked, the FIFO rule is applied.

For exploring the advantage of the proposed integrated approach (Rw-DR), nine different production control combinations are employed as comparison forms, which are integrated by three rework strategies (LSte, LSta, and, Rend) and three dispatching rules (FIFO, EDD, and CR). All of the combinations are displayed in Table 2. Each combination/comparison form and the Rw-DR shall be performed 30 simulation times for acquiring the analysis data.

Five performance indicators are used to evaluate the efficacy of each alternative, and are formularized and listed as follows:

1. *Mean flow time:* The average of the flow time of all the finished orders. Flow time is meaned the time between order releasing and order finishing. The definition is:

$$\text{Mean flow time} = \left(\sum_{i=1}^{n} f_i - r_i\right) \Big/ n$$

where r is the order releasing date, f is the order's finished date, n is the number of finished orders.
2. *On time delivery:* On time delivery is defined as the ratio of order's tardiness is smaller than 1 day. That is, the forecasted due date is not larger or smaller than one day of the actual flow time. The definition is:

$$\frac{\text{number of the orders that's tardiness is small than 1 day}}{n}$$

3. *Mean tardiness:* The average tardiness of all finished orders. The definition is:

$$\text{Mean tardiness} = \left\{\sum_{i=1}^{n}[\max(0, d_i - f_i) + \max(0, f_i - d_i)]\right\} \Big/ n$$

where d is the order's due date, f is the order's finished date, N is the number of finished orders.
4. *Work in process:* The number of works in process.
5. *Mean flow time of rework lots:* The average flow time of all the rework lots.

Table 3 summarizes the simulation results based on the mean of the responses for the five different performance measures. For confirming the statistical significance of the above results, the ANOVA technique was used. Rejecting the null hypothesis means that not all means of all responses are equal, and the Duncan's multiple-scope tests are used to conclude which pair of means are not equal. Five different ANOVA's were evaluated for the simulations: Mean flow time, On time delivery, Mean tardiness, Work in process, and Mean flow time of rework lots. The significance level for all analyses was set at $\alpha = .05$. Thirty duplicates were generated for each production control strategy using simulation. Analyses substantiated that the results are all statistically significant (Table 3).

For instance, the results of the mean flow time ANOVA analyzes that this model rejects the null hypothesis, and concludes that the mean flow time performance indicator responses for the production control strategies

Table 3
Comparisons between each combinational compared form and the proposed integrated approach

Performance indicators	Proposed integrated approach (Rw-DR)	Combinational compared forms									% improve Rw-DR vs. best compared form
		Best	←	—	—	—	—	—	→	Worst	
Mean flow time	48.76	Re-CR 50.85[a]	Le-CR 51.66	La-CR 51.31	Le-EDD 51.86	La-EDD 52.03	Re-EDD 53.21	Re-FIFO 60.89	La-FIFO 61.00	Le-FIFO 62.30	4.11
On time delivery	0.704	Re-CR 0.488[a]	Le-CR 0.442	La-CR 0.430	La-EDD 0.380	Re-EDD 0.333	Le-EDD 0.324	Le-FIFO 0.003	La-FIFO 0.001	Re-FIFO 0.002	44.26
Mean tardiness	2.41	Re-CR 3.19	La-CR 3.72[a]	Le-CR 3.98	Re-EDD 5.10	La-EDD 5.26	Le-EDD 6.22	Re-FIFO 13.01	La-FIFO 13.12	Le-FIFO 14.42	24.45
Work in process	675.59	Le-EDD 734.67[a]	Re-EDD 738.82	La-EDD 739.90	La-CR 748.44	Re-CR 749.06	Le-CR 749.04	Le-FIFO 750	La-FIFO 750	Re-FIFO 750	8.04
Mean flow time of rework lots	48.97	La-EDD 50.97[a]	Re-CR 51.06	La-CR 51.15	Le-CR 51.99	Re-EDD 52.10	Le-EDD 53.28	La-FIFO 60.92	Re-FIFO 61.02	Le-FIFO 62.43	3.92

[a] Statistically significant (P-value = 0.05).

are not all the same. Therefore, the Duncan's test is performed to group all strategies among this performance indicator, and the test results in Table 3 show that the Rw-DR is a single strategy in the first group. The Rw-DR's mean flow time is 2.09 (4.11%) days better than the best combinational compared form Re-CR's mean response of 50.85 days.

In addition, the on time delivery is 44.26% higher than the best combinational strategy response (Re-CR (0.488)) has exhibited, and is the most improvement when using the Rw-DR. Similarly, the simulation tests and statistic analyses use the on mean tardiness indicator for measuring performance. The response of Rw-DR is grouped with that of the Re-CR strategy, and is better than the response of other combinational strategies. All other responses for the Rw-DR have also shown statistically significant improvement over the use of the combinational strategies. Explicitly, mean tardiness, work in process, and mean flow time of rework lots has improved by 24.45%, 8.04%, and 3.92% respectively.

Combined, the responses are worst under all performance indicators if using the FIFO rule to collocate with any rework strategy. However, at the same time the combinational strategy (Re-CR) which combines the rendezvous reworked strategy and the critical ratio rule has better responses for performing the wafer fabrication.

Furthermore, we use the mean flow time and the standard deviation of the mean flow time to compare the production performances between the entire plant and the rework lots. Fig. 8 shows that only two production control strategies La-EDD and Re-EDD bring about different mean flow times between the entire plant and the rework lots, and the rework lots' responses are shorter than that those of the entire plant. Other mean flow time performances of combinational strategies (including Rw-DR) are too close for entire plant and rework lots. Adopting the Rw-DR, considering simultaneously planning of the original lots and rework lots, will reduce the mean flow times of the entire plant and rework lots at the same time. For the standard deviation of the mean flow time, the responses of Le-CR, Re-CR, and Rw-DR are not the same between the entire plant and the rework lots. Other combinational strategies take on approximate responses.

However, the standard deviation of the mean flow time of the proposed approach is higher than some combinational strategies, as shown in Fig. 9. There are two causes to generate this outcome: (1) According to different system situations, the dispatching rule of the proposed approach for reworking defective wafers shall have different dispatching behaviors. When the mother lot has a longer queuing time in the next station of the photolithography area, the mother lot shall stay behind waiting for the child lot, which causes the flow time of the rework lots to be increased. On the other hand, if the mother lot directly continues the follow-up processes without waiting for the child lot, the rework lots' flow time shall decrease when the mother lot's queuing time for waiting for processing at the next station of the photolithography area is shorter. (2) For the dispatching rule of the proposed approach, if there are no hot lots to be reworked, the original lots and the rework lots shall be ordered into the photolithography area processing. The rework lot needs to spend more

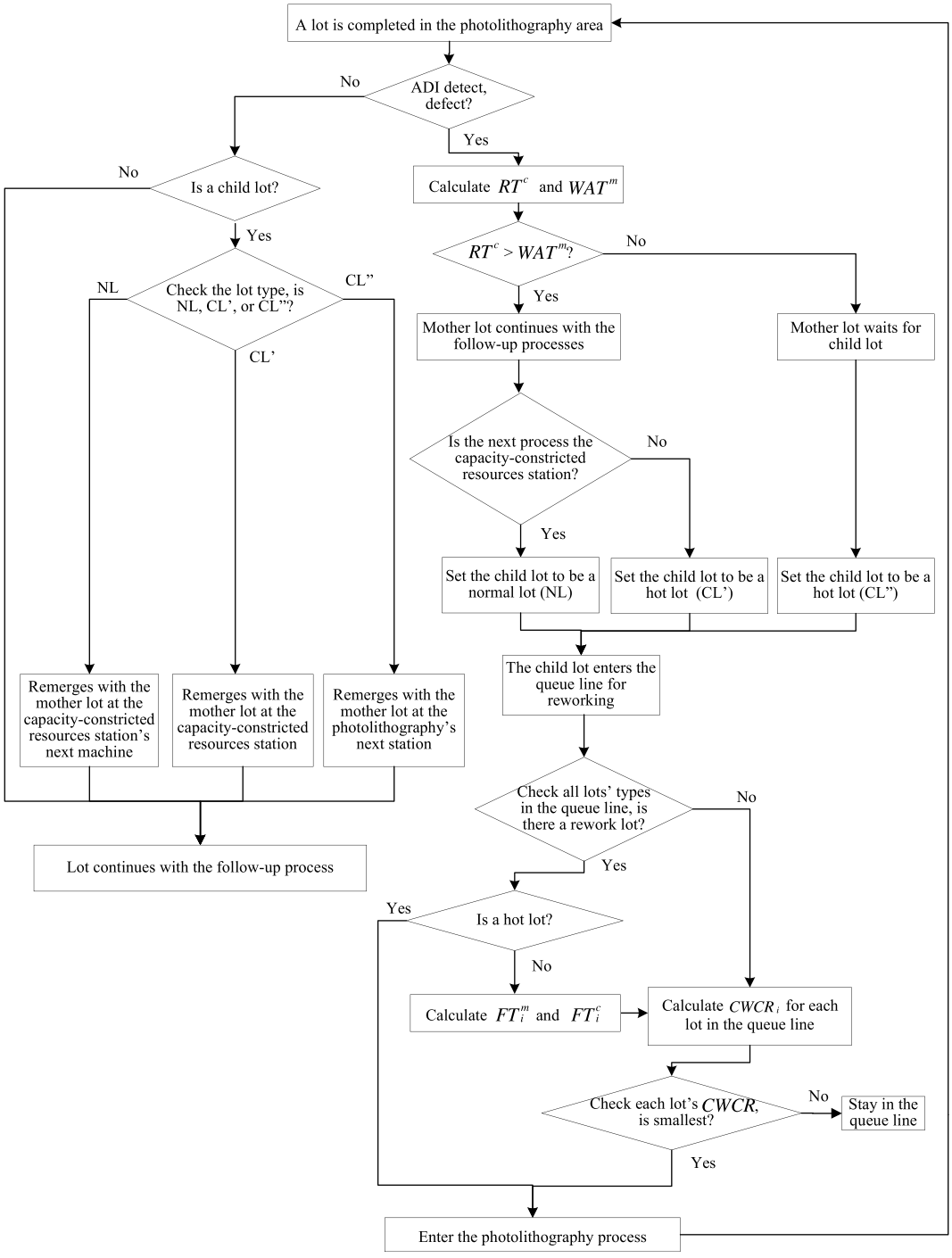

Fig. 8. Overall procedure of the developed dispatching methodology with the consideration of on-line rework.

time in the queuing line until the original lot processing is completed, if the original lot has the higher priority. These two reasons can cause the mean flow time to have some slight variations (Fig. 10).

In short, five performance indicators, mean flow time, on time delivery, mean tardiness, work in process, and mean flow time of rework lots, have improved notably by using the proposed approach (Rw-DR). These results can enhance the degree of satisfaction of customer demand.

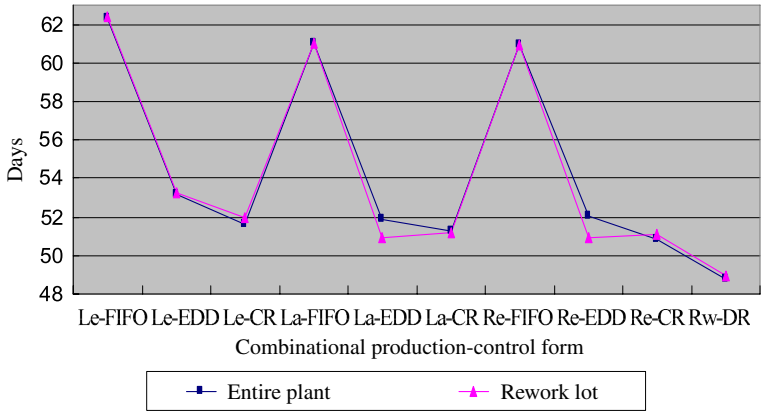

Fig. 9. Comparisons between the flow times of the entire plant and the rework lot.

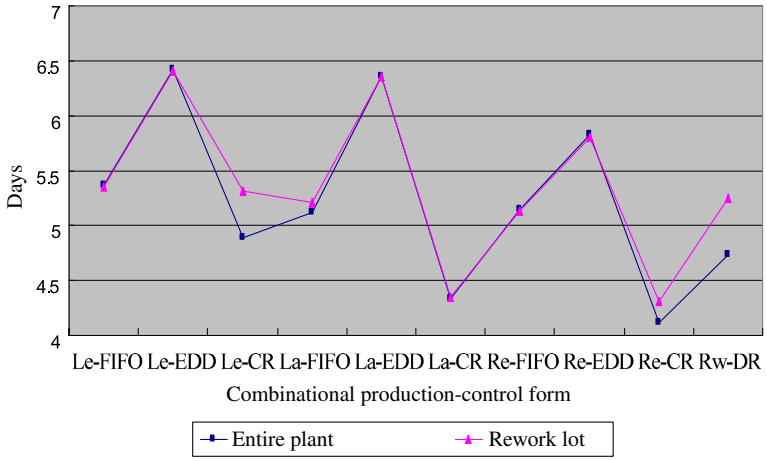

Fig. 10. Comparisons between the divisions of the flow time of the entire plant and the rework lot.

6. Conclusions

The analysis of dispatching is important for determining the dispatching rule for the rework lots in the photolithography area of wafer fabrication. In this paper, a dispatching rule (Rw-DR) was proposed that integrates the rework strategies while considering the capacity-constricted resource machine and taking into account both original lots and rework lots. A virtual wafer fabrication simulation model was tested. Different performance indicators (mean flow time, on time delivery, mean tardiness, work in process, and mean flow time of rework lots) were used to measures the Rw-DR achievements, and the results were compared to the use of the combinational strategies. Results showed that the performances of the proposed approach are improved under these indicators.

The results also showed that different combinational strategies present different outcomes when evaluated by different performance indicators. For example, La-EDD proves to be a remarkable strategy when assessed under the mean flow time of rework lots indicator, but is not so impressive when assessed using other measure indicators. This illustrates that using different combinational strategies to achieve quality results for different measures simultaneously is very difficult. However, the Rw-DR does achieve this goal, which by itself is sufficient demonstration that this approach is a quality methodology for dispatching wafer fabrication in the area of photolithography.

Appendix A. Methodologies of dispatching rules

1. *FIFO*. Select the lot, which arrived in the queue at the earliest time.
2. *EDD*. Select the lot, which has the earliest due date in the queue.
3. *CR*. Smallest CR, CR = (due date-total remaining PT − present date)/total remaining PT.
4. *WR* (*workload regulation*, Wein, 1988). WR regulating new wafer releases to maintain a constant amount of expected work at a bottleneck station. WR monitor the sum of remaining processing times at the bottleneck workstation for all lots in the fab and release a new lot when this sum falls below a critical value. Here, throughput can be controlled by changing the critical value.
5. *SA* (*starvation avoiding*, Glassey & Resende, 1988). SA released a new wafer lot to avoid starvation of a bottleneck workstation. SA starts a new lot to avoid idling the bottleneck workstation due to lack of work. More specifically, a new lot is released when virtual inventory at the bottleneck workstation falls down to a predetermined value. The virtual inventory at the bottleneck workstation can be estimated by the sum of the actual inventory at the workstation and WIP at upstream workstations that is expected to arrive at the bottleneck workstation within the lead time (L). Here the lead time is estimated with the sum of processing times operations that must be processed for a newly released lot before it visits the bottleneck for the first time. The predetermined value that triggers a lot release can be set to $\alpha \cdot L$, where α is a control parameter.
6. *SRPT*. Select the lot that has the shortest expected remaining processing time until it exist the fab.
7. *SA+* (Glassey & Resende, 1988). Assign high priority to jobs that are close to the bottleneck station and/or that contribute a large amount of work content to the station. The main characteristic of this SA-booster dispatching rule is its dynamic behavior. The rule combines two simple rules by mean of weights and dynamically changes the weights according to the state of the system. If the bottleneck is in no danger of starvation the dispatching rule gives more weight to the SRPT rule, while if there is imminent danger, more weight is given to a rule (SA+) that gives high priority to lots that are headed for the bottleneck station. The composite SA dispatching rule is a weighted mix of SRPT and SA+.
8. *ATC*. ATC is the abbreviation of apparent tardiness cost. The priority of lot i is equal to $\frac{D_i - b(R_i - TP_i) - TP_i - CT}{k * APT * TP_i}$. D_i, R_i, TP_i are the due date, remaining work, and total processing time of lot i. CT and APT are current time and average processing time of all lots in the fab. b and k are parameters for ACT.
9. *NACH* (Fowler, Hogg, & Philips, 1992). Next arrival control heuristic (NACH) was proved to be a robust heuristic in case forecasting data on future arrivals are used, i.e. estimated arrival moment for new lots. The decision rule of NACH in single product and single machine is:If $q \geqq C$ then loading the jobs.Else if $[q(t_1 - t_0) - (t_0 + T - t_1)] < 0$ (i.e. $\mathrm{NACH}_1 < 0$) then waiting for next arrival job.Else loading the jobs. NACH considers the next arriving job. There are two choices at time t_0, loading the jobs or waiting another new job. Waiting for next new arrival job is represented for a new decision phase.

References

Akcal, E., Nemoto, K., & Uzsoy, R. (2001). Cycle-time improvements for photolithography process in semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing, 14*(1), 48–56.

Andradottir, S., Ayhan, H., & Down, D. G. (2001). Server assignment policies for maximizing the steady-state throughput of finite queuing systems. *Management Science, 47*(10), 1421–1439.

Apolloni, S., Lando, M., & Savino, M. M. (2004). Advanced techniques of theory of constraints and activity based costing for scheduling of high technology production lines. In *Proceedings of the 2nd World Conference and 15th Annual POMS Conference* (in CD-R). Cancun: Mexico.

Chao, Q., & Sivakumar, A. I. (2004). Simulation based cause and effect analysis in semiconductor wafer fabrication. *Journal of the Institution of Engineers, 44*(4), 53–65.

Dabbas, R. M. (1999). *A new scheduling approach using combined dispatching criteria in semiconductor manufacturing systems*. Ph.D. thesis. Department of Industrial and Management Systems Engineering, Arizona State University.

Dabbas, R. M., & Fowler, J. W. (2003). A new scheduling approach using combined dispatching criteria in wafer fabs. *IEEE Transactions on Semiconductor Manufacturing, 16*(3), 501–510.

Dai, J. G., & Neuroth, S. (2002). DPPS scheduling policies in semiconductor wafer fabs. In G. T. Mackulak, J. W. Fowler, & A. Schomig (Eds.), *Proceedings of the international conference on modeling and analysis of semiconductor manufacturing* (pp. 194–199). Arizona: Tempe.

Fowler, J. W., Hogg, G. L., & Philips, D. T. (1992). Control of multiproduct bulk service diffusion/oxidation process. *IIE Transaction, 24*(4), 84–96.

Glassey, C. R., & Resende, M. G. C. (1988). Closed-loop job release control for VLSI circuit manufacturing. *IEEE Transactions on Semiconductor Manufacturing, 1*(1), 36–46.

Goldratt, E. M. (1990). *The haystack syndrome*. North River Press.

Hung, Y. L. (2003). *Dispatching rules with dedicated machines*. Master thesis, Department of Industrial Engineering and Management, National Chiao-Tung University.

Kidambi, M. R. (2001). *Impact of lot dedication on the performance of the FAB*. Master Thesis, Industrial and Systems Engineering, Virginia Polytechnic Institute and State University.

Kim, Y. D., Lee, D. H., Kim, J. U., & Roh, H. K. (1998). A simulation study on lot release control, mask scheduling, and batch scheduling in semiconductor wafer fabrication facilities. *Journal of Manufacturing Systems, 17*(2), 107–117.

Lee, J. S., & Park, K. S. (1991). Joint determination of production cycle and inspection schedule in a production systems. *Journal of Operational Society, 42*, 775–783.

Lu, S. C. H., Ramaswamy, D., & Kumar, P. R. (1994). Efficient scheduling policies to reduce mean and variance of cycle-time in semiconductor manufacturing plants. *IEEE Transactions on Semiconductor Manufacturing, 7*(3), 374–388.

Rooda, J. E., & Vervoort, J. (2003). *Analysis of manufacturing systems*. The Netherlands: Eindhoven. University of Technology, Available online: <http://se.wtb.tue.nl/documentation/lecnotes/mis/AMS.pdf/>.

Sha, D. Y., Hsieh, L. F., & Chen, K. J. (2001). Wafer rework strategies at the photolithography stage. *International Journal of Industrial Engineering, 8*(2), 122–130.

Sha, D. Y., & Hsu, S. Y. (2004). The effects of rework on the shop floor control in wafer fabrication. *Journal of Manufacturing Technology Management, 15*(1), 184–198.

Wein, L. M. (1988). Scheduling semiconductor wafer fabrication. *IEEE Transactions on Semiconductor Manufacturing, 1*(3), 115–130.

Weng, W. W., & Leachman, R. C. (1993). An improved methodology for real-time production decisions at batch-process work stations. *IEEE Transaction on Semiconductor Manufacture, 6*(3), 219–225.

Yan, H., Lou, S. X. C., & Sethi, S. P. (2000). Testing the robustness of various production control policies in semiconductor manufacturing. *Production and Operations Management, 9*, 171–183.

Yoon, H. J., & Lee, D. Y. (2000). A control method to reduce the standard derivation of flow time in wafer fabrication. *IEEE Transaction on Semiconductor Manufacturing, 13*(3), 389–392.

Zargar, A. M. (1995). Effect of strategies on cycle time. In *17th International Conference on Computers and Industrial Engineering* (pp. 239–243).