

PARETO CONTROL IN MULTI-OBJECTIVE DYNAMIC SCHEDULING OF A STEPPER MACHINE IN SEMICONDUCTOR WAFER FABRICATION

Amit Kumar Gupta

Planning and Operations Management Group
Singapore Institute of Manufacturing Technology
71 Nanyang Drive, SINGAPORE 638075

Appa Iyer Sivakumar

School of Mechanical and Aerospace Engineering
Nanyang Technological University
50 Nanyang Avenue, SINGAPORE 639798

ABSTRACT

This paper focuses on Pareto control in multi-objective dynamic scheduling of a stepper machine that is considered as a bottleneck machine in the semiconductor wafer fabrication process. We propose the use of compromise programming method for achieving Pareto control in the needs of conflicting objectives such as mean cycle time, cycle time variance and maximum tardiness. Using conjunctive simulated scheduling, at each decision instance in simulated time, a Pareto job is selected and loaded on the machine for processing. Using the real factory data, we demonstrate the concept of Pareto control in dynamic scheduling and show how a stepper machine can be controlled at specified needs of scheduling objectives. The results obtained from Pareto control approach are superior to the simulated results of actual operating heuristic in the factory.

1 INTRODUCTION

The semiconductor industry is considered as one of the most complex manufacturing systems today. Production scheduling in a semiconductor industry is highly complex due to a large number of process steps, highly reentrant process flows (multiple visits to the workcenters), different manufacturing routes, large variety of equipment types, high product mix variability and sequence dependent set-ups. The jobs at a workcenter may consist of hundreds of product types, each with different specifications. Shrinking product life cycles and customer demand for shorter cycle time and specialized delivery add further complexity. The entire semiconductor industry is sensitive to economic and trade patterns. It is typical in the semiconductor market to have large oscillations in demand. Rapid development of new technology is another factor that adds to this dynamic situation. This creates a need to consider multiple conflict-

ing objectives such as minimization of cycle time mean and its variance and maximization of delivery accuracy while scheduling. Considering these complexities, semiconductor manufacturing poses unique planning and scheduling challenges.

Real life scheduling problem requires the decision maker to consider a number of objectives before arriving at any decision. A solution which is optimal with respect to a given objective might have a poor performance on some other objective. The trade-off involved in considering several different objectives provides useful insights to the decision maker. Thus considering problems with more than one objective is more relevant in the context of real life scheduling problems.

A major portion of the past scheduling research deals mainly with single objective scheduling. Though, a vast amount of literature exists on the optimization of various single objective functions such as Makespan time, average cycle time, machine utilization and tardiness. Only a little effort has been made in the direction of simultaneously considering multiple objectives while scheduling, which may very well be contradicting each other in nature (T'kindt and Billaut 2002). In the field of semiconductor scheduling, the contradicting needs of the multiple objectives were taken into consideration only in a few research papers (Yang and Chang 1998, Dabbas et al. 2001, Gupta and Sivakumar 2005b). Pinedo and Chao (1999) also proposed "scheduling with multiple objectives", as a new direction of research in the field of scheduling as it may often be desirable to see the trade-offs between the different objectives of scheduling. Therefore, the primary focus of this paper is to develop Pareto control curves for conflicting objectives in multi-objective dynamic scheduling environment for serial processes. The relevance of the research findings in industrial application has been presented through a real-life case study for the concept of Pareto control in scheduling a stepper machine in semiconductor wafer fabrication.

2 PARETO CONTROL IN MULTI-OBJECTIVE DYNAMIC SCHEDULING

In semiconductor scheduling, following objectives are of high importance from the perspective of customer and producer:

- Minimizing mean cycle time (\bar{F}) in order to reduce the work-in-process (WIP) inventory and the cost of production.
- Minimizing variance of cycle time ($\sigma_{\bar{F}}^2$) for providing more reliable predictions of completion times to the customers.
- Minimizing maximum tardiness (T_{max}) for meeting due dates and reducing the penalty cost arisen out of unfulfilled demand.

The above objectives constitute the following two pairs of conflicting scheduling objectives:

- Minimizing mean cycle time and minimizing maximum tardiness (VanWassenhove and Gelders 1980, Koksalan 1999) = $\min(\bar{F}, T_{max})$
- Minimizing mean cycle time and variance of cycle time (Bagchi 1989) = $\min(\bar{F}, \sigma_{\bar{F}}^2)$

The motivation for using these two pairs of conflicting objectives is to simultaneously incorporate the manufacturer's as well as the customer's concerns in scheduling. For each pair, the interest lies not in just achieving a single optimal solution, but in defining a near-optimal trade-off curve so that the decision maker can operate the scheduling along the trade-off curve according to the dynamic needs of the manufacturing shop. Thus, there is a need to develop a Pareto control curves between these conflicting objectives.

From more than six decades, efforts have been invested by many researchers around the world towards the development of polynomial algorithms for solving complex scheduling problems. Exact algorithms are developed even for multiobjective scheduling problems as reviewed by Nagar, Haddock, and Heragu (1995) and T'kindt and Billaut (2002). However, most of these efforts are dedicated on finding pareto-optimal solutions for static scheduling problems, in which we assume that all jobs are available for scheduling at the start of the planning period, and their processing information, are known. The concept of pareto-optimal solutions has not yet been extended and applied to scheduling problems in dynamic cases, where jobs are assumed to arrive on continuous basis over time. The term "pareto-optimal solutions" as defined for static problems may not be applicable to scheduling problems studied in dynamic sense. In our earlier work (Ganesan, Gupta, and Sivakumar 2004), we introduced the concept of Pareto-control in dynamic scheduling, which is summarized herein.

Suppose there are two conflicting objective functions, F_1 and F_2 , in dynamic single machine scheduling. Assuming that these objectives are individually optimized by scheduling a job at each decision instance in simulation clock according to the corresponding criteria of $f_1(x)$ and $f_2(x)$, for $x \in X$, where x is the job index and X is the set of all available jobs at that particular decision instance. For biobjective scheduling, at each decision instance on the time horizon, a convex combination approach is used to define a joint criterion. This joint criterion is used to select a Pareto optimal job based on the relative importance of scheduling criteria $f_1(x)$ and $f_2(x)$. This results in a particular sequence of dynamically arriving jobs for scheduling and corresponding objective functions, F_1 and F_2 . Varying the relative importance of scheduling criteria at each decision instance will result in Pareto control of biobjective dynamic scheduling as shown in Figure 1.

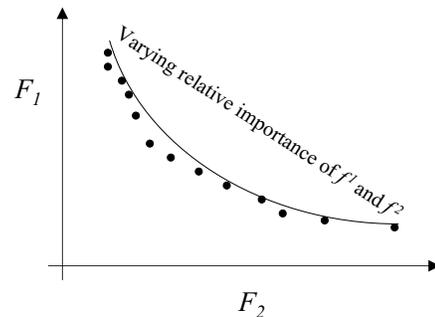


Figure 1: Pareto Control in Multiobjective Dynamic Scheduling

For scheduling dynamically arriving jobs on the machine, a discrete-event simulation-based approach of scheduling is used. In this approach, at each decision instance in the simulation clock, an optimal job is selected from the dynamic queue of jobs and scheduled on the machine for processing. The simulation clock is then moved forward to the next decision instance. Generally, scheduling and simulation are studied as two separate concepts; scheduling for sequencing jobs on the machine and simulation as a tool to validate the results. Here, the concepts of scheduling and simulation are combined and thus used to exploit the benefits of discrete-event simulation in dynamic scheduling of the shop. This combined approach is referred to as Conjunctive Simulated Scheduling (CSS) (Gupta and Sivakumar 2005a). Using CSS, a multiobjective dynamic scheduling problem is disintegrated into several decision instance based multiobjective optimization problems. Thus, an appropriate job, which satisfies multiple criteria of scheduling objectives need to be selected at each decision instance in simulation clock. Therefore, unlike finding a Pareto sequence in static scheduling problems, in multiobjective dynamic scheduling

problems a Pareto optimal job is selected for scheduling on the machine at each decision instance in simulation clock.

3 DETERMINATION OF PARETO OPTIMAL JOB

At each decision instance in dynamic scheduling, the Pareto optimal jobs are determined and loaded on the machine for processing. This results in a schedule of dynamic jobs, which gives a boundary solution for the dynamic biobjective scheduling problem (Gupta and Sivakumar 2002; Ganesan, Gupta, and Sivakumar 2004).

Various approaches exist to find the Pareto optimal solution for a multiobjective optimization problem as reported by Zeleny (1982), Tabucanon (1988), T'kindt and Billaut (2002). The most commonly used approach in multiobjective optimization is to transform the problem into single objective problem using convex combination of objectives and weights. The weights are decided based on the relative importance of each objective. In this research, the convex combination approach is used because of its simplicity of application in dynamic scheduling, low cost of computation, and ability to provide feasible solution in near-real-time.

Most of the convex combined functions are either in a linear fashion or in the form of some kind of the distance derivatives (Tabucanon 1988, Zeleny 1982). The prominent techniques in this relation are Weighted Aggregation (WA) and Compromise Programming (CP). In WA method, different objectives are weighted and summed up in a linear fashion to a single objective, formulated as: $\min F = \sum_{j=1}^2 w_j \cdot f_j(x)$, where w_j are non-negative weights with $\sum w_j = 1$. By varying these weights, all the Pareto optimal points can be achieved as each Pareto optimal solution point on a convex surface corresponds to a set of w_j (Geoffrion 1968). However, WA method suffers from the problems of (1) dimensional inconsistency among various objectives and (2) uneven spread of points on the Pareto set (Das and Dennis 1997). Motivated by the need of looking for a more powerful approach, Zeleny (1982) and others developed the CP method.

CP method identifies solutions which are closest to the utopia solution as determined by some measure of distance (Zeleny 1982, axioms of choice). The solutions identified as being closest to the utopia solution are called compromise solutions and constitute the compromise set. In the CP method, the point of interest is the comparison of distances of different efficient points ($f_j(x), j = 1, 2$) from the utopia solution which is the point of reference. Since the objectives may be of different dimensions, so the distance measure needs to be corrected to make the individual objectives mutually commensurable. It is therefore necessary to use relative rather than absolute deviations. The individual relative deviations can be raised to any power ($r > 0$) before these are summed and also the weights w_j ($0 < w_j < 1$ and $\sum_j w_j = 1$)

can be attached to the different relative deviations. For a multi-objective optimization problem, with utopia point $F^*(x^*) = [f_1(x^*), f_2(x^*)]$, the overall minimizing objective function can be expressed as follows:

$$\min F_r = \left[\sum_{j=1}^2 \left[w_j \cdot \frac{|f_j(x^*) - f_j(x)|}{f_j(x^*)} \right]^r \right]^{1/r} \quad (1)$$

In this formulation, the choice of exponent r reflects the user's concern with respect to the maximal deviation. Introduction of w_j allows the expression of the user's intuition concerning the relative importance of the various objectives. Thus, a double-weighting scheme exists. The parameter r reflects the importance of the maximal deviation and the parameter w_j reflects the relative importance of the j^{th} objective. From the application point of view, both these parameters give the benefit of controlling the generated solution over the Pareto boundary. In the above CP formulation, keeping the utopia point the same but by changing the weights, one may reach all the efficient points located on the boundary (Zeleny 1982). In the observations of Gupta and Sivakumar (2002), at any selection instance, all Pareto optimal jobs can be selected by the CP method by varying weights w_j within a particular spread of boundary. On the other hand, varying exponent r helps in controlling the spread of this boundary.

Operationally, three points of the compromise set are usually calculated, that is, those corresponding to $r = 1, 2$, and ∞ . When $r = 2$, the equation becomes simply the distance between two points, where each relative deviation is weighted in proportion to its magnitude. As r becomes larger and larger, the largest deviation receives more and more weight. For r approaching ∞ , the distance measure reduces to the Tchebycheff function (Zeleny 1982):

$$\lim_{r \rightarrow \infty} F_r = \max_{j \in \{1,2\}} \left[\left[w_j \cdot \frac{f_j(x^*) - f_j(x)}{f_j(x^*)} \right] \right] \quad (2)$$

This is because the relative contribution of the largest relative deviation when raised to a large exponent would be extremely larger than all the rest combined, and thus will dominate the distance determination.

Though the weights representing relative importance are used as the preference structure when applying CP, it has been mathematically proven that CP method is superior to WA method in locating the efficient solutions (Steuer 1986). Thus, CP method is widely applied in mechanical engineering design problems, where single stage multiobjective optimization problems exist. However, there are very few applications of CP to multiobjective scheduling problems. Taboun, Abib, and Atmani (1995) used CP to solve two biobjective scheduling problems with objectives

of (i) minimizing total tardiness and earliness *vs.* mean flowtime and (ii) minimizing total tardiness *vs.* total earliness. They modelled the scheduling problem using linear programming and used CP to generate efficient points of biobjective static scheduling problem having 13 jobs to schedule on a single machine. Koksalan (1999) exploited CP to develop a heuristic procedure for solving two different biobjective scheduling problems: (i) minimizing total flowtime and maximum tardiness and (ii) minimizing total flowtime and maximum earliness. Their application also addressed only static scheduling problems and presented an example of scheduling 15 static jobs on a single machine. However, to the best knowledge available, research on multiobjective dynamic scheduling has not been addressed yet in the existing literature.

Primarily, from the point of view of industrial application, it is important to develop a methodology for operating the shop in a user-controlled trade-off between the needs of conflicting scheduling objectives. The characterization of Pareto control curves for the conflicting objectives can be exploited to enhance the ability to choose and operate the shop in a user controlled way. By appropriate selection of parameters, the conflicting scheduling objectives can be controlled and maintained within a specified range for a particular operation. This Pareto control is highly relevant for the shop floor operations management.

In the following section, a real-life case study has been presented in order to show the effectiveness of Pareto control curves for the shop floor management. This case study shows the application of Pareto control in serial processing with respect to the two sets of conflicting objectives: (1) simultaneously minimizing mean cycle time and maximum tardiness and (2) simultaneously minimizing mean cycle time and cycle time variance.

4 CASE STUDY MODEL, DATA AND EXPERIMENTATION

In this case study, the concept of Pareto control in multi-objective dynamic scheduling has been applied to a stepper machine in semiconductor wafer fabrication. This case study is based on the real factory data from a wafer fabrication plant in Singapore.

The wafer fabrication process dominates the economics of IC production and it is the most technologically complex and capital-intensive stage in semiconductor manufacturing. In the wafer fab, the wafers are processed in order to build up layers of patterns to produce the required circuitry. This involves a complex sequence of processing steps with a number of operations that require different kinds of equipment.

Wafer fabrication process basically involves the following process steps: photolithography, etching, deposition, chemical mechanical polishing, ion implantation, diffusion.

Among these process steps, the photolithography process is considered the most crucial and bottleneck process. Due to the large number of photolithography steps needed in IC manufacturing, photolithography typically accounts for about 30% of the cost of manufacturing. Photolithography or optical lithography is basically a photographic process by which a light sensitive polymer, called a photoresist, is coated, exposed and developed to transfer a pattern from a photomask (also called reticle) to the surface of a wafer. The stepper machines are the pieces of equipment used for the photolithography process, which is the most expensive bottleneck resource in the wafer fabrication process.

The general sequence of processing steps for a typical photolithography process is as follows: substrate preparation, photoresist spin coat, prebake, exposure, development and postbake. This sequence is generally performed on several tools linked together into a contiguous unit called a lithographic cluster. The main processing steps in photolithography are coat, expose and develop. In the photolithographic cluster, there are two parallel coat tools, one expose tool and two parallel develop tools. The flow of these processing steps is presented in Figure 2. In this flow, the expose tool is a bottleneck as all the wafers have to necessarily pass through this single tool. This expose tool can be modeled as a single serial processing machine and the concept of Pareto control in serial processing can be applied in order to control the trade-off in multiobjective dynamic scheduling.

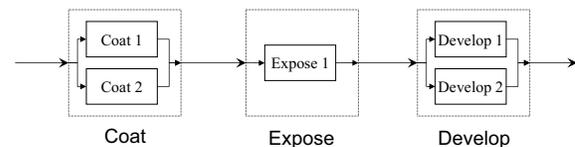


Figure 2: Flow of Processing Steps in a Photolithographic Cluster

A sample of the factory data is presented in Table 1. In this table, second column represents the Lot ID, third column gives the step begin time, which is basically the lot arrival time at the photolithographic cluster, fourth column shows the priority of the lot, fifth column gives the turn ratio of the lot, sixth column gives the lot quantity in terms of the number of wafers in the lot and finally seventh column represents the expose time per wafer in the lot. For the confidentiality reasons, the Lot IDs are changed as shown in the second column of Table 1.

The factory data is available for a period of two months, which involves the dynamic arrival of total 2508 lots. One requirement for applying the concept of Pareto control at the expose tool is the data for the arrival time at the expose tool, but this data is not available in the lot tracking system of the factory. In order to make the situation suitable to the

Table 1: Sample Presentation of the Factory Data

S.No.	Lot ID	Step Begin Time	Lot Priority	Turn Ratio	Current Lot Quantity	Expose Time (sec/wafer)
1	XYZ001	30/11/2005 15:52	4	6.37	1	78.5414
2	XYZ002	30/11/2005 20:26	4	0	25	72.2
3	XYZ003	30/11/2005 20:27	4	0	25	72.2
4	ABC001	30/11/2005 20:27	4	0	25	72.2
5	ABC002	30/11/2005 20:28	4	0	25	72.2
6	ABC003	30/11/2005 20:35	4	0	25	72.2
7	ABC004	30/11/2005 20:35	4	0	25	72.2
8	ABC005	30/11/2005 20:36	4	0	25	72.2
9	XYZ004	30/11/2005 20:36	4	0	25	72.2
10	ABC006	30/11/2005 20:36	4	0	25	72.2
11	XYZ005	30/11/2005 21:23	4	0	25	72.2
12	XYZ006	30/11/2005 21:24	4	0	25	72.2
13	XYZ007	30/11/2005 21:24	4	0	25	72.2
14	XYZ008	30/11/2005 21:24	4	0	25	72.2
15	XYZ009	30/11/2005 21:25	4	0	25	72.2
16	PQR001	30/11/2005 22:47	4	0	25	72.2
17	PQR002	01/12/2005 1:45	3	7.61	1	72.2
18	PQR003	01/12/2005 4:05	4	0	25	72.2
19	ABC007	01/12/2005 4:05	4	0	25	72.2
20	ABC008	01/12/2005 4:06	4	0	25	72.2
-	-	-	-	-	-	-
-	-	-	-	-	-	-
-	-	-	-	-	-	-
2508	PQR--	31/01/2006 19:29	2	6.99	25	80.1524

model, the lot arrival time at the photolithographic cluster is assumed to be the lot arrival time at the expose tool. Then, based on the advice of the industrial engineering experts at the factory, the expose processing time is multiplied with a random factor between 1 and 1.3 in order to balance the workload at the expose tool. Thus, the lot processing time (p_i) at the expose tool is computed according to the following equation:

$$\text{lot processing time } (p_i) = (1 + 0.3 * U(0,1)) * \text{expose time per wafer} * \text{lot quantity}$$

where, $U(0,1)$ is the uniform random number between 0 and 1.

Another requirement for applying the concept of Pareto control is the operation due date (d_i) for each lot at the expose tool, which is computed based on the following equation:

$$\text{operation due date } (d_i) = \text{arrival time} + (\text{lot priority} * \text{lot processing time}).$$

The actual operating heuristic in factory is priority minimization followed by turn ratio maximization followed by first arrival, which is represented as factory heuristic. Other simple dispatching rules considered in this study are FCFS, SPT and EDD. The simulation experiments on the factory data are conducted for two cases: first with just 1-month data and second with 2-month data.

5 CASE STUDY RESULTS AND DISCUSSION

The concept of Pareto control in serial processing is applied on the expose tool for the two pairs of conflicting objectives:

(1) mean cycle time and maximum tardiness and (2) mean cycle time and cycle time variance. The relative weights ($w_1, w_2; w_1 + w_2 = 1$) of the two criteria in the CP method are varied linearly from 0.05 to 0.95 at an interval of 0.05, computing total 19 points. The exponent r is varied at the levels of 0.33, 0.66, 1, 2, 3 and 10. The CP curves for these exponent settings are correspondingly represented by CP0.33, CP0.66, CP1, CP2, CP3 and CP10. Among these, CP0.33 and CP0.66 are studied to see the effects of exponent value less than unity and similarly CP2 and CP3 for the effects of exponent value greater than unity. However, CP10 is studied as an approximation of Tchebycheff function, as earlier explained in Equation (2) (Koksalan 1999).

5.1 Minimizing Mean Cycle Time (\bar{F}) and Maximum Tardiness (T_{max})

For a serial processing machine, the mean cycle time is minimized by the SPT rule and the maximum tardiness is minimized by the EDD rule. Thus, the two criteria of CP method are $f_1(x) = \text{minimize } (p_i)$ and $f_2(x) = \text{minimize } (d_i)$.

Thus, based on Equation (1), the Pareto objective function for the CP method is formulated:

$$\min F_r(i) = \left[\left(w \cdot \frac{|p_i - \min(p_i)|}{\min(p_i)} \right)^r + \left((1-w) \cdot \frac{|d_i - \min(d_i)|}{\min(d_i)} \right)^r \right]^{1/r} \quad (3)$$

Here, $F_r(i)$ represents the CP function value for the i^{th} job at the decision instance in the simulated time.

Then, among all the queued jobs in front of the machine, the best job is selected for loading on the serial processing machine. The job, which has the minimum value of Pareto objective function, $F_r(i)$, is selected for loading at every decision instance in the simulation clock. Thus, the k^{th} job is loaded on the machine for processing, where $k = \text{argmin}_{\forall i} [F_r(i)]$.

The Pareto curves for this pair of conflicting objectives are presented in Figures 3 (a) and (b), which respectively show the Pareto curves for 1-month and 2-month factory data. These plots show the characteristic curves of Pareto control for varying weights and exponent values in CP method. In these plots, the dispatching criterion, $\min(p_i)$, gives the extreme point corresponding to minimization of mean cycle time and similarly dispatching criterion $\min(d_i)$ gives the other extreme point corresponding to minimization of maximum tardiness. These two extreme points are the limits of Pareto control achievement, which are correspondingly referred to as SPT and EDD points.

5.2 Minimizing Mean Cycle Time (\bar{F}) and Cycle Time Variance (σ_F^2)

For this pair of conflicting objectives, the mean cycle time is minimized by the SPT rule and the cycle time variance is assumed to be minimized by the P_W rule based on simulation results. These two criteria of CP method are represented as $f_1(x) = \text{minimize}(p_i)$ and $f_2(x) = \text{minimize}(p_i/W_i)$, where W_i is the waiting time of the lot in front of the expose tool.

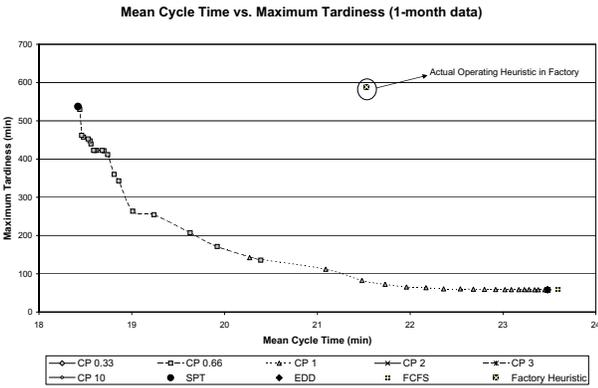
Thus, the Pareto objective function is formulated for the CP method based on Equation (1):

$$\min F_r(i) = \left[\left(w \cdot \frac{|p_i - \min(p_i)|}{\min(p_i)} \right)^r + \left((1-w) \cdot \frac{|\frac{p_i}{W_i} - \min(\frac{p_i}{W_i})|}{\min(\frac{p_i}{W_i})} \right)^r \right]^{1/r} \quad (4)$$

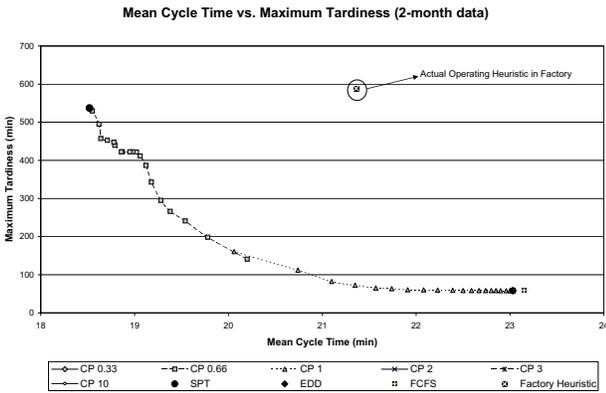
Then, among all the queued jobs in front of the machine, the job with the minimum value of Pareto objective function, $F_r(i)$, is selected for loading on the serial processing machine.

The Pareto curves for this pair of conflicting objectives are presented in Figures 4 (a) and (b) for 1-month and 2-month factory data. In these figures, the extreme points are achieved by the SPT and P_W rules. In Figures 4 (a) and (b), the points of CP10 curve give a wide spread of boundary between the SPT and P_W points. For CP3, the spread of points gets reduced. Further, as exponent value in CP method decreases, the spread of points gets reduced more and more. The points corresponding to $w = 0.05$ for various CP curves remain in the neighborhood of P_W point. On the other hand, the points corresponding to $w = 0.95$ for various CP curves move far away from SPT point as the exponent value decreases. Thus, the range of the achieved boundary gets reduced with the decrease in the exponent value of CP.

Under the given model and assumptions, the points corresponding to FCFS and the factory heuristic are far from the Pareto control curves in all these figures, Figure 3 (a) and (b) and Figure 4 (a) and (b). This observation reflects the importance of using the concept of Pareto control in order to trade-off the scheduling performance between the conflicting objectives. The characterization of Pareto control curves gives the ability to control the scheduling objectives within a specified range. Thus, the Pareto control curves can be used as a tool for the shop floor control management. By selecting an appropriate weight and exponent value, the decision maker can control the machine scheduling within the specific range of the scheduling objectives.



(a) Pareto Control Curve with 1-Month Data

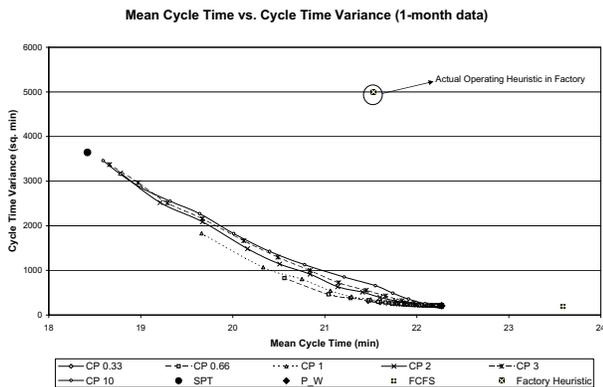


(b) Pareto Control Curve with 2-Month Data

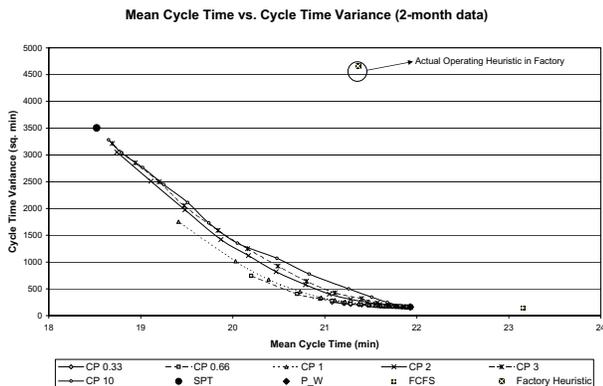
Figure 3: Case Study: Pareto Curves for Mean Cycle Time and Maximum Tardiness

In CP method, as the weight is varied linearly in a discrete fashion, different points are achieved on the boundary in a progressive manner from one end point to the other end point of the curve. CP method has a double weighing scheme which gives it flexibility to cover the whole boundary from one end point to the other end point. In addition, the variation of exponent value in CP method gives it the higher flexibility of Pareto control between a specified range of objective functions.

In Figure 3, all the points of CP0.33 coincide with the SPT point. For CP0.66 and CP1, the points get spread over a wider range and the curve shifts towards EDD point. For CP2, CP3 and CP10, all the points coincide with the EDD point.



(a) Pareto Control Curve with 1-Month Data



(b) Pareto Control Curve with 2-Month Data

Figure 4: Case Study: Pareto Curves for Mean Cycle Time and Cycle Time Variance

6 CONCLUSIONS AND FUTURE WORK

This paper presented the concept of Pareto-control in multi-objective dynamic scheduling of a single serial processing machine. A case study based on the data of a stepper machine in a wafer fab plant demonstrated the relevance of Pareto control in industrial application. The study showed that by varying the relative weights of selection criteria at each selection instance in simulation clock, we can achieve a Pareto-control curve for two conflicting objectives. CP method is used to determine Pareto jobs at every decision instance in the simulation clock. Using conjunctive simulated scheduling, a Pareto job is loaded on the machine for processing at each decision instance in simulated time.

The results obtained from Pareto control approach are superior to the simulated results of actual operating heuristic in the factory for both the pairs of conflicting objectives: (1) mean cycle time and maximum tardiness and (2) mean cycle time and cycle time variance. Also, it is demonstrated that by using the Pareto control approach, the stepper machine can be controlled at specified needs of scheduling objectives. Thus, the Pareto control curves can be used as a tool for the shop floor control management.

Further research can explore the application of the concept of Pareto control to the scheduling problems in serial and batch processing environments for the cases with more than two objectives. Pareto control concept also needs to be extended to scheduling problems with sequence dependent setup times in serial processing machines. This will lead to the development of a Pareto surface in 3-D space and thereafter controlling the schedule along the surface. Further analysis is also required on the mathematical understanding of the CP method for finding the correlation between the weight and the exponent, which will enhance the control aspect of Pareto curves between the conflicting objectives.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Lieu Sip Khean, Madan Mohan Chakravarthi, and Tan Hiap Keong from the Industrial Engineering Department of Chartered Semiconductor Manufacturing Ltd., Singapore, for providing the opportunity and support to build the real-life case study and test the concepts. This research work is partially supported by Singapore-MIT Alliance (SMA).

REFERENCES

- Bagchi, U. 1989. Simultaneous minimization of mean and variation of flow time and waiting time in single machine systems. *Operations Research* 37:118–125.
- Dabbas, R., H.-N. Chen, J. Fowler, and D. Shunk. 2001. A combined dispatching criteria approach to scheduling semiconductor manufacturing systems. *Computers and Industrial Engineering* 39:307–24.
- Das, I., and J. Dennis. 1997. A closer look at the drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. *Structural Optimization* 14 (1): 63–69.
- Ganesan, V., A. Gupta, and A. Sivakumar. 2004. Dynamic bi-objective Pareto-schedule generation and control in semiconductor supply chain. In *IEEE International Engineering Management Conference*, 1212–1216. Singapore, Oct 18–21.
- Geoffrion, A. 1968. Proper efficiency and the theory of vector optimization. *Journal of Mathematical Analysis and Application* 41:491–502.

- Gupta, A., and A. Sivakumar. 2002. Simulation based multi-objective schedule optimization in semiconductor manufacturing. In *Proceedings of the Winter Simulation Conference*, 1862–1870.
- Gupta, A., and A. Sivakumar. 2005a. Conjunctive simulated scheduling. *International Journal of Advanced Manufacturing Technology* 26 (11): 1409–1413.
- Gupta, A., and A. Sivakumar. 2005b. Multi-objective scheduling of two-job families on a single machine. *OMEGA: The International Journal of Management Science* 33:399–405.
- Koksalan, M. 1999. A heuristic approach to bicriteria scheduling. *Naval Research Logistics* 46:777–789.
- Nagar, A., J. Haddock, and S. Heragu. 1995. Multiple and bicriteria scheduling: a literature survey. *European Journal of Operational Research* 81:88–104.
- Pinedo, M., and X. Chao. 1999. *Operations scheduling with applications in manufacturing and services*. Singapore: Irwin McGraw-Hill.
- Steuer, R. E. 1986. *Multiple criteria optimization: theory, computation, and application*. Malabar, Florida: Krieger Publishing Company.
- Taboun, S., A. Abib, and A. Atmani. 1995. Generating efficient points of bicriteria scheduling problem by using compromise programming. *Computers and Industrial Engineering* 29 (1–4): 227–231.
- Tabucanon, M. 1988. *Multiple criteria decision making in industry*. Amsterdam, The Netherlands: Elsevier Science Publishers B.V.
- T'kindt, V., and J.-C. Billaut. 2002. *Multicriteria scheduling: theory, models and algorithms*. Heidelberg: Springer-Verlag.
- VanWassenhove, L., and L. Gelders. 1980. Solving a bicriterion scheduling problem. *European Journal of Operational Research* 4 (1): 42–48.
- Yang, J., and T.-S. Chang. 1998. Multiobjective scheduling for IC sort and test with a simulation testbed. *IEEE Transaction on Semiconductor Manufacturing* 11 (2): 304–315.
- Zeleny, M. 1982. *Multiple criteria decision making*. New York: McGraw-Hill Book Company.

AUTHOR BIOGRAPHIES

AMIT KUMAR GUPTA is a Research Engineer with the Planning and Operations Management group at the Singapore Institute of Manufacturing Technology (SIMTech), Singapore. He completed his Ph.D. studies in the School of Mechanical and Aerospace Engineering at Nanyang Technological University (NTU), Singapore. For his doctoral study, he received the Singapore Millennium Foundation Scholarship (2002) award. He obtained his bachelor degree in Mechanical Engineering from the Indian Institute of Technology (IIT), Delhi. His research interests are Operations Research

and Management, Production Scheduling, Simulation Modeling of Manufacturing Systems and Multi-objective Dynamic Scheduling in Semiconductor Manufacturing. He has been the Proceedings Chair for the 3rd International Conference on Modeling and Analysis of Semiconductor Manufacturing (MASM 2005), Singapore. His e-mail address is <amitkg@SIMTech.a-star.edu.sg>.

APPA IYER SIVAKUMAR is an Associate Professor in the School of Mechanical and Aerospace Engineering at the Nanyang Technological University, Singapore and Fellow of Singapore - Massachusetts Institute of Technology (MIT) Alliance (SMA). He was at Gintic Institute of Manufacturing Technology, Singapore prior to this appointment. His research interests are in the area of OR, Manufacturing Systems, Logistics, Supply chain, Simulation, Scheduling and Research Methodology. He held various management positions including technical manager and project manager for nine years at Lucas Systems and Engineering and Lucas Automotive, UK. He received a Bachelors of Engineering from University of Bradford, UK and a PhD in Manufacturing Systems Engineering from University of Bradford, UK. He has made many contributions at international conferences and journals. He has trained and supervised a number of PhD students and many Master's level students. He has been the technical chair and co-edited the proceedings of the 3rd and 4th International Conference on Computer Integrated Manufacturing (ICCIM '95 and ICCIM'97), Singapore. His e-mail address is <msiva@ntu.edu.sg>.