Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

MODELING THE IMPACT OF NEW PRODUCT INTRODUCTION ON THE OUTPUT OF SEMICONDUCTOR WAFER FABRICATION FACILITIES

Atchyuta Bharadwaj Manda
Reha Uzsoy

Edward P. Fitts Department of Industrial and Systems Engineering
North Carolina State University
Campus Box 7906
Raleigh, NC 27695-7906, USA

Karl G. Kempf

Decision Engineering Group
Intel Corporation
MS CH3-10
5000 W. Chandler Blvd.
Chandler, AZ 85226, USA

Sukgon Kim

Supply Chain Group
Maxim Integrated
160 Rio Robles
San Jose, CA 95135, USA

ABSTRACT

We consider the problem of managing output in semiconductor wafer fabrication facilities when a new product is introduced alongside a current product. We propose a mathematical model of the impact of the new product on the distribution of the effective processing time, and use a simple Excel simulation to illustrate the impact of different release control policies on output under product transitions.

1 INTRODUCTION

The introduction of new products is an important basis for competition in many industries (Levinthal and Purohit 1989; Padmanabhan et al. 1997). A new product often competes with earlier ones for customers in the marketplace, production and distribution capacity in the supply chain, and resources in the development organization. An unsuccessful product transition, or rollover, can severely affect a firm's performance, sometimes driving it out of business entirely (Billington et al. 1998; Erhun et al. 2007; Bilginer and Erhun 2010). The introduction of new manufacturing processes is critical to the introduction of new products, but disrupts production and product development. We will refer to product and process transitions collectively as *technology transitions*. Decisions are made under market uncertainties in the demand for new products, and hence processes; and technological uncertainties in the time and resources required to bring products and processes into high volume manufacturing.

The literature has treated management of technology transitions as a strategic problem, viewing them as infrequent events that can be managed as an exception to routine operations and planning, and in isolation from product lines not involved in the transition. In contrast, high tech industries such as semiconductor manufacturing face significant adverse impact on plant productivity from technology transitions (Leachman 1996; Gopal et al. 2011) which must be considered in operational decision making, with frequent product transitions being the norm rather than the exception. In these transitions a new device is introduced into a fab where it shares equipment capacity with previously introduced devices currently in demand in the market.

The adverse impact of new product introductions on the performance of semiconductor wafer fabrication facilities (fabs) is widely acknowledged. In an extensive study of competitive practices in the industry, Leachman (1996) identifies effective management of new product introductions as a major contributor to overall fab productivity. In this paper we focus on the operational problem of managing the release of wafers into the fab during a product transition, where an older product is being replaced in the market by a newer one. In general, newer products have higher profit margins than older ones, due either to reduced manufacturing cost achieved through die shrinkage, or additional features that command a price premium. Since the market price of new devices generally decreases rapidly over time as competitors introduce competing devices, maximizing the output of the new device early in the product lifecycle has important implications for revenue (Leachman 1996; Akcali et al. 2000; Leachman and Ding 2007). However, the transition from the older product to the newer one must proceed gradually for several reasons. First, the manufacturer's customers will not all adopt the new device simultaneously, but rather switch at different points in time as they exhaust inventories of the older product, incorporate the new device into their product designs, and observe the performance of the new device in the market. Hence demand for the new device will increase gradually over time. Second, introduction of the new device into the fab alongside the older device may have significant adverse effect on the output of both devices due to their sharing capacity. Thus increasing the starts of the new device too rapidly may result in excessive output of the new device early in its life cycle where its price is high but its demand is low, while reducing output of the older device that is still profitable and makes up most of the firm's demand. Thus the manner in which the mix of wafers of the two devices released into the fab is managed over time is critical to effective management of revenue during the transition.

In this paper we propose a simple probabilistic model of the interaction between old and new devices based on the effective processing time approach suggested by Hopp and Spearman (2008). In contrast to previous models of learning in semiconductor manufacturing, which have generally focused on the improvement of yield and throughput for either an individual device or a number of devices produced using the same process, this model explicitly captures the adverse impact of the new device on older devices whose learning has already taken place. We then use this model to simulate the impact of different release management policies to examine their impact on fab performance. We find that release policies based on maintaining a constant WIP level in the fab, such as CONWIP (Spearman et al. 1990) and Drum-Buffer-Rope (Goldratt and Fox 1986), may result in reduced output of both old and new devices during the transition. This suggests the need for optimization of the releases over time to maximize revenue while also meeting existing demand commitments.

In the following section we briefly review previous related work on the management of new product introductions and learning in manufacturing systems. Section 3 presents the model, and Section 4 implements it in a simple simulation of an aggregated wafer fab which must introduce a new device while ramping down another. Section 5 concludes the paper with a summary of the insights from this initial work and directions for future research.

2 PREVIOUS RELATED WORK

In this section we briefly review several streams of research related to the introduction of new products into manufacturing facilities.

2.1 Modeling Product Transitions

A growing body of research addresses the management of product transitions (Billington et al. 1998; Lim and Tang 2006; Erhun et al. 2007; Bilginer and Erhun 2010), examining different aspects of the problem. These models all study a centralized firm using aggregate formulations that do not consider the complex technical constraints faced by different units. They also do not consider the collateral impact of product transitions on products that are not involved in the transition, but share development or manufacturing resources with the transitioning product. In industrial practice, Intel's *Copy Exactly* technology transfer method (McDonald 1998) and the Tick-Tock cadence in which new device architectures and silicon compaction alternate have the effect of reducing the number of uncertainties to which a factory introducing a new product or process is exposed (Shenoy and Daniel 2006). Leachman (1996) presents industrial data supporting the effectiveness of this policy and the benefits of careful coordination between product development and manufacturing organizations in product transitions.

2.2 Manufacturing Systems with Learning

The initial effect of a product or process transition is usually a significant reduction in the throughput of the facility. Queueing models (Buzacott and Shanthikumar 1993; Hopp and Spearman 2008) and industrial experience both suggest that in multiproduct facilities, not only the products undergoing the technology transition but others sharing resources with them will be adversely affected. Over time, as experience with the new product or process accumulates, performance improves. Much existing work (Liao 1979; Reeves and Sweigart 1981; Hiller and Shapiro 1986) does not distinguish between production and engineering activity, but assumes that the resources required to produce a unit of good product will decline over time as a function of cumulative production. Later authors (Fine 1986; Fine 1988; Chand et al. 1996; Terwiesch and Bohn 2001) consider learning as a process of experimentation that uses production capacity to improve process capability. Leachman (1996) gives industrial data indicating that learning by doing in semiconductor manufacturing can be increased by allocation of engineering resources and organized problem-solving activities that reduce defect densities.

3 IMPACT OF NEW PRODUCT INTRODUCTIONS ON FAB THROUGHPUT

3.1 Modeling and Analysis

To motivate our approach, we shall treat the overall wafer fab as a single resource whose output can be modeled using a *clearing function* (Graves 1986; Karmarkar 1989). A clearing function (CF) represents the relation between the expected output of the production system it represents in a planning period as a function of some measure of the expected workload during that period. In this paper we follow previous production planning research (Kacar and Uzsoy 2010; Kacar et al. 2012; Kacar et al. 2013; Kacar et al. 2016) in using the expected workload, given by $\Lambda_t = W_{t-1} + R_t$, where W_{t-1} denotes the amount of work in process available to the resource at the start of planning period t, and R_t the amount of work planned to be released to the resource during period t. If we treat the fab as a single resource that can be modelled as an *M/G/1* queue, Missbauer (2002) has shown that the expected output X_t in period t, expressed in units of time, can be written as

$$\Theta_t = \frac{1}{2}\left[\Delta + k_t + \Lambda_t - \sqrt{\Delta^2 + 2\Delta k_t + k_t^2 - 2\Delta\Lambda_t + 2k_t\Lambda_t + \Lambda_t^2}\right]. \tag{1}$$

where Δ denotes the length of the planning period and

$$k_t = \frac{\sigma_t^2}{2a_t} + \frac{a_t}{2}, \tag{2}$$

where a_t and σ_t^2 denote the mean and variance of the effective processing time in period t, respectively. The expected number of units produced in a period will then be

$$X_t = \frac{\Theta_t}{a_t} \tag{3}$$

Thus we can model the effect of new product introductions on the expected throughput of the fab through changes in the distribution of the effective processing time over time, represented by its mean a_t

and its variance σ_t^2. This will result in a new clearing function for each planning period t, defining the expected output of the system in that period as a function of the workload Λ_t available to the system in that period. Kim and Uzsoy (2008; 2013) use a similar approach to represent the impact of learning over the product life cycle on the expected throughput of a production system. However, that work examined the impact of learning on process improvement benefiting all products in the system in a similar manner. This paper examines the impact of product mix on effective processing times, where the two products have different processing characteristics.

The objective of this work is to model a system into which a new product is introduced, changing the distribution of the effective processing time. We assume that all wafer lots of different products have a common underlying distribution of their natural processing time, which represents the behavior of the shared manufacturing technology. Based on our industrial observations as well as studies of industrial practice (Leachman 1996), we assume that differences between products arise from the disruptions they induce in the fab. Specifically, the impact of a new product on factory throughput is due to the increased level of engineering intervention it requires. Early in its life cycle, a new product will require frequent engineering interventions where problems are encountered performing a specific operation on a particular piece of equipment for one or more lots of the new product. This results in the equipment being placed on engineering hold, making it unavailable for regular production until manufacturing engineers have identified and resolved the issue. We will represent these engineering holds as non-preemptive disruptions that reduce the availability of the equipment and thus alter the distribution of the effective processing time. Specifically, when a new product is first introduced, it will require frequent engineering interventions, resulting in an abrupt increase in the mean and variance of the effective processing time. However, as manufacturing engineers identify and resolve problems with the new product, the frequency of engineering interventions will decrease over time, reducing the mean and variance of the effective processing time as continuous improvement activities are implemented.

We hasten to point out that engineering holds are by no means the only impact of the new product introduction. Individual lots of both old and new products, as opposed to equipment, may be placed on hold, altering the effective processing time distribution. Setups, equipment failures and yield excursions may also differ between the two products, again altering the effective processing time distribution. In general it is likely that the distributions of these parameters will have higher means and variances for the new product relative to the current one, and hence their impact on the effective processing time should be similar to that of engineering holds.

Let t_0 denote the expected natural processing time of a lot of wafers, and σ_0 its variance. Let the system encounter an engineering hold with mean duration P_t and variance $\sigma_P^2(t)$ on average once every Q_t lots in period t. Then the mean and variance of the effective processing time are given by $t_e = t_0 + P_t/Q_t$ and $\sigma_e^2(t) = \sigma_0^2 + \dfrac{\sigma_P^2(t)}{Q_t} + \dfrac{(Q_t - 1)}{Q_t^2}$, respectively (Hopp and Spearman 2008). The impact of new product introductions on output can then be captured through sequences P_t, $\sigma_P^2(t)$ and Q_t, $t = 1,...,T$ that determine the CF describing the output of the factory over time. Initially we will treat these sequences as scenarios exogenous to the planning models, yielding a CF of the form (1) with $a_t = t_0 + P_t/Q_t$ and

$$c_t^2 = \frac{1}{2}\left[c_a^2(t) + t_0^2 + \frac{\sigma_P^2(t)}{Q_t} + \frac{(Q_t - 1)}{Q_t^2}P_t^2\right] \tag{4}$$

where $c_a^2(t)$ denotes the squared coefficient of variation of the interarrival times, which we assume to be 1 in this paper. For our initial analysis we assume that the distribution of the time required to address a process problem, represented by P_t and $\sigma_P^2(t)$, is independent of the product causing them; the impact of new product introductions is isolated in the expected frequency $1/Q_t$ of engineering holds, which depends on the cumulative production of each product to date. Let $Q_{it}(X_{it})$ denote the mean number of lots between engineering holds for a product *i* whose cumulative production up to this time is X_{it} units. Let Q_s denote the expected number of lots between engineering holds for a stable product whose process is completely debugged. A model of the learning as a function of cumulative production is then given by

$$Q_{it}(X_{it}) = Q_i(0) + (Q_s - Q_i(0))e^{-\alpha X_{it}} \tag{5}$$

where α is a parameter determining the rate of improvement in the system, i.e., the rate at which problems with the new product are discovered and eliminated. We estimate the average number of lots between engineering holds for a system with N products as the weighted average of the mix of products making up the current workload Λ_t, given by

$$Q_t = \sum_{i=1}^{N} \left[\frac{Q_{it}\Lambda_{it}}{\sum_{i=1}^{N} \Lambda_{it}} \right] \tag{6}$$

The following section presents a series of simple simulation experiments that examine the implications of this model for the management of wafer fabs in the presence of product transitions, particularly for different release control policies that may be viable under different situations.

4 SIMULATION EXPERIMENTS

In order to illustrate the performance of the model proposed above, we implement the model in a system dynamics simulation of a wafer fab represented with a single aggregate clearing function. We consider a sequence of discrete time periods t = 1,...,T, and two products whose natural processing time and initial engineering holds (non-preemptive disruptions) occur following the distributions in Table 1. We consider a planning period of three months (129,600 minutes), due to the substantial amount of time involved in identifying problems with a new product, implementing remedial actions and observing results. Hence the system can produce an average of 1524 lots in each period.

Table 1: Product data for simulation examples.

	Parameter	Value
Natural Processing Time	Mean t_0	85 mins.
	Std. Dev. σ_0	68 mins.
	Coeff. of Variation c_0	0.8
Engineering Hold	Mean Duration P	900 mins.
	Std. Dev. σ_P	720 mins.
	Coeff. of Variation c_P	0.8

We assume that engineering holds for each of the two products occur following a Poisson process, with the mean number of lots between consecutive holds for Product 1, the older, stable product, given by Q_1 = 50 lots. Engineering holds for Product 2, the new product that is to be debugged, initially occur with a mean frequency of once every $Q_2(0)$ = 10 lots, eventually improving to $Q_2(\infty) = Q_s$ = 50 lots. In all experiments we fix the value of the improvement rate parameter to α = 0.0001 and simulate for a horizon of 65 periods. This results in the mean number of lots between engineering holds for Product 2, $Q_2(t)$, evolving over time as shown in Figure 1.

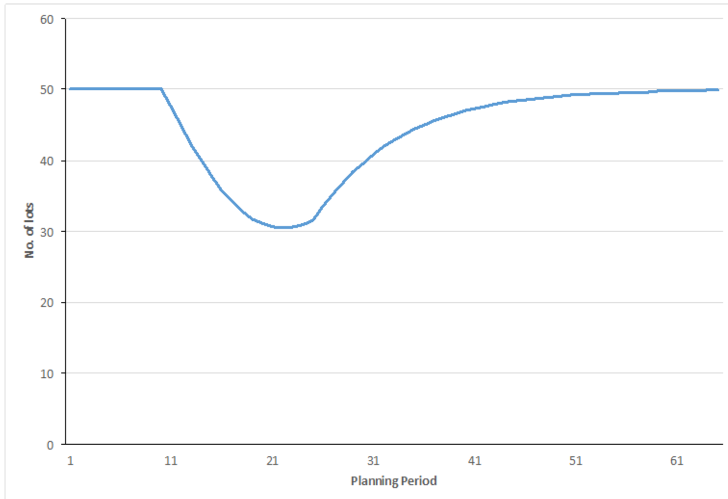

Figure 1: Evolution of number of lots between engineering holds.

The simulation is implemented in Microsoft Excel with no random variables. The number of unprocessed lots W_{it} of each product i in the system at the end of period t is given by the balance equation

$$W_{it} = W_{i,t-1} + R_{it} - X_{it} \tag{7}$$

where R_{it} denotes the number of lots of product i released in period t and X_{it} the output of product i in period t. The total workload of the system at the start if each period is given by

$$\Lambda_t = \Lambda_{1t} + \Lambda_{2t} = (W_{1,t-1} + R_{1t}) + (W_{2,t-1} + R_{2t}) \tag{8}$$

i.e., the total number of lots of both products available in WIP at the end of the previous period and the total planned releases of both products. The clearing function (1) is then used to compute the total output X_t of the system in period t in units of lots, assuming that the mix of output matches the mix of workload.

4.1 Base Experiment

In our base experiment we adopt an intuitive but non-optimized release policy of the type described to us by several industrial collaborators. For the first ten periods Product 1 is released at the rate of 1200 lots per period, giving a utilization of approximately 0.93. Starting in period 11, the releases of Product 1 are reduced by 7% (approximately 78 lots) in each period and replaced by a similar number of Product 2 lots such that the total releases into the system remain constant, with no restriction on the amount of WIP that may accumulate in the line. This profile under which the releases of Product 1 are replaced by releases of Product 2 between periods 11 and 25 defines the transition profile between the two products. The results of this run are shown in Figures 2 and 3.

Figure 2 shows that once lots of Product 2 begin to enter the fab in period 11, output begins to decline at an increasing rate due to the increased mean and variance of the effective processing time caused by the engineering holds on Product 2. After some time the rate of decrease begins to slow as manufacturing engineering reduces the frequency of engineering holds for Product 2. Eventually, in period 23, the engineering improvements overcome the negative effect of the additional engineering holds to the point that additional releases of Product 2 increase output rather than reducing it. Fab output increases past the initial level of 1200 lots, due to the accumulation of WIP in the fab caused by the increased effective processing time. Eventually the distribution of the effective processing time for the fab returns to the original value, as Q_2 reaches its final value of Q_s = 50, attaining the same level of engineering control as Product 1.

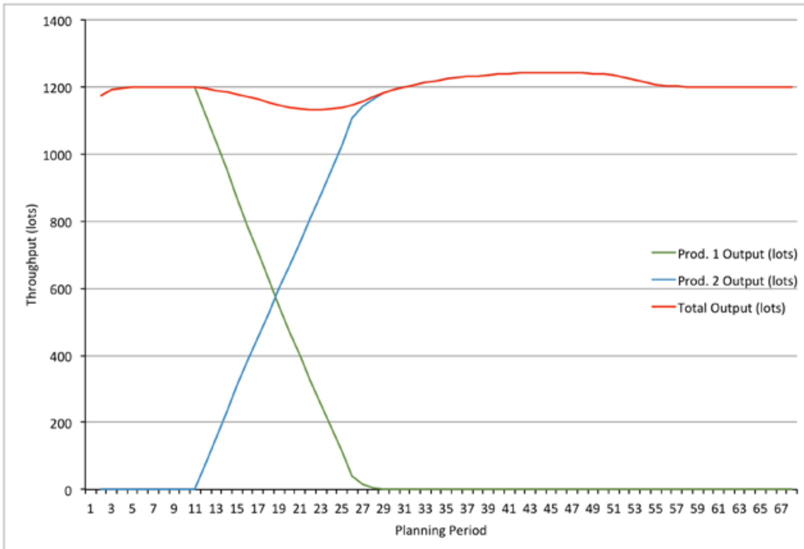

Figure 2: Evolution of fab output over time in baseline experiment.

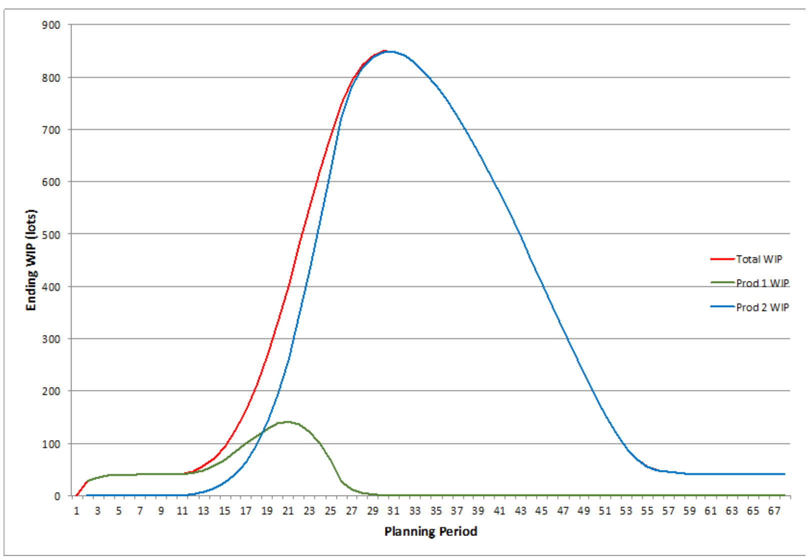

Figure 3: Evolution of WIP over time in baseline experiment.

It is interesting to note that although fab throughput initially decreases well below the initial level of 1200 lots per period with the introduction of Product 2, between periods 11 and 29, it eventually exceeds it quite significantly for an extended period of time, between periods 29 and 55. This behavior is due to the accumulated WIP in the fab raising its throughput by reducing idle time at bottleneck resources. Hence if cycle time is not a major issue, the total revenue from both products delivered by the end of the simulation may be satisfactory.

4.2 Controlled Release Policies - CONWIP

The behavior of the WIP over time is shown in Figure 3, which would cause concern to most fab managers - a very significant accumulation of WIP takes place between periods 11 and 30, which is not completely eliminated until period 55. By Little's Law (Hopp and Spearman 2008), this temporary increase in WIP implies a major increase in cycle time, with potential adverse consequences for delivery performance of both products. This is caused by the increase in both the mean and variance of the effective processing time in the fab due to the introduction of Product 2 and the increased engineering hold episodes it causes. In fact, the utilization of the system is very close to 1 between periods 13 and 51 for this reason. Most current fabs are managed with release policies that try to maintain stable cycle times by monitoring the current WIP level in the fab. Examples of such policies are the Workload Regulating policy(Wein 1988; Lu et al. 1994), the Bottleneck Starvation Avoidance policy (Glassey and Resende 1988), CONWIP (Spearman et al. 1989; Spearman et al. 1990) and Drum-Buffer-Rope (Goldratt and Fox 1986). In this paper we implement the CONWIP policy of Spearman et al. (1990) due to its simplicity and the fact that it captures the basic logic of several such approaches: to maintain constant average cycle times by maintaining constant WIP levels as far as possible.

Under the CONWIP system, a target workload level for the system is specified in units of time, representing the maximum workload at the bottleneck resource allowed to enter the fab. This target workload level is held constant by releasing a number of lots into the fab whose workload is exactly equal to that of the lots completed in the previous period. To accomplish this, the total workload, in units of time, whose processing was completed in the previous period is calculated to obtain the workload of new lots that can be released into the fab in the next period. The mix of lots of each product to enter the fab is determined based on the transition trajectory, the effective processing time parameters for the next period are calculated, and the system output in time units determined by the clearing function (1). In this experiment we explore the behavior of the system under this policy where we set a target workload of 123,600 minutes, corresponding to releasing 1173 lots of Product 1 in each period, and allow the system to reach steady state. Starting in period 11, we maintain the target workload level by releasing lots equal to the total output X_{t-1} of the system, in time units, in the previous period. The number R_{2t} of lots of Product 2 required by the transition profile are released in period t, together with as many lots of Product 1 as the target workload can accommodate. Hence under CONWIP the releases of Product 2 required by the transition profile are maintained while those of Product 1 are reduced. Clearly the use of a single target workload level for the entire planning horizon is not optimal, given the changing distribution of the effective processing times. However, our purpose in this experiment is to examine the basic behavior of policies of this type, which we plan to refine in future work. We consider two different CONWIP policies in our experiments, denoted by CONWIP Low and CONWIP High, in which we set the target workload levels to 123,600 and 133,900 minutes, respectively, to show the impact of different target workload levels. Since the throughput of the individual products is determined by the transition profile and is thus quite intuitive, we focus on the total throughput of both products in our discussion of this experiment.

Figure 4 compares the total output of the fab over time under the two CONWIP policies and the baseline experiment discussed above. It is apparent that CONWIP Low restricts throughput to a lower level than the other policies, and is outperformed in terms of throughput by the baseline policy. Increasing the target workload level under CONWIP High yields higher throughput than CONWIP Low.

Figure 5 shows the evolution of the WIP over time. The baseline policy results in a very large increase in WIP between periods 11 and 53, which is greatly reduced by the CONWIP policies. CONWIP Low is clearly sacrificing output by keeping the target workload level too low. CONWIP High, in contrast, achieves somewhat higher output than CONWIP Low, and recovers from the new product introduction more rapidly, stabilizing at a higher throughput than the Base case. Figure 6 clearly shows that the Base case results in unacceptable levels of WIP, and, by implication, cycle time; utilization is close to 1 for the greater part of the simulation horizon. Both CONWIP policies result in a slight decrease in utilization as the new product begins to enter the system, as throughput is reduced, but both recover quite rapidly. The interesting result is the comparison of the Base case and CONWIP High. CONWIP High represents a higher initial release rate of work into the fab than the Base Case, resulting in higher utilization in the early periods. However, once Product 2 begins to be released, CONWIP High experiences a slight decrease in utilization while that of Base increases dramatically. The decrease in output is due to the releases for each period being computed based on the mean effective processing time in the previous period, which will exceed that of the current period. Thus, for all practical purposes, CONWIP High maintains almost constant utilization, avoiding the excessive WIP of the Base case. It yields lower throughput than Base through period 53, but then stabilizes at a higher throughput.

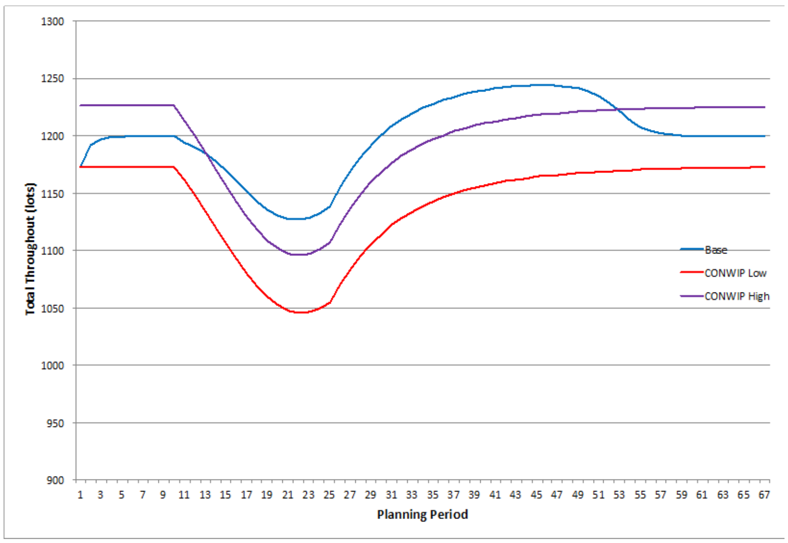

Figure 4: Throughput comparison of CONWIP and base policies.

5 CONCLUSIONS AND FUTURE DIRECTIONS

The preliminary models and experiments presented above produce results consistent with industrial observation. Undoubtedly the mechanism we model is only one of several that may be active in the fab at the same time, including equipment-related improvements that affect all products using the equipment in a similar manner and the acquisition of manufacturing knowledge by equipment operators and manufacturing engineers. The important insight is that the introduction of a new product alters the distribution of the effective processing time of lots at the processing equipment, both in expectation and in variance, requiring this change to be taken into account in the release of work into the fab. In particular, maintaining a constant release rate while replacing lots of the older, more stable product with those of a newer, less stable one in some proportion will lead to reduced throughput and increased cycle times and WIP in the short term, suggesting the need for careful management of releases during product transitions.

The release control policies explored in this paper are simplistic in nature, but suggest several directions for future research. The time-varying nature of the effective processing time distribution suggests the use of CONWIP policies with time-varying target workload levels that allow the impact of the product mix to be taken into account. We conjecture that such a CONWIP policy would initially raise the target workload level to allow increased throughput with an increase in cycle time, reducing it gradually in subsequent periods as the learning takes place and the effective processing time distribution returns to normal. Finally, the mathematical model can be used to develop clearing functions that can incorporate the changing impact of product mix and engineering improvements over time, allowing optimization models of the type suggested by Asmundsson et al. (2009) and tested by Kacar et al. (2013; 2016) to incorporate these effects in at least an approximate manner. The explicit control of the cycle time of individual products, as might be required in high variety fabs, constitutes an additional extension.

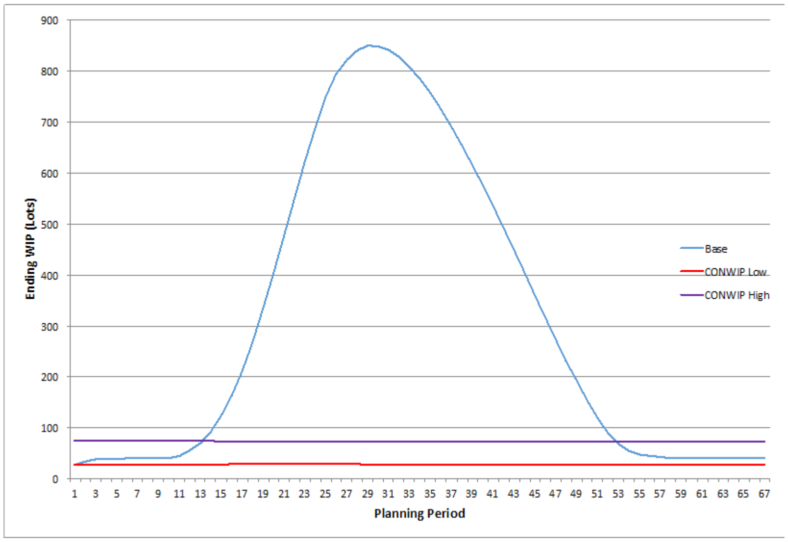

Figure 5: WIP comparison for CONWIP and baseline policies.

REFERENCES

Akcali, E., K. Nemoto, and R. Uzsoy. 2000. "Quantifying the Benefits of Cycle-Time Reduction in Semiconductor Wafer Fabrication." *IEEE Transactions on Electronics Packaging Manufacturing* 23(1): 39-47.

Asmundsson, J. M., R. L. Rardin, C. H. Turkseven, and R. Uzsoy. 2009. "Production Planning Models with Resources Subject to Congestion." *Naval Research Logistics* 56(2): 142-157.

Bilginer, O., and F. Erhun (2010). Managing Product Introductions and Transitions. *Wiley Encyclopedia of Operations Research and Management Science*. J. J. Cochran. New York, John Wiley: 1-12.

Billington, C., H. L. Lee, and C. S. Tang. 1998. "Successful Strategies for Product Rollovers." *Sloan Management Review*(Spring): 23 - 30.

Buzacott, J. A., and J. G. Shanthikumar. 1993. *Stochastic Models of Manufacturing Systems*. Englewood Cliffs, NJ, Prentice-Hall.

Chand, S., H. Moskowitz, A. Novak, I. Pekhi, and G. Sorger. 1996. "Capacity Allocation for Dynamic Process Improvement with Quality and Demand Considerations." *Operations Research* 44(6): 12.

Erhun, F., P. Concalves, and J. Hopman. 2007. "The Art of Managing New Product Transitions." *Sloan Management Review* 48(3): 73 - 80.

Fine, C. H. 1986. "Quality Improvement and Learning in Productive Systems." *Management Technology* 32(10): 1301-1315.

Fine, C. H. 1988. "A Quality Control Model with Learning Effects." *Operations Research* 36(3): 437-444.

Glassey, C. R., and M. G. C. Resende. 1988. "Closed-Loop Job Release Control for VLSI Circuit Manufacturing." *IEEE Transactions on Semiconductor Manufacturing* 1: 36-46.

Goldratt, E., and R. E. Fox. 1986. *The Race*. Croton-on-Hudson, NY, North River Press.

Gopal, A., M. Goyal, S. Netessine, and M. Reindorp. 2011. *The Impact of New Product Intorduction on Plant Productivity in the North American Automotive Industry*. Faculty and Research Working Paper, INSEAD.

Graves, S. C. 1986. "A Tactical Planning Model for a Job Shop." *Operations Research* 34: 522-533.

Hiller, R. S., and J. F. Shapiro. 1986. "Optimal Capacity Expansion Planning When There Are Learning Effects." *Management Science* 32(9): 11.

Hopp, W. J., and M. L. Spearman. 2008. *Factory Physics : Foundations of Manufacturing Management*. Boston, Irwin/McGraw-Hill.

Kacar, N. B., D. F. Irdem, and R. Uzsoy. 2012. "An Experimental Comparison of Production Planning Using Clearing Functions and Iterative Linear Programming-Simulation Algorithms." *IEEE Transactions on Semiconductor Manufacturing* 25(1): 104-117.

Kacar, N. B., L. Moench, and R. Uzsoy. 2013. "Planning Wafer Starts Using Nonlinear Clearing Functions: A Large-Scale Experiment." *IEEE Transactions on Semiconductor Manufacturing* 26(4): 602-612.

Kacar, N. B., L. Moench, and R. Uzsoy. 2016. "Modelling Cycle Times in Production Planning Models for Wafer Fabrication." *IEEE Transactions on Semiconductor Manufacturing* 29(2): 153 - 167.

Kacar, N. B., and R. Uzsoy. 2010. Estimating Clearing Functions from Simulation Data. *Winter Simulation Conference*. B. Johansson, Jain, S., Montoya-Torres, J., Hugan, J., Yucesan, E. Baltimore, MD: 1699-1710.

Karmarkar, U. S. 1989. "Capacity Loading and Release Planning with Work-in-Progress (Wip) and Lead-Times." *Journal of Manufacturing and Operations Management* 2(1): 105-123.

Kim, S. and R. Uzsoy. 2008. "Integrated Planning of Production and Engineering Process Improvement." *IEEE Transactions on Semiconductor Manufacturing* 21: 390-398.

Kim, S., and R. Uzsoy. 2013. "Modeling and Analysis of Integrated Planning of Production and Engineering Process Improvement." *IEEE Transactions on Semiconductor Manufacturing* 26(3): 414-422.

Leachman, R. C. 1996. *The Competitive Semiconductor Manufacturing Survey: Third Report on Results of the Main Phase*. The Competitive Semiconductor Manufacturing Program. Berkeley, CA, University of California, Berkeley.

Leachman, R. C., and S. Ding. 2007. "Integration of Speed Economics into Decision-Making for Manufacturing Management." *International Journal of Production Economics* 107: 39-55.

Levinthal, D. A., and D. Purohit. 1989. "Durable Goods and Product Obsolescence." *Marketing Science* 8(1): 35-56.

Liao, W. M. 1979. "Effects of Learning on Resource Allocation Decisions." *Decision Sciences* 10: 10.

Lim, W. S., and C. S. Tang. 2006. "Optimal Product Rollover Strategies." *European Journal of Operational Research* 174: 905 - 922.

Lu, S., D. Ramaswamy, and P. R. Kumar. 1994. "Efficient Scheduling Policies to Reduce Mean and Variance of Cycle Time in Semiconductor Plants." *IEEE Transactions on Semiconductor Manufacturing* 7: 374-388.

McDonald, C. J. 1998. "The Evolution of Intel's Copy Exactly! Technology Transfer Method." *Intel Technology Journal(*Q4): 1-6.

Missbauer, H. 2002. "Aggregate Order Release Planning for Time-Varying Demand." *International Journal of Production Research* 40: 688-718.

Padmanabhan, V., S. Rajiv, and K. Srinivasan. 1997. "New Products, Upgrades and New Releases: A Rationale for Sequential Product Introduction." *Journal of Marketing Research* 34(4): 456-472.

Reeves, G. R., and J. R. Sweigart. 1981. "Product-Mix Models When Learning Effects Are Present." *Management Science* 27(2): 9.

Shenoy, S., and A. Daniel. 2006. *Intel Architecture and Silicon Cadence: The Catalyst for Industry Innovation*. White Paper, Intel Corporation.

Spearman, M. L., W. J. Hopp, and D. L. Woodruff. 1989. "A Hierarchical Control Architecture for Constant Work-in-Process (CONWIP) Production Systems." *Journal of Manufacturing and Operations Management* 2/3: 147-171.

Spearman, M. L., D. L. Woodruff, and W. Hopp. 1990. "CONWIP: A Pull Alternative to Kanban." *International Journal of Production Research* 28(5): 879-894.

Terwiesch, C., and R. E. Bohn. 2001. "Learning and Process Improvement During Production Ramp-Up." *International Journal of Production Economics* 70: 1-19.
Wein, L. M. 1988. "Scheduling Semiconductor Wafer Fabrication." *IEEE Transactions on Semiconductor Manufacturing* 1: 115-130.

AUTHOR BIOGRAPHIES

ATCHYUTA BHARADWAJ MANDA is a masters student in the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. He holds a Bachelor of Technology degree in Mechanical Engineering from GITAM University. His research interests are in production systems and supply chain management. His email address is amanda@ncsu.edu.

REHA UZSOY is Clifton A. Anderson Distinguished Professor in the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. He holds BS degrees in Industrial Engineering and Mathematics and an MS in Industrial Engineering from Bogazici University, Istanbul, Turkey. He received his Ph.D. in Industrial and Systems Engineering in 1990 from the University of Florida. His teaching and research interests are in production planning, scheduling, and supply chain management. He was named a Fellow of the Institute of Industrial Engineers in 2005, Outstanding Young Industrial Engineer in Education in 1997, and has received awards for both undergraduate and graduate teaching. His email address is ruzsoy@ncsu.edu.

KARL G. KEMPF is a Senior Fellow and Director of Decision Engineering at Intel Corporation in Chandler Arizona where he has been involved in designing and implementing decision support systems from strategic product design to tactical supply chain execution. He is a member of the National Academy of Engineering, a Fellow of both the IEEE and INFORMS, and has (co) authored more than 150 publications on various topics in decision science. He holds a B.A. in Physics, a B.S. in Chemistry, a Ph.D. in Applied Mathematics, and pursued post-doctoral studies in Artificial Intelligence. His email address is karl.g.kempf@intel.com.

SUKGON KIM received the B.S. and M.S. degrees in industrial engineering from Hanyang University, Seoul, Korea, in 1990. He received the Ph.D. degree in industrial engineering from Purdue University,West Lafayette, IN, in 2007. He worked at Samsung semiconductor in Korea for built planning and scheduling systems for twelve years. He also taught at Northern Illinois University as an assistant professor. He currently works for Maxim Integrated in its Supply Chain group, where he can be reached at sukgon@gmail.com.