Proceedings of the 2015 Winter Simulation Conference
L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds.

SIMULATION STUDIES ON MODEL SELECTION IN PM PLANNING OPTIMIZATION

Minho Lee
James R. Morrison

Department of Industrial and Systems Engineering
KAIST
335 Gwahangno Yuseong-gu
Daejeon, 305-701, REPUBLIC of SOUTH KOREA

Adar A. Kalir

Fab/Sort Manufacturing Division
Intel Corporation
2 HaZoran St., Qiriat-Gat 82109, ISRAEL

ABSTRACT

In semiconductor manufacturing, preventive maintenance (PM) is complicated and essential. Since tool down time contributes significantly to manufacturing flow variability and thus mean cycle time, effective PM planning is important. Here we extend existing PM planning methods to allow for four categories of PM models. We study the quality of these PM models and the resulting optimized PM plans via simulation. We observe that the approximate mean cycle time formulae for these models are generally of good accuracy. Our studies show that good PM plans suggested from the use of these approximate formulae remain good plans in the true system. Finally, we study the implications of using optimized PM plans generated from one type of model in another type of model. The results suggest that good PM plans are relatively insensitive to which of the four PM models are selected.

1 INTRODUCTION

Preventive Maintenance activities (PMs) prevent unplanned downtime by imposing planned downtime. Well-planned PMs increase overall equipment availability and equipment reliability and decrease unplanned tool-failures at the expense of planned downtime. In high technology industries such as semiconductor manufacturing, the complexity and diversity of PM tasks have increased. PMs are critical for manufacturing performance and require careful consideration.

The vast majority of research on PMs has focused on the scheduling of an existing PM plan into the daily manufacturing operations. However, in Kalir (2013), the question of how frequently PMs should be planned to mitigate manufacturing disturbances was studied. They studied a G/G/m queueing model subject to one PM (e.g., a monthly PM) and attempted to determine if the PM activities should be split (e.g., into two half sized biweekly PM events). Morrison, Kim, and Kalir (2014) extended this approach to allow for multiple PMs with different cycles (e.g., monthly, quarterly and yearly PM cycles). That approach allowed for more practical PM planning. However, in both of these studies, the PMs were considered to be time-based and preemptive.

Wu (2014) classified tool down events (e.g., tool failures, interruptions and PM events) into four categories: time-based preemptive (TB/P), run-based preemptive (RB/P), time-based non-preemptive (TB/NP) and run-based non-preemptive (RB/NP). Wu (2014) also provided approximate equations for the mean cycle time in G/G/1 queues subject to these classes of interruptions.

Here, we extend the nonlinear optimization models for PM planning as given in Morrison, Kim and Kalir (2014) to account for all four classes of interruptions from Wu (2014). We study the quality of the mean cycle time approximations via simulation at the optimal PM cycles suggested by the nonlinear program. The approximation accuracy is acceptable. We study the quality of the optimal PM cycle decisions suggested by the nonlinear program (based on the approximate mean cycle time equations) by simulating how well they perform relative to a simulation based optimal setting for the PM plans. The PM cycle recommendation from the nonlinear program performs nearly as well as that obtained via detailed simulation. Finally, we consider how the PM cycle recommendation given from one of the PM types (from Wu (2014)) performs when used in a system that operates under a different PM type. We observe that the mean cycle time is relatively invariant to which model provided the PM recommendation. That is, it does not really matter which of the four classes of PMs are used. The resulting cycle time only changes by a few percent. This is because the mean cycle time manifold is quite flat in a relatively large region around the good PM cycle decisions.

The organization of the paper is as follows. In Section 2, we review the classification of Wu (2014) and provide some preliminary concepts. The quality of the mean cycle time approximations and the optimization results are studied via simulation in Section 3. We study the sensitivity of the results to the model used in Section 4. Concluding remarks are presented in Section 5. Detailed models are provided in the Appendix.

2 PRELIMINARIES

We first review four types of equipment models and approximations for their mean cycle time as detailed in Wu (2014). A model for PM planning from Morrison, Kim and Kalir (2014) which was generalized from the work of Kalir (2013) is reviewed. Finally, PM data that is similar to real data that we will use throughout our study is introduced.

2.1 The Four Categories of PMs

In Wu (2014), four classes of queueing models for failure-prone tools were identified based on the manner in which the failures occur. We review these four categories of models as well as approximations for their mean cycle time as discussed in Wu (2014). We use the word interruption synonymously with the word failure. In our context, we will focus on PM events as our interruptions or failures.

As discussed in Wu (2014), according to Buzacott and Hanifin (1978), interruptions can be classified as run-based and time-based. A run-based interruption occurs after the tool has been actively processing for some duration of time. They can only happen while the tool is processing. A time-based interruption occurs some duration of time after the tool has returned from the previous interruption. They can occur at any time regardless of whether the tool is busy or idle.

These interruptions can further be classified as preemptive or non-preemptive (Wu (2014)). Preemptive interruptions can occur anytime (even during processing), while non-preemptive interruptions can only occur before or after processing. That is, when a preemptive interruption occurs, the tool immediately enters the interrupt state even if it was processing at that time. If a non-preemptive interruption occurs, the tool waits until it finishes the processing of its current job before entering the interrupt state.

There are thus four categories of interruptions: TB/P (Time-Based Preemptive), RB/P (Run-Based Preemptive), TB/NP (Time-Based Non-Preemptive), and RB/NP (Run-Based Non-Preemptive).

Approximations for the mean cycle time of a G/G/1 queue subject to the four types of interruptions were reviewed and developed in Wu (2014). We have slightly modified the form of those approximations to match our intent to use them to reflect different categories of PMs (our interruptions). The details are provided in the Appendix. Here, we give the mean cycle time approximation results with limited commentary.

The mean cycle time for a G/G/1 queue subject to TB/P, RB/P, TB/NP and RB/NP PM events (our interruptions) are provided in equations (1), (2), (3) and (4), respectively.

$$E(CT\_TBP) \approx \frac{1}{\mu A} + \frac{1}{\mu A}(\frac{\rho}{1-\rho})(\frac{1}{2})(C_A^2 + C_S^2 + \frac{(C_{AR}^2+C_R^2)A(1-A)m_R\mu}{\rho}), \quad (1)$$

$$E(CT\_RBP) \approx \frac{1}{\mu A} + \frac{1}{\mu A}(\frac{\rho}{1-\rho})(\frac{1}{2})(C_A^2 + C_S^2 + (1 + C_R^2)A(1-A)m_R\mu). \quad (2)$$

$$E(CT\_TBNP) \approx \frac{1}{\mu} + \frac{\frac{m_R}{m_T}\frac{C_{AR}^2+C_R^2}{2}m_R + \frac{\lambda}{\mu}\frac{C_A^2+C_S^2}{2}\frac{1}{\mu}}{(1-\frac{m_R}{m_T}-\frac{\lambda}{\mu})(1-\frac{m_R}{m_T})}. \quad (3)$$

$$E(CT\_RBNP) \approx \frac{1}{\mu A} + \frac{1}{\mu A}(\frac{\rho}{1-\rho})(\frac{1}{2})(C_A^2 + C_G^2). \quad (4)$$

Here, μ, A, and ρ are the tool service rate, tool availability, and tool loading, respectively. The variables C_S, C_G, C_A, C_R, and C_{AR} are the coefficients of variation of the service times, effective service times, interarrival times, repair times, and times from the start of one PM to the start of the next, respectively. Finally, m_R and m_T are the mean times to repair the tool (the mean PM duration) and mean time between the start of one PM and the start of the next PM (we call this the mean PM cycle). Note that these approximations are valid only so long as $\rho < 1$.

2.2 PM Planning Optimization

We now turn our attention to PM plan optimization as detailed in Morrison, Kim and Kalir (2014). The goal of the optimization will be to minimize the mean cycle time for a G/G/1 queue with PMs by selecting the mean time between PM events (m_T for a single PM cycle). There may be n PMs that the tool requires. For example, the tool may require three different PMs with different mean PM cycle durations, e.g., a monthly PM, a quarterly PM and a yearly PM. In that case, we use m_{Ti} as the mean duration of PM cycle i. These mean durations are the decision variables in the PM plan optimization; see Morrison, Kim and Kalir (2014). There may be lower and upper bounds L_i^{min} and L_i^{max} that limit the acceptable values for m_{T_i}. In Morrison, Kim and Kalir (2014) an approximate method to combine these separate failure cycles into a single one is provided. From this, the net mean time between PMs m_T can be calculated for use in the G/G/1 approximations provided in Wu (2014). Please refer to Morrison, Kim and Kalir (2014) for the details as they are omitted here for brevity.

Depending on which category of interruption the PMs fall into, the objective function for the optimization will use the approximate mean cycle time approximation of equations (1-4). There the variables ρ, A, m_R, m_T, and C^2_R are functions of the decision variables m_{Ti}, $i = 1, \ldots, n$. The nonlinear program proposed in Morrison, Kim, and Kalir (2014) for G/G/1 queues with PMs is provided in (5-8).

$$\text{Min } E(CT) \quad (5)$$

Subject to

$$0 \le \rho(m_T) \le 1, \quad (6)$$

$$L_i^{min} \le m_{T_i} \le L_i^{max}, \forall i = 1, \cdots, n, \quad (7)$$

$$m_{T_i} \ge 0, \forall i = 1, \cdots, n. \quad (8)$$

We will study the performance of these approximations and this optimization as a function of the four categories of PMs.

2.3 PM Data Used in this Paper

Throughout the remainder of the paper, we use PM data that is similar to real PM data from semiconductor manufacturing equipment. Table 1 provides the parameters for two PM cycles and the toolset. The distribution, mean value, and coefficient of variation of each random variable are given. The superscript 0 indicates that the data was the **original data** for the PM prior to optimization (the original setting for the PMs). Type 1 and Type 2 PMs are represented by the subscripts 1 and 2, respectively. The variables $m_{T_i^0}$, $m_{PM_i^0}$, and m_{F^0} are the mean duration of the original PM cycles, mean duration of the activities for PM type i (time the tool is down on average for the PM of type i), and time the tool is available for production after the PM is complete, respectively. SU_i is the mean duration of setup for a type i PM (it is not a decision variable). Note that our assumption of exponential interarrival times in Table 1 gives us an M/G/1 failure prone queue.

Table 1: Input parameters for our M/G/1 queue with two types of PM.

RV	Distribution	Mean (hours)	C_{RV}
$m_{T_1^0}$ $m_{T_2^0}$		240 720	
$m_{PM_1^0}$ $m_{PM_2^0}$	Erlang (1/33, 2) Erlang (1/15, 2)	66 30	$1/\sqrt{2}$ $1/\sqrt{2}$
SU_1 SU_2		3 4	
m_{F^0}	Exp (1/119.75)	119.75	1
Interarrival $(1/\lambda)$	Exp (0.13)	7.69	1
Service$(1/\mu)$	U (3.425, 4.075)	3.75	0.05

3 EXAMPLES AND MODEL VERIFICATION

In Section 3.1, we provide some examples of the results of the optimization using the four different categories of Wu (2014) for the PMs. In Section 3.2, we then seek to verify the quality of the approximations of (1-4) by comparing the mean cycle time approximations with simulation. Finally, in Section 3.3, for TB/NP and RB/NP models, we seek to verify the quality of the results of the optimization by comparing a good PM plan suggested via the approximate equations with good PM plans suggested via detailed simulation based search.

3.1 Examples of the Optimization over the Four Categories of PM

We consider examples of the behavior of the optimization model in Section 2.2 across the 4 types of interruption models (TB/P, RB/P, TB/NP, and RB/NP). The PM data is as in Table 1. The results suggest that, while the optimal cycle time can differ according to models, the percentage improvement in cycle time as well as the optimal decisions are surprisingly close.

The results of the optimization are presented in Table 2 according to the types of PM: TB/P, RB/P, TB/NP, and RB/NP. Mean downtime (m_R), mean uptime (m_F), availability (A) and utilization (ρ) of the original cycle (before optimization) and optimal cycle are distinguished by superscript 0 and *, respectively. Since the original PM cycles ($m_{T_1^0}$, $m_{T_2^0}$) are the same regardless of models, all variables of in the original cycle cases are also same, except for the mean cycle time value (which use the four different approximations (1-4)). We used Excel solver and the Matlab Optimization Toolbox to solve the nonlinear program of Section 2.2.

Table 2: Results for the PM plan optimization in an M/G/1 queue for the four categories of PM.

Original Cycle					Optimal Cycle				
Variables	**TB/P**	**RB/P**	**TB/NP**	**RB/NP**	**Variables**	**TB/P**	**RB/P**	**TB/NP**	**RB/NP**
$\boldsymbol{m_{T_1^0}}$	240				$\boldsymbol{m_{T_1^*}}$	**55.3597**	**62.0322**	**56.1993**	**58.6370**
$\boldsymbol{m_{T_2^0}}$	720				$\boldsymbol{m_{T_2^*}}$	**424.6187**	**475.1779**	**430.9775**	**449.4445**
$\boldsymbol{m_{R^0}}$	60.2500				$\boldsymbol{m_{R^*}}$	18.6240	20.4907	18.8589	19.5408
$\boldsymbol{m_{F^0}}$	119.7500				$\boldsymbol{m_{F^*}}$	30.3507	34.3785	30.8575	32.3289
$\boldsymbol{A^0}$	0.6653				$\boldsymbol{A^*}$	0.6197	0.6266	0.6207	0.6233
$\boldsymbol{\rho^0}$	0.7328				$\boldsymbol{\rho^*}$	0.7866	0.7781	0.7854	0.7822
$\boldsymbol{E(CT^0)}$	**73.5657**	**57.4852**	**143.4754**	**109.0027**	$\boldsymbol{E(CT^*)}$	**42.2181**	**36.6919**	**79.9766**	**64.6088**

We observe several things.

- For the optimal cycles, the availability decreases and the utilization increases in all cases.
- While the absolute values vary, cycle time improvements are similar and vary from 36.17% to 44.26%.
- In all cases, the optimal decisions ($m_{T_1^*}$, $m_{T_2^*}$) are surprisingly similar. The $m_{T_1^*}$ values vary from 55.4 to 62.0 hours. The $m_{T_2^*}$ values vary from 424.6 to 475.2 hours.

These results suggest the optimal decisions are somewhat invariant to the model choice. We will further investigate this point in our study on the sensitivity to model selection in Section 4.

3.2 Comparison of Approximation for Mean Cycle Time with Simulation

Wu (2014) conducted extensive simulation experiments to validate the approximations for the mean cycle time in the four classes of models he identified. In general the approximations performed well. Here, in the context of our problem we similarly compare the mean cycle time values obtained via the approximations and via simulation. We check the quality of the approximations at the PM cycle values suggested as optimal by the nonlinear program of Section 2.2 (these are the cycles given in the optimal cycle section of Table 2.).

We use the Autosched AP simulation software. We simulate for 23000 days (5,000 days are for warm-up and 18,000 days are for data collection) and 30 replications. The results are provided in Table 3. Table 3 provides the percent difference between simulation and the approximations (1-4). The model sensitivity (mean simulation cycle time divided by the mean cycle time approximation) is provided as $\frac{Simulation}{Optimization}$. A model sensitivity value close to 1 is good.

Table 3: Results of simulation at the optimized PM cycles.

Types of interruption	**Simulated Cycle Time**	**Approximate Cycle Time**	**Diff (%)**	$\boldsymbol{\frac{Simulation}{Optimization}}$
TB/P	39.1678	42.2181	7.7879	0.9277
RB/P	69.1614	36.6919	46.9474	1.8849
TB/NP	79.4182	79.9766	0.7031	0.9930
RB/NP	66.9974	64.6088	3.5653	1.0370

The TB/NP model appears the best in this example with a sensitivity near 1. The run based model results (RB/P and RB/NP) are worse than their time based counterparts (TB/P and TB/NP). This agrees with what was observed in Wu (2014) and they suggested it is worse for low tool utilizations.

3.3 Optimization Model Validation

We next consider the quality of the optimal PM cycles suggested by the nonlinear optimization of Section 2.2 This is done by searching for the optimal PM cycles via simulation and comparing the result to that suggested by the approximate cycle time based nonlinear optimization. Since most real PMs are non-preemptive, we consider the TB/NP and RB/NP models.

For each type of model, we search for the optimal PM cycles via simulation by considering numerous cases for m_{T1} and m_{T2}. We consider m_{T1} = 40, 45, 50, 55, 60, 65, 70, 155, and 240 hours and m_{T2} = 120, 235, 350, 400, 450, 500, 550, 700, and 850 hours and all combinations formed by those values. We have more points near the optimal PM cycles given in Table 2. We do not search below 40 hours for m_{T1} and 120 hours for m_{T2} since these lower bounds were given in Morrison, Kim and Kalir (2014) and represent some practical fab lower limits. We thus consider 81 pairs of m_{T1} and m_{T2} values. In all cases we simulate for 20 replications and 20000 days (5,000 days of warm-up and 15,000 days of data collection).

In both the TB/NP and RB/NP models, we compare the results for the simulated mean cycle time values for these 81 cases with the values from the approximations of (3-4). We will see that for each m_{T1} and m_{T2} value, the simulated and approximate mean cycle times are very close. The best m_{T1} and m_{T2} values obtained from the 81 cases in the simulation (which have the smallest mean cycle time value) are nearly equivalent with those given by the best m_{T1} and m_{T2} values obtained from the approximations.

3.3.1 Time-Based Non-Preemptive Model

We first consider an M/G/1 queue with TB/NP preventive maintenance events. The mean cycle time values for the simulation and approximations for the 81 pairs of m_{T1} and m_{T2} values considered are provided in Tables 4 and 5, respectively. From among the 81 decisions considered, the simulation best decision is m_{T1} = 60 and m_{T2} = 400 hours. The resulting mean cycle time is 78.12 hours. For the approximation, the best decision from among those considered is m_{T1} = 55 and m_{T2} = 450 hours with an approximate cycle time of 80.01 hours. Had we used this decision in the real system, the resulting mean cycle time (from the simulation Table 4) would have been 80.01 hours. This is only 2.4% away from the best value obtained in Table 4. As such, the result suggested by the approximate values is still a very good one. This is because the simulated mean cycle time values in Table 4 are quite invariant to moderate changes in the decisions m_{T1} and m_{T2} around the best value. This result suggests that the approximate equation based optimization of Section 2.2 can give good decisions even though there are some errors introduced by the approximation. That is, the shapes of the m_{Ti} to mean cycle time curves obtained via simulation and approximation, refer to Figure 1, are similar and flat near the optimal decisions.

Table 4: Result of simulated mean cycle time with TB/NP PMs.

Simulation		**m_{T_2} (hours)**								
		120	**235**	**350**	**400**	**450**	**500**	**550**	**700**	**850**
m_{T_1} (hours)	**40**	111.33	86.48	85.58	86.45	86.88	86.50	86.60	86.04	89.62
	45	99.42	86.27	82.91	82.65	82.69	82.95	83.00	80.91	82.87
	50	100.42	83.14	81.02	79.52	81.77	81.41	82.15	84.65	85.05
	55	91.80	84.03	80.09	80.92	80.01	80.21	83.48	83.75	80.72
	60	94.23	84.60	79.46	**78.12**	80.45	79.77	81.69	81.88	84.92
	65	98.71	82.93	82.95	79.87	82.91	80.27	81.41	82.63	84.72
	70	93.99	82.50	85.18	82.72	81.81	82.77	81.58	82.98	85.72
	155	130.64	116.45	105.71	114.89	108.73	113.90	109.94	109.85	113.75
	240	183.45	156.88	138.29	144.48	150.66	144.83	138.10	155.40	146.39

Table 5: Result of approximate mean cycle time with TB/NP PMs.

Approximation		**m_{T_2} (hours)**								
		120	**235**	**350**	**400**	**450**	**500**	**550**	**700**	**850**
m_{T_1} (hours)	**40**	105.07	88.43	85.39	85.06	85.03	85.19	85.49	86.92	88.81
	45	99.67	84.99	82.28	82.00	81.98	82.14	82.43	83.76	85.51
	50	96.86	83.32	80.79	80.53	80.51	80.66	80.93	82.19	83.83
	55	95.57	82.72	80.29	80.03	**80.01**	80.14	80.40	81.58	83.14
	60	95.24	82.81	80.43	80.17	80.14	80.26	80.50	81.62	83.11
	65	95.57	83.38	81.02	80.75	80.70	80.81	81.03	82.10	83.53
	70	96.35	84.30	81.92	81.64	81.58	81.68	81.88	82.91	84.28
	155	131.09	116.15	112.65	112.04	111.70	111.55	111.53	111.99	112.89
	240	174.22	154.59	149.61	148.63	148.00	147.61	147.40	147.42	148.01

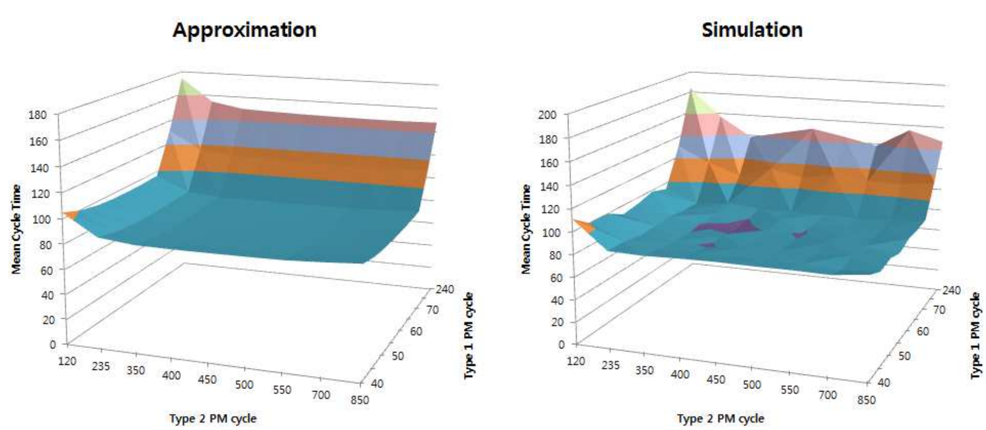

Figure 1: Graphs of the approximate and simulated mean cycle time for TB/NP model.

3.3.2 Run-Based Non-Preemptive Model

We now consider an M/G/1 queue with RB/NP preventive maintenance events. The mean cycle time values for the simulation and approximations for the 81 pairs are provided in Tables 6 and 7, respectively.

From among the 81 decisions considered, the simulation best decision is m_{T1} = 55 and m_{T2} = 450 hours. The resulting mean cycle time is 63.71 hours. For the approximation, the best decision is m_{T1} = 60 and m_{T2} = 450 hours with an approximate cycle time of 64.62 hours. Had we used this decision in the real system, the resulting mean cycle time (from the simulation Table 6) would have been 65.91 hours. There is only 3.5% of difference from the best value obtained in Table 6. As such, the result suggested by the approximation values is still a very good one. The reasons are similar to those given for the TB/NP model in Section 3.3.1 above; see Figure 2.

Table 6: Results of the simulated mean cycle time with RB/NP PMs.

Simulation		**m_{T_2} (hours)**								
		120	**235**	**350**	**400**	**450**	**500**	**550**	**700**	**850**
m_{T_1} (hours)	**40**	88.49	77.55	74.17	69.01	74.55	73.11	71.09	71.23	75.35
	45	82.68	73.20	70.95	72.00	68.65	74.20	68.32	69.75	71.16
	50	80.46	67.04	65.27	67.55	65.57	70.21	66.08	66.29	68.60
	55	82.33	68.12	66.90	67.32	**63.71**	68.75	67.25	67.02	67.53
	60	77.87	68.27	65.64	67.09	65.91	69.59	65.17	67.04	66.21
	65	73.80	65.78	65.58	64.81	68.28	70.26	64.68	65.50	66.45
	70	76.04	68.56	65.96	65.32	67.04	69.39	65.38	66.84	66.72
	155	94.87	89.82	86.23	85.27	88.30	87.02	86.63	85.41	83.75
	240	121.28	112.29	115.58	108.47	108.92	117.17	108.93	110.79	113.05

Table 7: Results of the approximate mean cycle time with RB/NP PMs.

Approximation		**m_{T_2} (hours)**								
		120	**235**	**350**	**400**	**450**	**500**	**550**	**700**	**850**
m_{T_1} (hours)	**40**	83.49	71.73	69.40	69.08	68.97	69.00	69.12	69.87	70.94
	45	79.22	68.93	66.89	66.62	66.54	66.58	66.72	67.44	68.46
	50	76.84	67.45	65.58	65.34	65.27	65.32	65.45	66.15	67.13
	55	75.59	66.76	65.00	64.77	64.71	64.76	64.89	65.57	66.51
	60	75.06	66.60	64.90	64.68	**64.62**	64.67	64.79	65.45	66.36
	65	75.01	66.80	65.13	64.92	64.86	64.90	65.02	65.66	66.54
	70	75.31	67.25	65.61	65.39	65.33	65.37	65.49	66.10	66.96
	155	96.57	87.49	85.38	85.02	84.84	84.77	84.78	85.14	85.76
	240	124.27	112.84	109.98	109.44	109.10	108.91	108.83	108.96	109.42

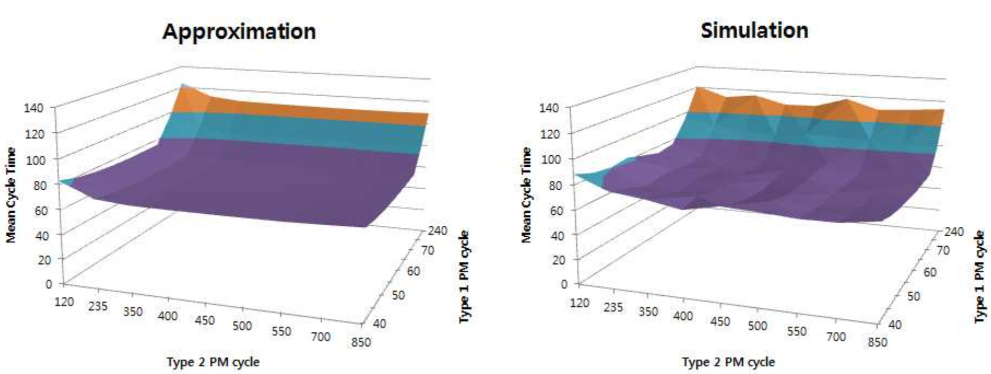

Figure 2: Graphs of approximate and simulated mean cycle time for RB/NP model.

4 SENSITIVITY TO MODEL SELECTION

Our goal here is to determine how much it matters which model we select relative to the optimal decision. That is, if we obtain optimal PM cycles from the nonlinear program using equation (1) as its objective function, will those PM cycle decisions still be good ones if the true system behaves as in equation (2) or (3) or (4). We refer to this as sensitivity to model selection. As we have already seen in Section 3.1, the optimal decisions provided when using the different models are similar. Here we study this question more deeply.

Table 8 provides optimal PM cycles when using approximations (1), (2), (3) and (4) in the nonlinear program objective function corresponding to PMs of type TB/P, RB/P, TB/NP and RB/NP, respectively. We include them here for convenience in clearer form than in Section 3.1.

Table 8: Optimal solution of each model.

Types of interruption	**Type 1 PM (hours)**	**Type 2 PM (hours)**
TB/P	55.3597	424.6187
RB/P	62.0322	475.1779
TB/NP	56.1993	430.9775
RB/NP	58.6370	449.4445

The results of our study are provided in Table 9. There we compare the results of the mean cycle for various models when using optimal m_{Ti} values suggested for another model. The "Actual Model" is the way the system behaves. For the actual model, we conduct 25,000 days of simulation (5,000 days of warm-up and 20,000 days for collecting data) to determine the mean cycle time using the m_{Ti} value suggested as optimal in Table 8 for that model. This value is referred to as "E_{sim}[CT] with optimized m_{Ti} in actual model". For example, in every RB/P "Actual Model" row is the simulated mean cycle time one would achieve by using the optimal PM cycles from the nonlinear program with equation (2) for the objective function; its value is 69.1614 hours for all such rows. We also provide the mean cycle time obtained from simulation of the "Actual Model" when the m_{Ti} values are set to those provided as optimal for the "Test Model" system. This called "E_{sim}[CT] with optimized m_{Ti} from test model".

For example, with "Test Model" TB/P and "Actual Model" RB/NP, the first mean cycle time value was obtained from simulation of the RB/NP system at the m_{Ti} values given in Table 8 for that same system. The second mean cycle time value was obtained from simulation of the RB/NP system with the m_{Ti} values given in Table 8 for the **TB/P system**.

The percent difference between these two cycle time values is provided as "% difference". This difference varies from 0.15 % to 4.57 %; less than 5%. From this, we conclude that the optimal decision is relatively invariant to the model used.

Table 9: Simulated mean cycle time when applying optimal m_{Ti} values from one model to a different one.

Test Model	Actual Model	E_{sim}[CT] with optimized m_{Ti} in actual model	E_{sim}[CT] with optimized m_{Ti} from test model	Diff (%)
TB/P	RB/P	69.1614	68.3693	1.1453
	TB/NP	79.4182	80.4664	1.3198
	RB/NP	66.9974	66.8762	0.1809
RB/P	TB/P	39.1678	40.4041	3.1564
	TB/NP	79.4182	81.9266	3.1585
	RB/NP	66.9974	65.0629	2.8874
TB/NP	TB/P	39.1678	38.9253	0.6191
	RB/P	69.1614	67.5705	2.3003
	RB/NP	66.9974	64.9584	3.0434
RB/NP	TB/P	39.1678	39.0886	0.1511
	RB/P	69.1614	66.0004	4.5705
	RB/NP	79.4182	82.0235	3.2805

5 CONCLUDING REMARKS

As tool down time contributes significantly to cycle time in semiconductor manufacturing, it is essential to carefully consider PM plans. Recent efforts in Kalir (2013) and Morrison, Kim and Kalir (2014) have investigated nonlinear programming methods to optimize PM plans. They considered PMs of the TB/P class. However, as studied in Wu (2014), tool down events can be classified into four categories: TB/P, RB/P, TB/NP and RB/NP. Here, we sought to extend the PM planning results into these other categories and investigate via simulation the quality of these models.

As in Wu (2014), we found that the approximate mean cycle time formulae for M/G/1 queues under the four categories of PMs are in general of good quality. The run based models (RB) were in general less accurate than their time based (TB) counterparts. In two examples considered, we observed that a decision suggested as good by the approximate mean cycle time equations was also good in the true (simulated) system. This was due to the flatness of the objective function curve near the good decisions. Finally, we studied the sensitivity of the mean cycle time to the model chosen for use in the nonlinear program. We observed that no matter which model is the true model, it does not really matter if you use a different model in the nonlinear program. The decisions provided by the nonlinear program were relatively good for all models in the cases considered.

Future directions include increasing the number of replications in our simulations and conducting studies for a larger number of test systems. It would be good to consider a variety of distributions for our random variables. We will also seek properties that provide insight for PM planning and lead to rules of thumb.

A APPENDICES

The results given here are modified versions of those developed and provided in Wu (2014). We adjust them to fit our PM context.

A.1 Time-Based Preemptive

TB/P events have impact on the service time of a customer, so we consider the mean generalized service time of a customer. $E(G)$ is the mean generalized service time which considers the availability of a machine resulting from downtimes (E(G) = $\frac{1}{\mu A}$). In Wu (2014), the mean QT (queue time) and CT (cycle time) are modeled as $E(QT) = \frac{\rho_G E(R_G)}{(1-\rho_G)} + (1 - A_{NP})E(R_D)$ and $E(CT) = E(QT) + E(G)$.

In our model, the mean downtimes ($E(D)$) and uptimes ($\frac{1}{\eta}$) of a TB/P interruption are treated as m_R and m_F, respectively. Thus, the coefficient of variation of generalized service time is $C_G^2 = C_S^2 + \frac{(1+C_R^2)A(1-A)m_R\mu}{\rho}$ and the residual downtime is $E(R_D) = \frac{(1+C_R^2)}{2}m_R$. The mean cycle time approximation is

$$E(CT\_TBP) \approx \frac{1}{\mu A} + \frac{1}{\mu A}\left(\frac{\rho}{1-\rho}\right)\left(\frac{1}{2}\right)\left(C_A^2 + C_S^2 + \frac{(C_{AR}^2+C_R^2)A(1-A)m_R\mu}{\rho}\right). \qquad \text{(A.1)}$$

A.2 Run-Based Preemptive

When we treat the PMs as RB/P, we do not allow the PMs to occur when the machine is idle. Since RB/P interruptions do not occur during a machine idle period, the mean queue time is $E(QT) = \frac{\rho_G E(R_G)}{(1-\rho_G)}$. The coefficient of variation of the generalized service time is $C_G^2 = C_S^2 + (1 + C_R^2)A(1 - A)m_R\mu$ and the mean cycle time equation of RB/P is

$$E(CT\_RBP) \approx \frac{1}{\mu A} + \frac{1}{\mu A}\left(\frac{\rho}{1-\rho}\right)\left(\frac{1}{2}\right)\left(C_A^2 + C_S^2 + (1 + C_R^2)A(1 - A)m_R\mu\right). \qquad \text{(A.2)}$$

A.3 Time-Based Non-Preemptive

For TB/NP interruptions, Wu (2014) models the interruption as a high priority customer. As such, when a PM arrives, it waits until the tool finishes its current customer and immediately takes over the tool (it may wait behind another queued PM). As our cycle time of interest is for the customers, we need only the mean queue time and service time of customers. From Wu (2014), the mean queue time equation is $E[QT_2] = \frac{\rho_1 E(R_{S1}) + \rho_2 E(R_{S2})}{(1-\rho)(1-\rho_1)}$, where index 1 for high priority job (PM) and 2 for low priority job (customer) and the mean cycle time equation for customers is $E(CT) = E(S_2) + E(QT_2)$.

We model the arrival rate of the PMs as $\lambda_1 = \frac{1}{m_T}$ (since the mean PM cycle is m_T) and the arrival rate of customers is $\lambda_2 = \lambda$. The system loading caused by PMs and customers are $\rho_1 = \lambda_1 E(S_1) = \frac{m_R}{m_T}$ ($E(S_1) = m_R$), $\rho_2 = \lambda_2 E(S_2) = \frac{\lambda}{\mu}$ ($E(S_2) = 1/\mu$), respectively. The total system loading is $\rho = \rho_1 + \rho_2 = \frac{m_R}{m_T} + \frac{\lambda}{\mu}$. The mean cycle time approximation for customer in a G/G/1 queue with TB/NP PM is

$$E(CT\_TBNP) \approx \frac{1}{\mu} + \frac{\frac{m_R}{m_T}\frac{C_{AR}^2+C_R^2}{2}m_R + \frac{\lambda}{\mu}\frac{C_A^2+C_S^2}{2}\frac{1}{\mu}}{(1-\frac{m_R}{m_T}-\frac{\lambda}{\mu})(1-\frac{m_R}{m_T})}. \qquad \text{(A.3)}$$

A.4 Run-Based Non-Preemptive

Wu (2014) treated RB/NP interruption as set-ups and classified it into state-induced set-ups and product-induced set-ups. In this research, we consider only product-induced setups as a model for our PMs. According to Wu (2014), the probability of performing RB/NP interruption after any job is $\frac{1}{N_P}$, where N_P is the number of jobs being processed between successive interruptions on average. Under this assumption, the mean generalized service time is $E(G) = E(S + T_P) = E(S) + \frac{E(P)}{N_P}$, where P is the product-induced set-up times and T_P is the product-induced setup-time experienced by each job.

By treating PMs as RB/NP product-induced set-ups, we obtain a mean cycle time approximation as follow. Calculate the mean downtime as $E(P) = m_R$, $N_P = \lambda\, m_T$ and $E(G) = E(S) + \frac{m_R}{\lambda\, m_T} = \frac{1}{\mu} + \frac{(1-A)}{\lambda}$. The mean queue time equation is the same as the RB/P interruption model, we have

$$E(CT\_RBNP) \approx \frac{1}{\mu A} + \frac{1}{\mu A}\left(\frac{\rho}{1-\rho}\right)\left(\frac{1}{2}\right)(C_A^2 + C_G^2). \tag{A.4}$$

However, the coefficient of variation of the generalized service time is different from the RB/P model. The coefficient of variation of the generalized service time is $C_G^2 = \frac{\frac{C_S^2}{\mu^2} + \frac{m_R^2 C_R^2}{\lambda\, m_T} + \frac{\lambda\, m_T - 1}{m_T^2 \lambda^2} m_R^2}{\left(\frac{1}{\mu A}\right)^2}$.

REFERENCES

Buzacott, J. A., and L. E. Hanifin. 1978. "Models of Automatic Transfer Lines with Inventory Banks – A Review and Comparison." AIIE Transactions 10 (2): 197–207.

Kalir, A. A. 2013. "Segregating Preventive Maintenance Work for Cycle Time Optimization." IEEE Transactions on Semiconductor Manufacturing. 26:1, 125 – 131.

Morrison, J. R., H Kim, and A. A. Kalir. 2014. "Mean Cycle Time Optimization in Semiconductor Tool Sets via PM Planning with Different Cycles: A G/G/m Queueing and Nonlinear Programming Approach." In Proceedings of the 2014 Winter Simulation Conference, edited by A. Tolk, S. D. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 2466-2477. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Wu, K. 2014. "Classification of Queueing Models for A Workstation with Interruptions: A Review." International Journal of Production Research, 52:3, 902-917, DOI: 10.1080/00207543.2013.843799.

AUTHOR BIOGRAPHIES

MINHO LEE is an MS candidate in the department of Industrial and Systems Engineering at KAIST, South Korea. His email address is mncf89@kaist.ac.kr.

JAMES R. MORRISON is an Associate Professor in the Department of Industrial and Systems Engineering, KAIST, South Korea. He holds B.S. degrees from the University of Maryland at College Park, USA and M.S. and Ph.D. degrees in Electrical and Computer Engineering from the University of Illinois at Urbana-Champaign, USA. He is a co-Chair of the IEEE RAS Technical Committee on Semiconductor Manufacturing Automation. His email address is james.morrison@kaist.edu.

ADAR A. KALIR is a Senior Principal Engineer with the Fab/Sort Manufacturing Division of Intel Corporation, Qiriat-Gat, Israel. He holds a B.S. and M.S. in Industrial Engineering from Tel-Aviv University, Israel, and a Ph.D. in Industrial Engineering and Operations Research from Virginia Tech, USA. He also serves as an Adjunct Professor at Ben-Gurion University, Israel. His e-mail address is adar.kalir@intel.com.