Proceedings of the 2013 Winter Simulation Conference
R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, eds

ESTIMATING WAFER PROCESSING CYCLE TIME USING AN IMPROVED G/G/M QUEUE

Roland E.A. Schelasin

Texas Instruments Incorporated
Industrial Engineering
5 Foden Road
South Portland, ME 04106, USA

ABSTRACT

Wafer processing cycle times have been successfully calculated using the basic G/G/m queue by relying on historical data to determine the required variability component. The basic equation was found to work well for a highly utilized factory but provided less accurate results at low factory utilization points. Implementation of an improved G/G/m queue as suggested by existing research has resulted in improved correlation with factory performance even during times of lower factory utilization. An overview of the original implementation is presented, followed by the equation for the improved G/G/m queue, its implementation, and subsequent validation results.

1 INTRODUCTION

Queuing theory has been widely dismissed as an insufficient means to estimate production cycle time in semiconductor manufacturing due to the inherent complexity of the re-entrant and lengthy process flow associated with silicon wafer processing and the inherent difficulty of determining the variability component required by the equations that have been developed. Previous work by Schelasin (2011) has shown that by taking advantage of historical data the Kingman equation can be solved to backward calculate variability representative of known factory states. This variability can then be utilized to forward calculate estimated cycle time variations that would result from changes in equipment utilization due to loading, process or equipment parameter changes.

Initial results showed encouraging correlation with actual factory cycle time performance and were effectively used to assist in making tactical as well as strategic factory management decisions for several weeks. However, when factory utilization decreased cycle time estimates were found to be less accurate than they were during the original period of higher factory utilization. Additional review of already existing research revealed that improvements to the G/G/m queue were suggested by Ward Witt (1993) to address this problem. This paper provides a review of the original work done using the more basic equation for the G/G/m queue, presents the improved equation proposed by Ward Whitt, its implementation, and how its use has improved cycle time estimates when compared to actual factory performance.

2 ORIGINAL WORK DONE USING THE BASIC G/G/M QUEUE

2.1 Original Equation

The original work used the basic equation for the G/G/m queue as found in Hopp and Spearman (2003). Using *Kendall's notation* the Kingman equation is also known as the G/G/m queuing system. The fundamental equation for the wait time in queue for this system is defined as:

978-1-4799-2076-1/13/$31.00 ©2013 IEEE

$$CT_q(G/G/m) = \left(\frac{c_a^2 + c_e^2}{2}\right)\left(\frac{u^{\sqrt{2(m+1)}-1}}{m(1-u)}\right)t_e \quad (1)$$

where

CT_q = time waiting in queue;
c_a = coefficient of variation of inter-arrival times into the queue;
c_e = coefficient of variation of effective process times at each machine;
m = number of machines in the equipment set;
u = utilization of the equipment set; and
t_e = effective process time.

2.2 Original Implementation

In practical applications machine count m and effective process time t_e (consisting of raw process time RPT, Load and Unload time L/UL, and Move time between operations) are known. Utilization u can be calculated using these values as part of standard static capacity modeling algorithms. In today's highly automated factories variability coefficient c_e can be easily calculated from historical data. Variability coefficient c_a remains as the only unknown element in Equation (1) that is difficult to determine. Li et al. (2005) suggest using the Kingman or VUT equation to backward calculate variability V from historical data. The implementation of this approach follows the steps depicted in the overview flow chart in Figure 1.

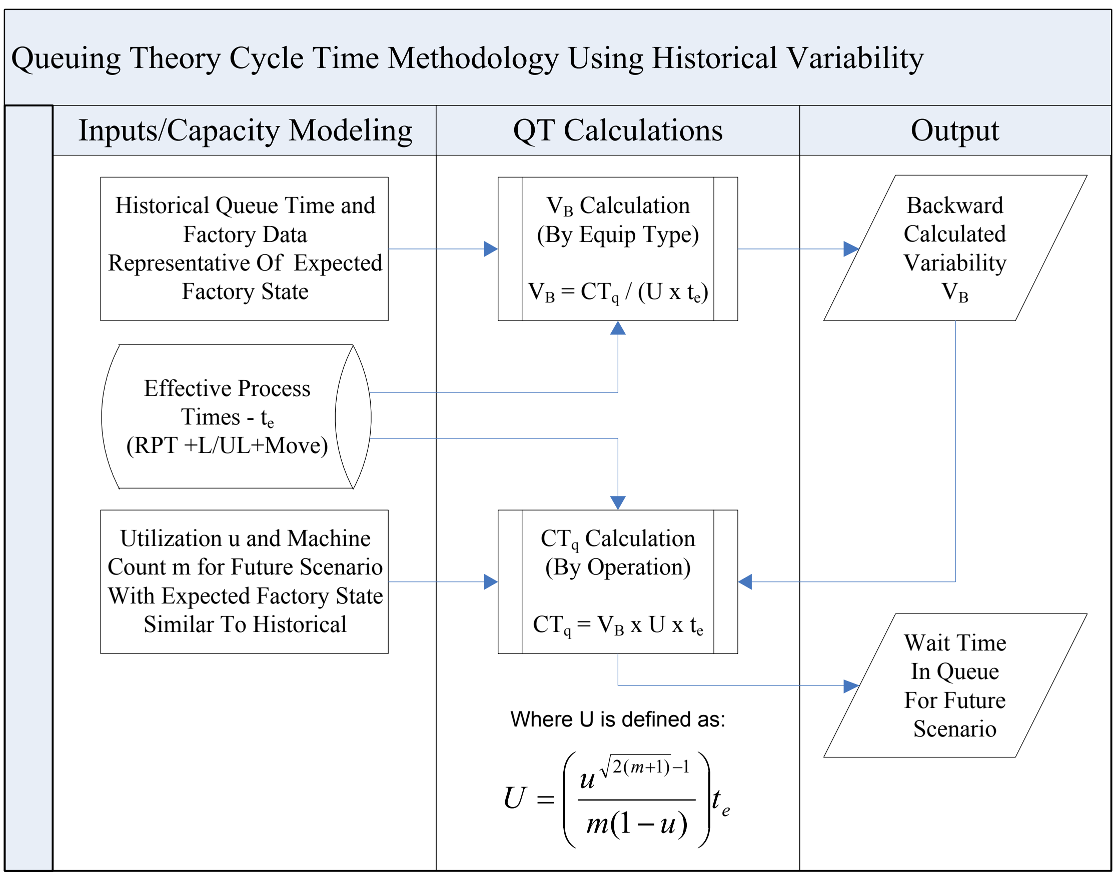

Figure 1: Queuing theory cycle time implementation – Overview.

This methodology is employed to backward calculate each toolset's specific variability component using toolset specific historical machine count, historical toolset utilization, operation specific historical queue times, and operation specific effective process times. Operation specific cycle times are subsequently forward calculated for each production flow operation using the backward calculated toolset specific variability, scenario specific toolset machine counts and utilization, and operation specific effective process times. These operation specific cycle time estimates are then added together for each technology flow to produce overall technology cycle time estimates for the entire process flows through the factory.

In order to account for different equipment types and tool configurations that run the same process, toolset machine counts are adjusted within the static capacity model using an equivalency factor that is based on the model throughput standard for that toolset.

2.3 Original Validation Results

Cycle time estimates for major technologies running in the factory were calculated using the initial equation for the G/G/m queue. Comparison of results over the course of several weeks showed very encouraging correlation to actual factory performance as shown in the first half of the charts in Figure 2.

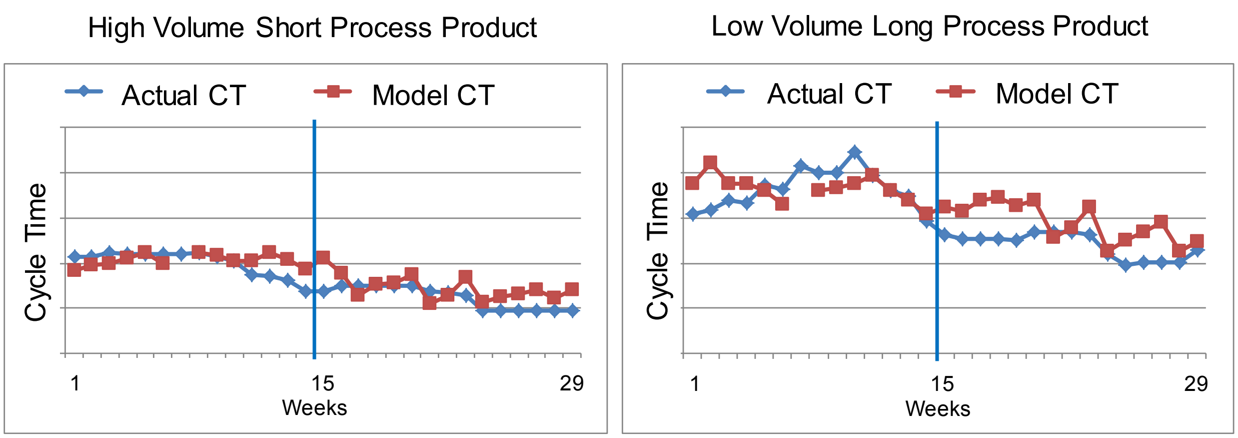

Figure 2: Cycle time validation chart – Initial results.

As can also be seen in Figure 2 cycle time results started to drift in week 10 and did not correlate very well for the second half of the charts starting in week 15. It was found that cycle time estimates became less accurate as factory utilization decreased.

3 THE IMPROVED G/G/M QUEUE

As suggested by Ward Whitt (1993) the fundamental high-traffic G/G/m equation performs well in cases of high machine utilization but results in decreased accuracy when machine utilization is lower. Whitt proposes improvements to the G/G/m queue by adding correction factors based on known closed-form solutions. He applies his own modifications of the accurate approximations for the M/D/m and D/M/m systems developed by Cosmetatos (1975) to obtain a better approximation. The result leads to much improved queue time values for utilizations below seventy percent especially for low machine counts. Both of these circumstances were affecting the factory being modeled when cycle time estimates based on the more fundamental queuing theory equation started to no longer validate as well against actual factory performance as they did initially. The improved equation for the G/G/m queue as suggested by Whitt was used to replace the more basic equation originally implemented and is made up of the components shown in Equations (2) through (9) below:

$$CT_q(\mathrm{u}, {c_a}^2, {c_e}^2, m) \approx \Phi(\mathrm{u}, {c_a}^2, {c_e}^2, m)\, CT_q(G/G/m) \qquad (2)$$

where

$$\Phi(\mathrm{u}, {c_a}^2, {c_e}^2, m) = \begin{cases} \left(\frac{4({c_a}^2 - {c_e}^2)}{4{c_a}^2 - 3{c_e}^2}\right)\Phi_1(m,u) + \left(\frac{{c_e}^2}{4{c_a}^2 - 3{c_e}^2}\right)\psi\left(\frac{{c_a}^2 + {c_e}^2}{2}, m, u\right), & {c_a}^2 \geq {c_e}^2 \\ \left(\frac{{c_a}^2 - {c_e}^2}{2{c_a}^2 + 2{c_e}^2}\right)\Phi_3(m,u) + \left(\frac{{c_e}^2 + 3{c_a}^2}{2{c_a}^2 + 2{c_e}^2}\right)\psi\left(\frac{{c_a}^2 + {c_e}^2}{2}, m, u\right), & {c_a}^2 \leq {c_e}^2 \end{cases} \quad (3)$$

where

$$\Phi_1(m,u) = 1 + \gamma(m,u) \quad (4)$$

and

$$\psi\left(\frac{{c_a}^2 + {c_e}^2}{2}, m, u\right) = \begin{cases} 1, & \frac{({c_a}^2 + {c_e}^2)}{2} \geq 1 \\ \Phi_4(m,u)^{\left[2\left(1 - \frac{{c_a}^2 + {c_e}^2}{2}\right)\right]}, & 0 \leq \frac{({c_a}^2 + {c_e}^2)}{2} \leq 1 \end{cases} \quad (5)$$

where

$$\gamma(m,u) = min\left\{0.24, \frac{(1-u)(m-1)[(4+5m)^{0.5} - 2]}{16mu}\right\} \quad (6)$$

and

$$\Phi_4(m,u) = min\{1, (\Phi_1(m,u) + \Phi_3(m,u))/2\} \quad (7)$$

where

$$\Phi_3(m,u) = \Phi_2(m,u)e^{[-2(1-u)/3u]} \quad (8)$$

where

$$\Phi_2(m,u) = 1 - 4\,\gamma(m,u). \quad (9)$$

4 IMPLEMENTATION OF THE IMPROVED G/G/M QUEUE

As was done in the original implementation, in order to effectively use the improved equation as part of the process to predict CT_q for hypothetical scenarios assuming a similar factory state but with different machine count m and utilization u, the G/G/m queue needs to be backward calculated for variability since only CT_q, t_e, c_e^2, m, and u are known from historical data for actual scenarios. Therefore Equation (3) needs to be solved for c_a^2. Based on experience with the factory being modeled, several approximations can be made to simplify this process.

Historical data for the factory being modeled indicate that c_a^2 is greater than or equal to c_e^2. This reduces Equation (3) to:

$$CT_q(\mathrm{u}, {c_a}^2, {c_e}^2, m) \approx \left(\frac{4({c_a}^2 - {c_e}^2)}{4{c_a}^2 - 3{c_e}^2}\right)\Phi_1(m,u) + \left(\frac{{c_e}^2}{4{c_a}^2 - 3{c_e}^2}\right)\psi\left(\frac{{c_a}^2 + {c_e}^2}{2}, m, u\right). \quad (10)$$

Equation (10) can be further simplified based on historical data. For the majority case in the factory being modeled, the variability term $(c_a^2+c_e^2)/2$ is greater than or equal to one. This allows simplification of Equation (5). After appropriate substitutions of Equations (1) through (9) into Equation (10) the final equation for the improved wait time in queue becomes:

$$CT_q \approx \left[\left(\frac{4({c_a}^2 - {c_e}^2)}{4{c_a}^2 - 3{c_e}^2}\right)\left(1 + min\left\{0.24, \frac{(1-u)(m-1)[(4+5m)^{0.5}-2]}{16mu}\right\}\right) + \left(\frac{{c_e}^2}{4{c_a}^2 - 3{c_e}^2}\right)\right]\left(\frac{{c_a}^2 + {c_e}^2}{2}\right)\left(\frac{u^{\left(\sqrt{2(m+1)}-1\right)}}{m(1-u)}\right) t_e. \tag{11}$$

Since for the factory being modeled c_e^2 is much smaller than c_a^2 additional approximations are made to solve Equation (11) for c_a^2. The resulting Equation (12) provides backward calculated values for c_a^2 that have been found to work well for both low as well as high factory utilization points.

$${c_a}^2 \approx \left(\frac{2\ CT_q}{t_e\left(\frac{u^{\left(\sqrt{2(m+1)}-1\right)}}{m(1-u)}\right)\left(1 + min\left\{0.24, \frac{(1-u)(m-1)[(4+5m)^{0.5}-2]}{16mu}\right\}\right)}\right) - {c_e}^2. \tag{12}$$

Equation (12) can now be used to backward calculate c_a^2 based on known values for CT_q, t_e, c_e^2, m, and u. Using the backward calculated c_a^2 and known values for t_e and c_e^2 allows calculation of CT_q for scenarios with different values of m and/or u, thereby providing a method to estimate the impact on cycle time due to tool count as well as volume and/or mix changes for future factory scenarios. While cycle time estimates will reflect the assumption that the future factory state modeled will be similar to that which produced the back-calculated factory variability, this method has been sufficiently accurate to greatly assist in the decision making processes used in factory capacity management.

5 VALIDATION RESULTS FOR THE IMPROVED G/G/M QUEUE

The charts in Figure 3 show how cycle time estimates compare with actual factory performance both before and after the change to the G/G/m equation used. Data is presented for two of the more prevalent technology flows in the factory.

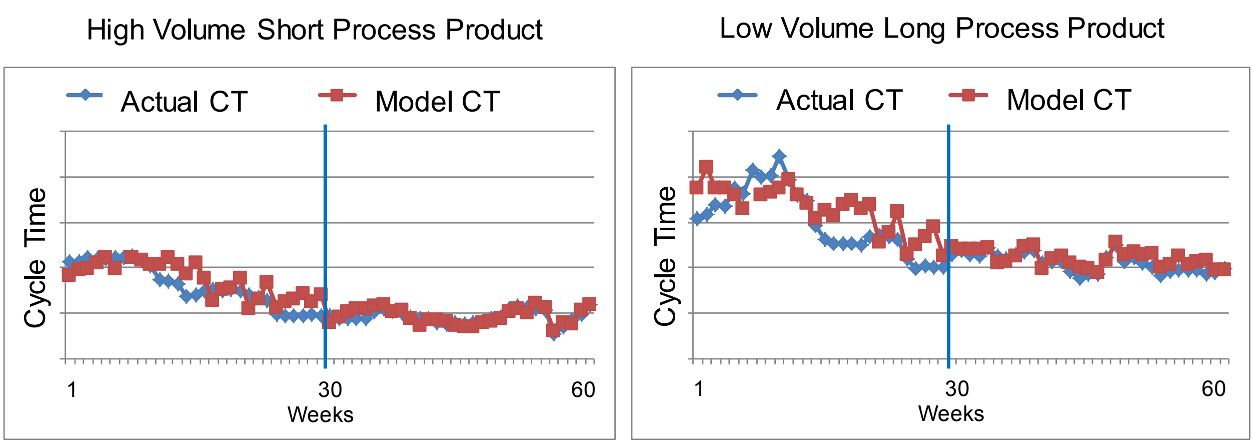

Figure 3: Cycle time validation chart – Initial results and improved G/G/m queue results.

As shown previously initial cycle time estimates using the simpler equation for the G/G/m queue correlated well through week 14 but then drifted away from actual cycle time performance as factory utilization dropped. Cycle time estimates improved significantly once the improved equation for the G/G/m queue as presented in this paper was implemented starting in week 30 indicated by the vertical blue line in each graph.

The largest relative improvement in accuracy was noted in the cycle time estimates for the low volume long process product. This product was running on a newer technology and as a result was running on a number of technology unique toolsets that had a low machine count as well as low utilization.

Figure 4 shows the average percent delta to actual cycle times. As can be seen from this graph the accuracy of the calculated cycle time estimates when compared to actual factory performance not only returned to its original level, but actually improved considerably. The dashed line in Figure 4 starting in week 1 shows how after the initial implementation using the simpler G/G/m queue equations estimated cycle times were on average within just over 6% of actual cycle times. The dotted line starting in week 15 shows how the delta to actuals increased at lower factory utilization to almost 11%. The equations for the improved G/G/m queue were implemented in week 30 as indicated by the vertical blue line. With the implementation of the improved G/G/m queue equations accuracy improved markedly to being on average within just over 3% of actual cycle times. As factory utilization increased starting around week 60 cycle time estimate accuracy continued to be better than before and even further improved.

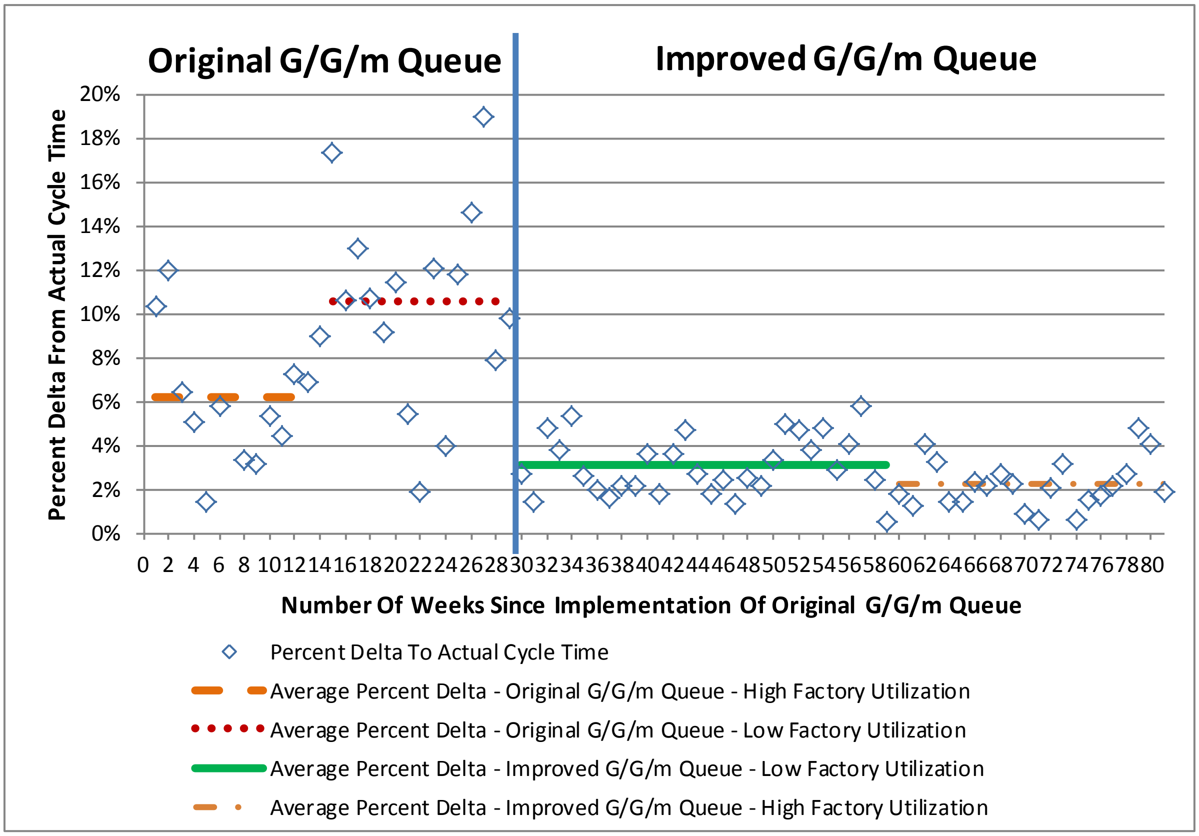

Figure 4: Percent delta between estimated and actual cycle times.

6 CONCLUSION

The original implementation of the basic G/G/m queue algorithms to calculate wafer processing cycle time provided good results when compared to actual factory performance but only as long as factory utilization was high. An improved G/G/m queue equation as suggested by existing research which includes approximations based on closed-form solutions has been successfully used to improve on the accuracy of cycle time wafer processing estimates during time periods of both low as well as high factory utilization.

ACKNOWLEDGMENTS

The author expresses appreciation to the following individuals for their insight and assistance during the development of the methodology outlined, its implementation, validation, application, and refinement:

Paul Fearon, Texas Instruments Incorporated, South Portland

Scott Hopkins, Texas Instruments Incorporated, South Portland

REFERENCES

Cosmetatos, G. P. 1975. "Approximate Explicit Formulae for the Average Queueing Times in the Process (M/D/r) and (D/M/r)." *INFOR* 13: 328-331.

Hopp, W. J., and M. L. Spearman. 2001. *Factory Physics: Foundations Of Manufacturing Management*. 2nd ed., New York: McGraw-Hill, Inc.

Li, N., L. Zhang, M. Zhang, and L. Zheng. 2005. "Applied Factory Physics Study on Semiconductor Assembly and Test Manufacturing." In *ISSM 2005 IEEE International Symposium on Semiconductor Manufacturing Conference Proceedings*, 307-310. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Schelasin, R. 2011. "Using Static Capacity Modeling and Queuing Theory Equations to Predict Factory Cycle Time Performance in Semiconductor Manufacturing." In *Proceedings of the 2011 Winter Simulation Conference*, Edited by S. Jain, R.R. Creasey, J. Himmelspach, K.P. White, and M. Fu, eds., 2045-2054. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Whitt, W. 1993. "Approximations For The GI/G/m Queue." *Production and Operations Management* 2(2): 114-161.

AUTHOR BIOGRAPHY

ROLAND SCHELASIN is the site operational modeling and simulation engineer at Texas Instruments Incorporated (TI) in South Portland, Maine. He started working in the semiconductor industry in 1988 at IBM in Essex Junction, Vermont. Collaborating with IBM Research in Yorktown Heights, New York, he pioneered the development and use of discrete event simulation models for complex equipment within the semiconductor industry and has authored and co-authored several papers on this subject. In 1996 he joined National Semiconductor Corporation (NSC) to manage the modeling and simulation efforts for NSC's first 200MM wafer fabrication facility. He joined TI in 2011 as part of TI's acquisition of NSC and has been elected a Member of Group Technical Staff within TI's Technology and Manufacturing Group. He is a member of the Institute of Industrial Engineers (IIE) and has presented at IIE's annual conference. As part of the Industrial Engineering organization at TI he continues to provide modeling and simulation expertise to develop decision making tools that are used to manage and improve equipment and overall factory effectiveness. Email: Roland.Schelasin@ti.com.