Proceedings of the 1999 Winter Simulation Conference
P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, eds.

A COMPARISON OF THE EXPONENTIAL AND THE HYPEREXPONENTIAL DISTRIBUTIONS FOR MODELING MOVE REQUESTS IN A SEMICONDUCTOR FAB

Siroos Sokhan-Sanj
Gabriel Gaxiola

PRI Automation, Inc.
Automation Planning and Design
1250 S. Clearview Ave. #104
Mesa, AZ 85208, U.S.A.

Gerald T. Mackulak
Fredrik B. Malmgren

Department of Industrial Engineering
Arizona State University
Tempe, AZ 85287-5906, U.S.A.

ABSTRACT

Variability in any manufacturing process negatively impacts performance since it leads to system disruption. Semiconductor manufacturing, with its characteristic re-entrant flow, typically experiences extreme variability. The Automated Material Handling System (AMHS) in a semiconductor fab is subject to this variability and yet must still complete deliveries within a specified time limit. When designing the AMHS the variability used in the simulation model will have a direct impact on the equipment set selected. Sizing a system based on the average case scenario creates a system incapable of meeting the extreme conditions often encountered in reality. The challenge for the modeler of a semiconductor fab is to accurately represent this variability. This paper discusses how the hyperexponential distribution more accurately represents the variability in semiconductor fabs than the typically used exponential distribution.

1 INTRODUCTION

One of the characteristics of the semiconductor industry is the complex re-entrant flow of wafers throughout the fab and the number of process steps required in the production of an Integrated Circuit (IC). These conditions cause high variability in Inter Arrival Times (IAT) between processes. Due to the high cost of semiconductor processing tools, it is important to maintain high utilization levels for these tools. One way of maintaining high tool utilization is to have an efficient AMHS that will insure that the right lot is at the place at the right time.

The AMHS must deliver product within a defined time window. If the lot misses this time window, the tool would starve for material, resulting in lower utilization and hence lower profit margin. The challenge for the modeler is to design an AMHS that can hit the time window no matter the level of congestion in the real fab.

When planning and designing an Automated Material Handling System (AMHS) for a semiconductor facility, advanced modeling tools are very useful. One of the most useful tools is discrete event simulation. However, in simulation as in all types of modeling, good input data is required for the model to produce valid results. An important realization for a model builder is that the output from a model is only as accurate as the input data used to drive it.

One of the most critical input parameters to a simulation model is the distribution used to model entity creations. It has been hypothesized that the exponential distribution provides the proper amount of variability common to most types of arrival processes. In semiconductor fabs this is not the case. Arrivals generated from an exponential distribution exhibit the same average characteristics, but have significantly less variability than actual processes.

For this reason, modelers typically increase the mean used with the exponential distribution to provide a safety factor. This has the effect of increasing bandwidth but not really in a realistic statistical way (upper limit is higher but average number of arrivals is also higher). Using this safety factor might result in an oversized design, and a pessimistic performance prediction. On the other hand, if this safety factor was not introduced, the design could have been undersized and optimistic predictions would have been made. The criticality of a valid and accurate AMHS design is evident when knowing that the cost of a complete AMHS solution for a semiconductor fab (300mm) will be close to $100 million. When planning and designing an AMHS it is essential to use representative distributions and parameters for all of the processes. If the data and distributions used in the modeling effort are inaccurate, the installed system will not perform as predicted in the model.

2 PROBLEM STATEMENT

To validate and further improve the accuracy of the simulation modeling of real fab AMHS, data have been collected and analyzed (see Figure 1). Comparing the number of requested moves per hour from the real fab with the simulated data based on exponential Time Between Creation (TBC) (see Figure 2) illustrates differences. The real fab move requests clearly exhibit much higher variability.

Table 1 shows how the simulated system was able to replicate an average number of moves/hr very close to that observed in the real fab (78.4 moves/hr in the actual fab vs. 79.1 moves/hr in the model). However, the level of variation is quite different (standard deviation of 14.37 in the actual fab vs. 8.42 in the model). Some of the reasons for the high variability are due to the surges in the movement rate caused by operator shift changes and batching of lots. It was concluded that the variability in the modeling stage must be increased to map the real fab environment.

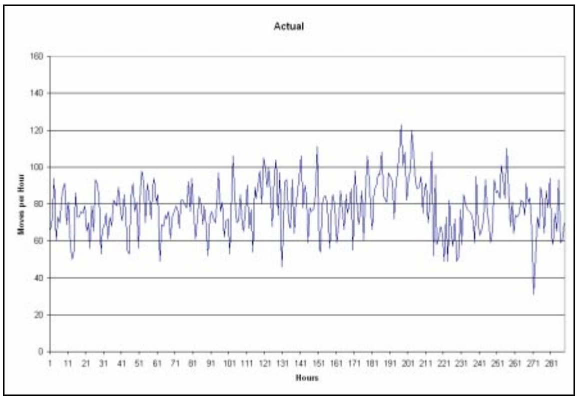

Figure 1: Hourly Moves Performed in Real Fab

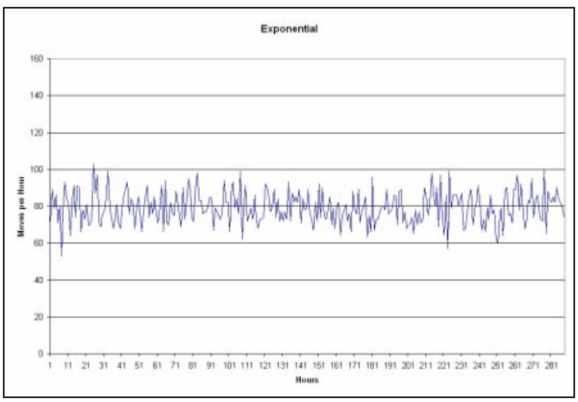

Figure 2: Hourly Moves Obtained Using Exponential TBC

Table 1: Actual vs. Simulated Moves/Hr and Standard Deviation

	Actual	Exponential
Average Moves/Hour	78.04	79.10
St. Dev. Moves/Hour	14.37	8.42

Using the exponential distribution to model the TBC of requests of lots in a semiconductor facility does not induce the correct amount of variability. One of the alternatives used in the analysis of simulation models has been the application of safety factors to account for any variations in the system not originally considered. These safety factors have a big impact on the sizing of the AMHS, and can be hard to justify since these might result in an oversized design. Figure 3 below shows a sample cumulative delivery time chart for systems run with and without safety factors. It can be seen that the performance is seriously impacted by the addition of the safety factor.

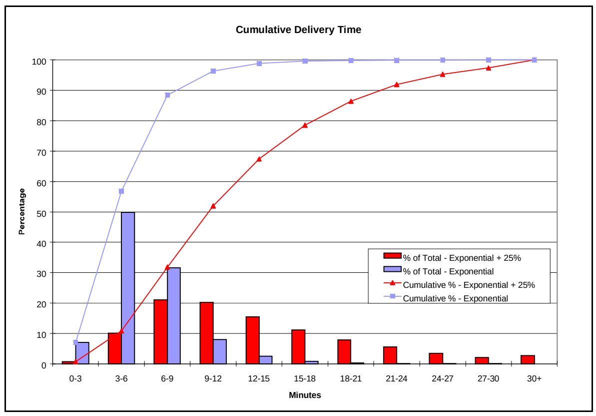

Figure 3: Performance Measure of Simulated Data

Another approach to increase the variability has been to induce a surge in the movement rate. These surges are induced with regular time intervals, representing surges that might occur due to shift changes. Again, the variability was not representative of what was seen from the actual data. In the actual data, no pattern due to shift surges was found.

Other avenues that allow a better representation of the real variability occurring in the fab had to be found.

3 THE SIMULATION MODEL

This study is concerned with evaluation of different distributions with respect to how they affect the AMHS performance and how the simulation results match the actual AMHS performance in the fab. An accurate estimate can be obtained only if the simulation of the

AMHS is an accurate reflection of the actual system. For this reason, PRI Automation created an accurate AutoMod (Phillips 1998) model of their AeroTrak™ interbay delivery system (Colvin and Mackulak 1997).

PRI Automation's AMHS consists of independent vehicles that travel on a track suspended from the ceiling of the fab. The simulation model exactly mimics the control logic software used to control the movement of these vehicles. In this simulator, the requests for moves are created at every stocker and a destination stocker is given for the move. In order to simulate the movement between each stocker pair, random samples are generated from an exponential distribution. Previous studies have shown that the distribution used for generating these move requests have a large impact on simulated performance. (Wu et al. 1999)

3.1 Evaluation of Distributions

A special purpose simulator was created to evaluate different types of distributions for generating movement requests. This simulation was built to model the move request pattern that is created in the AeroTrak™ simulator, and to find a distribution that more closely resembles fab variability.

A hyperexponential distribution is a weighted average of two or more exponential distributions with different mean values (Gross et al. 1985). The probability density function (pdf) for the hyperexponential distribution is the following:

$$fx(x) = \sum_{i=1}^{k} \alpha i \lambda i \, \ell^{-\lambda i x} \qquad (1)$$

$$x \geq 0, \alpha i \geq, \lambda i \geq 0$$

$$\text{Where:} \quad \sum_{i=1}^{k} \alpha i = 1 \qquad (2)$$

One of the characteristics of the hyperexponential distribution is that it has higher variability than the exponential distribution. It has an expectation

$$E\left[\chi^{2}\right] = 2\sum_{i=1}^{k} \frac{\alpha i}{\lambda i^{2}}, \qquad (3)$$

and hence its coefficient of variation (C.V.) is larger than 1. C.V. is an essential measure of variability, and is defined as standard deviation divided by the average (σ / μ). This non-dimensional ratio allows for a consistent comparison among different types of distributions.

In an attempt to mimic the variability associated with actual fab move requirements, a hyperexponential of two weighted exponential distributions was examined. Initially, 50% of the creations were made with a lower mean for the TBC and 50% of the creations were made with a higher mean for the TBC than the mean used in the original exponential distribution. This split results in an exponential pdf with one mean 50% of the time and another mean in the remaining 50%. In essence, numbers are drawn from different exponential distributions. In this case there are two brackets k=2 (with a 50/50 split), but the number of brackets could be higher if additional variability is needed. Additionally, the percentages drawn from each distribution can be weighted as described.

Table 2: Inputs and Outputs from the Special Purpose Simulator

		Simulator Input		Simulator Output		
	Theoretical Mean	Multiplier 1	Multiplier 2	Average	St.Dev	C.V.
Exponential	10	1.00	1.00	9.98	10.02	1.00
Hyper 0.50	10	0.50	1.50	10.00	10.26	1.03
Hyper 0.10	10	0.10	1.90	10.02	13.17	1.31
Hyper 0.05	10	0.05	1.95	10.02	14.85	1.48
Hyper 0.01	10	0.01	1.99	10.03	16.74	1.67

By varying the mean values and ratios, different levels of variability can be obtained. By changing the multipliers the variability can be expanded until the desired level of variability is reached. Table 2 above illustrates this concept and how the C.V. increases when changing the multipliers. Figure 4 below shows the cumulative TBC of the exponential distribution and the hyperexponential distribution with different means. It can be seen that the hyperexponential is a bimodal distribution that introduces more variability than the exponential.

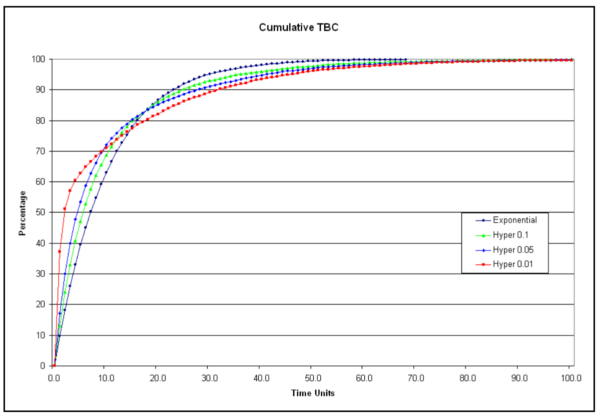

Figure 4: Cumulative Time Between Creation for the Tested Distributions

3.2 Matching Simulation Output with Real Fab Data

To match the simulation output with actual data from the fab, the means of the hyperexponential were varied to obtain a movement pattern similar to the one in the fab. Table 3 below shows the results obtained from these runs. It can be seen that using the two multipliers with values of 0.03 and 1.97 will result in the most similar C.V. for the moves per hour. Using these multipliers also presented similar delivery time statistics for the simulated versus the actual data. With this increased confidence in the simulation model, future modeling of potential expansions or proposed changes of this AMHS will be done using these multipliers.

Table 3: Simulation Results from Effort to Match the Actual Variability

	Actual	Exponential	Hyper 0.05	Hyper 0.04	Hyper 0.03	Hyper 0.02	Hyper 0.01
Average Moves/Hour	78.04	79.10	78.08	77.42	78.12	79.16	78.32
St Dev Moves/Hour	14.37	8.42	13.18	13.53	14.32	14.41	13.88
C.V. Moves/Hour	0.184	0.106	0.169	0.175	0.183	0.182	0.177
Average Delivery Time	5.99	4.92	5.82	5.86	6.03	6.45	6.69
St Dev Delivery Time	2.49	1.59	2.36	2.43	2.63	2.92	2.91

Figure 5 shows the hourly move rates performed by the AMHS in the simulation model. Comparing this chart to the ones in the introduction, it is easy to see that the variation is much higher.

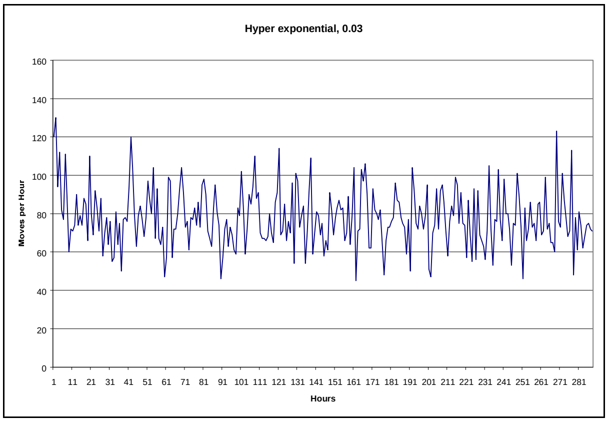

Figure 5: Graph Showing the Hourly Movement Pattern

Further analysis showed that increasing the mean of TBC with a safety factor resulted in an oversized design. Additionally, the predicted performances were much worse than what was seen in the fab.

3.3 Comparison of the Effect of Hyperexponential Versus Other Distribution Types

A previous study (Wu et al. 1999) showed that using different distributions for generation of movement rates results in a significant difference in the simulated performance output. Previously, distributions with the following characteristics were evaluated:

Constant distribution with C.V. = 0.0
Normal distribution with C.V. = 0.5 when $\sigma = 0.5\mu$
Exponential distribution with C.V. = 1.0

A hyperexponential distribution with C.V. = 1.6 was simulated using the same AMHS layout. The delivery time distribution is shown in Figure 6, and it can be seen that the distribution has a longer tail than the other distributions. This is caused by the large variability in the movement rates, which results in a large system surge and subsequent delayed delivery times. A similar analysis was performed using other AMHS systems, and the same general performance characteristic is obtained.

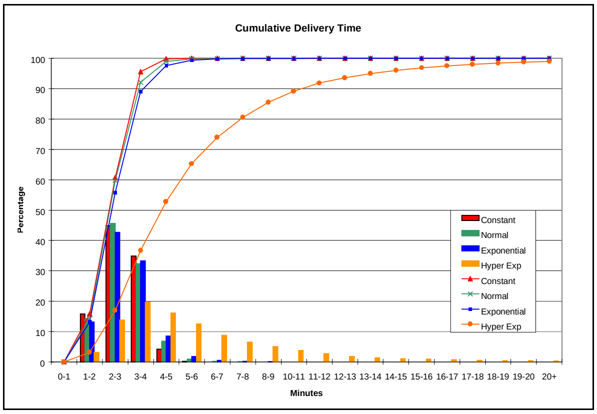

Figure 6: Comparison of System Performance with Different Input Distributions

4 CONCLUSIONS

By evaluating different distributions for lot creations, it has been shown that the hyperexponential distribution induces more variability than the commonly used exponential distribution. The level of variability that the hyperexponential creates can be modified to accomplish the level of variability desired by the modeler. This feature has successfully been utilized to match simulation output with AMHS performance measures collected from a fab. Additionally, by using the hyperexponential, the necessity of using a safety factor has been eliminated.

Future work in this area will be to compare simulation results with actual data from other fabs. A collection of multipliers for use in the hyperexponential will be used to try to find a heuristic for appropriate use of hyperexponential distribution in the design process.

REFERENCES

Colvin, T.D. and Mackulak, G.T. 1997. A Concurrent Engineering Approach for Semiconductor Material Handling System Design: A Case Study. In *Proceedings of the 2nd International Conference on Industrial Engineering Applications and Practice*, ed. J.J. Chen, and A. Mital, 123-128. San Diego, California.

Gross, Donald and Harris, Carl M. 1985. *Fundamentals of Queuing Theory, 2nd edition*. New York: John Wiley and Sons Inc.

Phillips, T. 1998. AutoMod by AutoSimulations. In *Proceedings of the 1998 Winter Simulation Conference*, ed. D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan, 213-218. Washington, DC.

Wu, S., Sokhan-Sanj, S., Mackulak, G.T. and Gaxiola, G. 1999. The Impact of Manufacturing Variability on Automated Material Handling System Performance. In *Proceedings of the 1999 International Conference on Semiconductor Manufacturing Operational Modeling and Simulation*, ed. J. W. Fowler, J. K. Cochran, C. M. Hilton, 81-85. San Francisco, California.

AUTHOR BIOGRAPHIES

SIROOS SOKHAN-SANJ, Senior Simulation Engineer, joined PRI Automation in 1996. He has developed simulation and animation models, and other computer tools, and applied them to the design of AMHS for semiconductor manufacturing fabs. He is currently conducting AMHS design and simulation analyses for 300mm fabs. Mr. Sokhan-Sanj holds a Masters Degree in Industrial Engineering from Arizona State University and has authored/co-authored several papers for industrial engineering and simulation conferences.

GABRIEL GAXIOLA, Senior Industrial Engineer, joined PRI Automation in 1996. He has several years experience in the design and simulation of AMHS for semiconductor manufacturing fabs. Mr. Gaxiola has also held Industrial Engineering positions at Motorola and AMP. He holds a Masters Degree in Industrial Engineering from Arizona State University

GERALD T. MACKULAK is an Associate Professor in the Department of Industrial Engineering at Arizona State University. His research interests center on improvements to simulation methodology, material handling modeling, and production scheduling in high technology manufacturing. He holds BSIE, MSIE and Ph.D. degrees from Purdue University.

FREDRIK B. MALMGREN is a graduate student in the Department of Industrial and Management Systems Engineering at Arizona State University Arizona State University. His is currently working as a research assistant for PRI Automation. His interest of research is simulation of material handling systems, simulation output analysis, and scheduling of semiconductor fab operations. He holds a BSIE from University of Kalmar in Sweden.