Proceedings of the 1996 Winter Simulation Conference
ed. J. M. Charnes, D. J. Morrice, D. T. Brunner, and J. J. Swain

APPLICATION OF A 2-STAGE GROUP-SCREENING DESIGN TO A WHOLE-LINE SEMICONDUCTOR MANUFACTURING SIMULATION MODEL

Theodora Ivanova

Lucent Technologies
9333 S. John Young Parkway
Orlando, FL 32819

Mansooreh Mollaghasemi

Department of Industrial
Engineering and Management
Systems
University of Central Florida
Orlando, FL 32816

Linda Malone

Department of Industrial
Engineering and Management
Systems
University of Central Florida
Orlando, FL 32816

ABSTRACT

The focus of the paper is on the application of an experimental design methodology to a semi-conductor manufacturing simulation model. A complex whole-line simulation model of a semiconductor fab is built. Seventeen input factors are set for investigation through a 2-stage group-screening experimental design. A multiple response regression metamodel is built to define the relationships between the significant input factors and the four response variables of interest. The combination of simulation modeling methods with experimental design and regression analysis techniques allows the development of a flexible tool for capacity analysis of a semiconductor manufacturing facility.

1 INTRODUCTION

As semiconductor companies look for ways to increase their competitiveness, many are turning to simulation modeling to help them control their facilities.

One of the major manufacturers of Application Specific Integrated Circuits (ASIC), has formed an operations research team whose main task is to create simulation models and to assist the company's management in making its future business decisions.

At present, simulation is the only tool that is capable of modeling the complex, often random nature of the semiconductor manufacturing environment. Simulation modeling, however, has certain drawbacks, such as the lack of optimization capability. Also, the simulation model is often referred to as a "black box", because the explicit relationships between its input and output parameters are typically unknown. That is why simulation modeling becomes most effective in combination with other analysis methods, such as experimental design and regression analysis.

Experimental design allows examination of the input factor effects on the system response variables. In cases where the effects of less than 11 input factors are studied, Biles (1984) recommends the application of fractional factorial designs for the simulation experiments. Research presented by Hood and Welch (1990, 1993) shows the application of fractional factorial Resolution III and IV designs in modeling the logistics of semiconductor manufacturing lines. In cases where more than 11 input factors are studied, the recommended type of design is a group-screening design. A 2-stage group-screening procedure was introduced by Watson (1961) and further developed for multiple-stage designs by Patel (1962) and Li (1962). Significant contribution to the group-screening design method has been made by Mauro and Smith with their numerous papers on the robustness and effectiveness of the method (Mauro and Smith 1982, 1984, and Mauro 1984).

Based on the experimental design results, regression analysis equations are built to define the relationships between the input factors and the measures of performance. The regression metamodel concepts were introduced to simulation by Kleijnen (1979). A long-term advocate for the implementation of multiple response regression metamodels to simulation output analysis is Friedman (1984, 1987, 1989).

Although group-screening design combined with regression metamodel analysis appears well suited for the analysis of large-scale semiconductor manufacturing simulation models, there is a limited number of papers dealing with this type of experimental design application. The objective of the present study is to build a whole-line simulation model and to estimate the future Work-In-Process (WIP) levels, cycle times and throughputs for two basic semiconductor products. Further, the most significant input factors for the production measures of performance are to be identified through the application of group-screening design to the simulation model.

The organization of the paper is as follows: in Section Two, an overview of the theoretical aspects of the group-screening experimental design and multiple response regression metamodels is presented. Section Three presents the whole-line simulation model definitions and the output analysis steps. Section Four includes the application of a 2-stage group-screening design to the simulation model and the multiple response regression metamodel analysis. Finally, Section Five summarizes the results from present research.

2 THEORETICAL BACKGROUND

2.1 Two-Stage Group Screening Design

Watson (1961) suggests that the k input factors in a model can be separated into g groups of f factors each, by any method. Each group is then considered as a single factor called group-factor. At the upper level of a group-factor, all factors in that group are at their high levels. The lower level of a group-factor is determined by setting all individual factors at their low levels. If a group-factor is found to be significant, a second stage of the design is set, where the original factors from the significant groups are tested individually. If after the first stage there is still a considerable number of important factors left in the experiment, further regrouping might be applied and the group-screening process will then have more than two stages (Li 1962 and Patel 1962). Kleijnen (1987) recommends keeping the "unimportant" factors at fixed levels during the next stages of the experiment.

2.2 Multiple Response Regression Metamodel

After every group-screening design stage, the most significant input factors are determined by the use of regression analysis. Multiple response regression metamodels relate each system response to the most significant input factors. The simplest multiresponse metamodel is the additive first-order (linear) model. Multiple linear regression equations are typically built after each experimental stage and a global F-test is used to analyze the hypothesis that all regression model coefficients, β_{km}, equal zero. Then, through individual t-tests, the significant input factors are determined. The insignificant factor deletion procedure is iterative, (i.e. one factor is deleted at each step, after which the t-tests are run again). After including only the significant group-factors, a linear regression metamodel is built. The significant input factors are used as individual or group-factors in the second stage of the group-screening design. At the second experimental stage, similar tests are performed to determine the significant input factors and so forth.

3 THE WHOLE-LINE SIMULATION MODEL

3.1 Model Assumptions and Definitions

The ManSim/X simulator, developed by Tyecin Systems Inc., was used to build the fab simulation model. ManSim/X has been specifically designed for capacity analysis and production planning of semiconductor manufacturing facilities.

The whole-line simulation is a model of a 6" semiconductor wafer fab with more than 250 machines and operators, grouped into multiple work areas. Two basic recipes for two products are included in the model. Different operational rules are used to control the interactions between the model elements.

3.2 Model Validation and Output Results

The model validation processes included variable reasonableness tests, conceptual and operational validity tests, comparisons with mathematical models, etc. Based on the WIP autocorrelation functions for five runs, a warm-up period of 90 days was determined. All statistical calculations used in further simulation runs were based on the truncated "steady-state" time series with a length of 270 days. An overall confidence level of 0.80 was set for four system measures of performance, namely cycle times for Product 1 and Product 2, and throughputs for Product 1 and Product 2. By using Bonferroni's inequality, the individual confidence level for each response was set at 0.95.

Queue size analysis showed that implanters, steppers and etchers are the three most critical production facility groups. The comparatively large queue sizes at these workstations, which form even in the case of stable WIP output time-series, remind of the danger that the workstations could easily become a fab "bottleneck" at certain conditions. Therefore, there was a need for further study to identify the factors which are significant for the performance of these three workstations and for the overall factory performance.

4 GROUP-SCREENING EXPERIMENTAL DESIGN

4.1 Group-Screening Design - Stage I

The objective of the group-screening experimental work was to determine the importance of certain input factors on the four simulation model responses, namely, the cycle times for Products 1 and 2, and throughputs for Products 1 and 2. Seventeen input factors were selected at the beginning of the experiment. Fifteen input factors are related to the three most critical wafer fab facility groups, namely implanters, steppers and etchers and two input factors are related to the overall fab performance. Following is a list of the input factors for Stage I screening process:

X1 = MTBF (Mean Time Between Failures) for steppers
X2 = MTBF for implanters
X3 = MTBF for etchers
X4 = MTTR (Mean Time to Repair) for steppers
X5 = MTTR for implanters
X6 = MTTR for etchers
X7 = Lot Dispatch Rule for steppers (the rule by which a lot is chosen from the queue in front of a machine)
X8 = Lot Dispatch Rule for implanters
X9 = Lot Dispatch Rule for etchers
X10 = Number of steppers
X11 = Number of implanters
X12 = Number of etchers
X13 = Operator/machine Ratio for steppers
X14 = Operator/machine Ratio for implanters
X15 = Operator/machine Ratio for etchers
X16 = Lot Release Rule (the rule which organizes the lot release into production)
X17 = Hot Lots percentage for both products.

The seventeen input factors were tested for significance through a group-screening design. By using factor grouping rules (Watson 1961), seven group-factors were formed at the first design stage, as shown in Figure 1.

A two-level fractional factorial 2_{IV}^{7-3} design with 16 design points and 5 replicates for each was planned. A full factorial design was set for the first 4 variables (A, B, C, D). The rest of the design input variables E, F and G were defined as design generators, where

$$E = ABC\ ;\ F = BCD\ ;\ G = ACD \qquad (3)$$

The defining relation for this Resolution IV design is :

$$I = ABCE = BCDF = ADEF = ACDG = BDEG = ABFG = CEFG \qquad (4)$$

Table 1 defines the low and high levels for each group factor. The low level for each factor was chosen to be more constraining to the simulation model compared to the high input factor level. Trial runs were performed to make sure that the model is stable under the low factor level setting. Then, the high factor levels were set as an improvement over the base level for each factor.

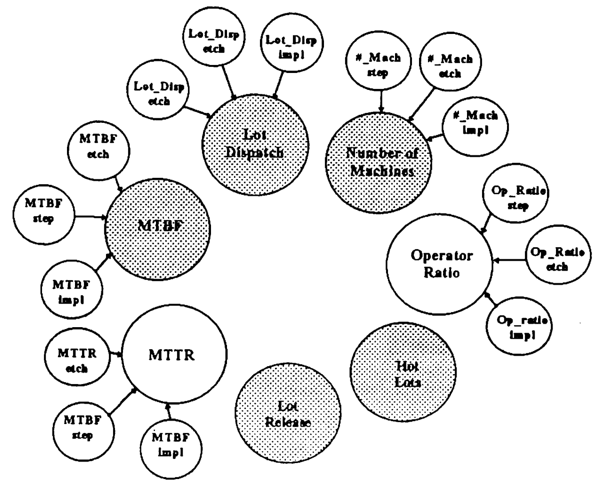

Figure 1: Group-Screening Design - Stage I

This method for setting the low and high factor levels ensures that there is sufficient resource capacity and that the model is stable for all experimental runs (Hood and Welch 1992).

Table 1: Group-Factor Levels

Group - Factor Description	Name	Low Levels (-1)	High Levels (+1)
MTBF	A	base	2 * base
MTTR	B	base	.5 * base
Lot Release Rule	C	FIFO	Fewest Lots at Next Queue
Number of Machines	D	base	base + 1
Operator / machine Ratio	E	base	1.5 * base
Lot Dispatch Rule	F	Random	Constant
Hot Lots	G	10%	5%

As a next step the JMP software, a product of SAS Institute Inc., was used to build a multiple response regression metamodel based on the Stage I results and to determine the significant group-factors. It was assumed that the simulation model output results could be generalized in a linear regression metamodel with no interactions between the group screening factors and no quadratic terms. The qualitative nature of some of the input factors and the narrow range between the low and high factor levels for others, let us maintain the linearity assumption.

The F- test results appear in Table 2. All probability values p are less than 0.05; therefore, all four models are statistically significant. The regression coefficients for the four response variables were tested one at a time using t tests and are presented in Table 3. The shaded regression coefficients have p values of less than 0.05 and are considered statistically significant.

In conclusion, at the end of stage I of the group-screening design, input group factors MTBF, Lot Dispatch Rule, Number of Machines, Lot Release Rule and Hot Lots were declared statistically significant for at least one of the four output responses. Therefore, these group factors were further investigated in Stage II of the experimental design. On the other hand, the MTTR and Operator-to-Machine Ratio group-factors were found insignificant with respect to all responses and were dropped from the next experimental stage.

Table 2: Least-Squares Analysis Table by Response Variable - Stage I

Response Variable	Source	d.f.	Sum of Squares	Mean Square	F Ratio	Prob>F	RSquare
Cycle Time for Product 1 (hrs)	Model	7	21379.22	3054.17	7.86	0.0047	0.87
	Error	8	3108.655	388.58			
	Total	15	24487.88				
Cycle Time for Product 2 (hrs)	Model	7	17057.48	2436.78	7.944	0.0045	0.87
	Error	8	2453.95	306.74			
	Total	15	19511.43				
Throughput Product 1 (wafers)	Model	7	4407.69	629.671	15.939	0.0004	0.93
	Error	8	316.04	39.505			
	Total	15	4723.73				
Throughput Product 2 (wafers)	Model	6	5270.347	878.391	8.681	0.0025	0.85
	Error	9	910.67	101.186			
	Total		6181.018				

Table 3: Linear Regression Metamodel Coefficients - Stage I

			Input Factors							
			Intercept	MTBF	MTTR	Lot Dispatch Rule	Number of Machines	Operator/ Machine Ratio	Lot Release Rule	Hot Lots
			Regression Coefficients							
Response Variable	Average Value	Standard Deviation	$\beta 0$	$\beta 1$	$\beta 2$	$\beta 3$	$\beta 4$	$\beta 5$	$\beta 6$	$\beta 7$
Cycle Time for Product 1 (hours)	762.2	23.55	**762.24**	**-18.84**	-1.53	**-14.23**	**-12.63**	-3.47	-6.71	-4.8
Cycle Time for Product 2 (hours)	698.4	21.31	**698.38**	**-19.05**	-2.99	**-11.01**	**-11.51**	-0.74	-5.56	-5.08
Thruput for Product 1 (wafers)	1294.7	23.41	**1294.7**	2.23	-0.59	**8.51**	-0.59	-2.11	**3.67**	1.44
Thruput for Product 2 (wafers)	2183.8	26.45	**2183.8**	2.08	0.29	3.2	**6.91**	2.775	**6.75**	**8.8**

Note: The shaded cells mark the significant input factors regression coefficients.

4.2 Group-Screening Design - Stage II

In Stage II, the five significant group factors were separated into individual factors. The separation of the three significant group-factors (MTBF, Lot Dispatch Rule and number of machines) and the two single factors (Lot Release Rule and "hot lots") resulted in 11 individual input factors to be examined in the second stage of the experimental design, as follows:
A = MTBF for steppers
B = MTBF for implanters
C = MTBF for etchers
D = Lot Dispatch Rule for steppers
E = Lot Dispatch Rule for implanters
F = Lot Dispatch Rule for etchers
G = Number of steppers
H = Number of implanters
I = Number of etchers
J = Lot Release Rule
K = Hot Lots percentage for both products.
To obtain a Resolution IV experimental design and to minimize the number of simulation runs, a Plackett-Burman design with 24 runs was performed at this stage. Each run was replicated 5 times. A total of 120 simulation runs were performed, equivalent to almost 120 hours of computer run time. As in Stage I of the experiment, the global F-test for the model adequacy indicated that all four models are significant (see Table 4). Table 5 displays the regression coefficients for all input factors, where the significant input factor coefficients are shaded. As shown in Table 5, the MTBF at etchers, the Lot Dispatch Rule at implanters and etchers, the number of machines at steppers and implanters, and the Lot Release Rule have significant positive effects on the two cycle time variables, (i.e. cycle times decrease when these factors are set at their high levels - see Table 1). The "hot lots" percentage does not have a significant effect on the average cycle time for the products, which could be expected. Although the "hot lots" cycle time decreases, the cycle time for the "regular" lots increases, therefore, the average product cycle time does not change. Factors which have significant influence on the throughput levels are the MTBF at implanters, Lot Dispatch Rule on the implanters and etchers, number of machines in the steppers, etchers and implanters groups, the Lot Release Rules and the percentage of Hot Lots.

Table 4: Least-Squares Analysis Table by Response Variable - Stage II

Response Variable	Source	d.f.	Sum of Squares	Mean Square	F Ratio	Prob>F	RSquare
Cycle Time for Product 1 (hrs)	Model	10	22229.76	2222.98	11.49	< .0001	0.896
	Error	13	2514.34	193.41			
	Total	23	24744.11				
Cycle Time for Product 2 (hrs)	Model	10	25557.18	2555.72	11.23	< .0001	0.898
	Error	13	2958.51	227.58			
	Total	23	28515.69				
Throughput Product 1 (wafers)	Model	8	3362.83	420.35	8.26	0.0003	0.815
	Error	15	763.49	50.89			
	Total	23	4126.32				
Throughput Product 2 (wafers)	Model	12	17936.24	1494.69	20.67	< .0001	0.957
	Error	11	795.54	72.32			
	Total	23	18731.79				

Table 5: Linear Regression Metamodel Coefficients - Stage II

			Input Factors											
			Intercept	MTBF_ step	MTBF_ impl	MTBF_ etch	Lot Disp_step	Lot Disp_impl	Lot Disp_etch	Number_ Mach_ step	Number_ Mach_ impl	Number_ Mach_ etch	Lot Release	Hot Lots
			Regression Coefficients											
Response Variable	Average Value	Standard Deviation	$\beta 0$	$\beta 1$	$\beta 2$	$\beta 3$	$\beta 4$	$\beta 5$	$\beta 6$	$\beta 7$	$\beta 8$	$\beta 9$	$\beta 10$	$\beta 11$
Cycle Time for Product 1 (hours)	747	26.2	**747**	-1.32	-1.42	**-7.25**	**-7.04**	-4.15	-5.29	**-8.62**	**-13.11**	3.13	**-11.8**	-5.56
Cycle Time for Product 2 (hours)	684.1	22.1	**684.1**	1.12	-2.81	**-8.72**	**-9.73**	-0.7	**-7.78**	**-11.52**	**-9.58**	-0.33	**-9.26**	-3.46
Thruput for Product 1 (wafers)	1289.2	27.11	**1289.2**	0.41	1.27	-2.23	-0.28	**4.65**	**3.27**	0.56	1.72	**5.03**	**3.34**	0.81
Thruput for Product 2 (wafers)	2181.6	32.63	**2181.6**	-0.49	**7.65**	1.44	-2.08	**-3.88**	**10.46**	**5.02**	**3.87**	**4.57**	**4.41**	**8.62**

Note: The shaded cells mark the significant input factors regression coefficients.

Note that the decrease in the percentage of "hot lots" from 10% to 5% has a significant positive influence on the throughput of Product 2. The only factor which has no significance on any of the four response variables is the MTBF on steppers. The Lot Release Rule, on the other hand, has a significant influence on all four responses. A conclusion could be made that the higher the number of response variables, the harder it becomes to identify factors that are totally insignificant for all response variables.

5 CONCLUSIONS

A whole-line simulation model of an ASIC wafer fab was built and validated. This model is a flexible tool for capacity analysis of the semiconductor manufacturing facility. Additionally, a 2-stage group-screening experiment was designed to study the interactions between the input factors and the multiple measures of performance. The experience with performing group-screening design on a simulation model with multiple responses leads us to believe that although group- screening design is efficient in cases with a large number of input factors and one response variable, it is not as efficient when multiple response variables are involved. At the end of stage I, only two out of the seven group-factors were declared insignificant which brings us back to the still considerable number of eleven individual factors at the second stage. Therefore, it could be concluded that the greater the number of response variables, the less efficient the group-screening design method becomes.

REFERENCES

Biles, William E. 1984. Design of Simulation Experiment. *Proceed. of the 1984 Winter Simulation Conference*, 99 - 104.

Friedman, Linda Weiser. 1987. Design and Analysis of Multivariate Response Simulations: The State of the Art. *Behavioral Science*: 32, 138 - 148.

Friedman, Linda Weiser. 1984. Multivariate Simulation Output Analysis: Past, Present and Future. *Proceed. of the 1984 Winter Simulation Conference*, 277 - 279.

Friedman, Linda Weiser. 1989. The Multivariate Metamodel in Queuing System Simulation. *Computers and Industrial Engineering:* 16, No. 2, 329 - 337.

Hood, Sarah J. and Peter D. Welch. 1992. Experimental Design Issues in Simulation with Examples from Semiconductor Manufacturing. *Proceed. of the 1992 Winter Simulation Conference*, 255 - 313.

Hood, Sarah J. and Peter D. Welch. 1990. The Application of Experimental Design to the Analysis of Semiconductor Manufacturing Lines. *Proceed. of the 1990 Winter Simulation Conference*, 303 - 309.

Kleijnen, Jack P. C. 1979. Regression Metamodels for Generalizing Simulation Results. *IEEE Trans. on Systems, Man, and Cybernetics:* 9, No. 2, 93 - 96.

Kleijnen, Jack P. C. 1978 *Statistical Tools for Simulation Practitioners*, Marcel Dekker Inc.

Li, C. H. 1962. A Sequential Method for Screening Experimental Variables. *Journal of American Statistical Association:* 57, 455 - 477.

Mauro, Carl A. and Dennis E. Smith. 1982. Factor Screening in Computer Simulation. *Simulation*, 49 - 54.

Mauro, Carl A. 1984. On the Performance of Two-Stage Group Screening Experiments. *Technometrics*: 26, No. 3, 255 - 264.

Patel, M.S. 1962. Group-Screening with more than two stages. *Technometrics*: 4, 209 - 217.

Watson, G. S. 1961. A Study of the Group Screening Method. *Technometrics*: 3, No. 3, 371 - 388.

AUTHOR BIOGRAPHIES

THEODORA IVANOVA is a member of Technical Staff in the Manufacturing Systems Engineering Department at Lucent Technologies in Orlando. She received a B.A. in Electrical Engineering from St. Petersburg, Russia, and a M.S. degree in Operation Research from the University of Central Florida in Orlando, Florida. Her present research interests include simulation modeling and analysis, production planning and control of semiconductor manufacturing system, and experimental design. She is a member of IEEE.

MANSOOREH MOLLAGHASEMI received her Ph.D. in Industrial Engineering from the University of Louisville in 1991. She also holds a B.S. and an M.S. in Chemical Engineering from the University of Louisville. She is currently an Associate Professor in the Department of Industrial Engineering at the University of Central Florida. She is actively involved in funded research through Lucent Technologies, Rhone-Poulenc Rorer Pharmaceuticals, and the US Army. Her research interests involve simulation modeling and analysis, multiple criteria decision making, decision support systems, and multiple response simulation optimization. She is a member of IIE and INFORMS.

LINDA C. MALONE is an associate professor in the Industrial Engineering Department of the University of Central Florida. She earned a B.S. in mathematics from Emory and Henry College, a M.S. in mathematics from the University of Tennessee in Knoxville, and a Ph.D. in statistics from Virginia Polytechnic Institute and State University in 1979. She is an Associate Editor of the *Journal of Statistical Computation and Simulation*. Her major interests are in Design of Experiments, Response Surface Analysis, and Regression. She has a number of publications and has co-authored an introductory statistics text book.