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

COMPARISON OF A TWO-STAGE GROUP-SCREENING DESIGN TO A STANDARD 2^{k-p} DESIGN FOR A WHOLE-LINE SEMICONDUCTOR MANUFACTURING SIMULATION MODEL

Theodora Ivanova

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

Linda Malone
Mansooreh Mollaghasemi

Department of Industrial Engineering and
Management Systems
University of Central Florida
Orlando, FL 32816, U.S.A.

ABSTRACT

The focus of the paper is on the comparison of results obtained using and not using group screening in an experimental design methodology applied to a semiconductor manufacturing simulation model. A whole-line simulation model of a semiconductor fab is built. The model includes more than 200 tools used in manufacturing 2 products with around 250 steps each. Output analysis results for the equipment utilization and queue sizes have identified the three most critical equipment groups in the fab. Seventeen input factors are set for investigation through a 2-stage group-screening experiment and a fractional factorial using all 17 factors. The result illustrates that the final models can be quite different. While group screening used with simulation can be an appealing, flexible, tractable tool for capacity analysis of a semiconductor manufacturing facility, one must be concerned with the fact that the two techniques can give different answers to the users. Additionally, researchers need to address the proper choice of significance level for group screening.

1 INTRODUCTION

As semiconductor companies look for ways to increase their competitiveness, many are turning to mathematical and simulation modeling to help them gain control of 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 and mathematical models and assist the management of the plant in making its business decisions. Management would like to optimize several important factory performance measures, including Work-In-Process (WIP) levels, product cycle times, product throughputs, equipment utilization and idle times. These measures of performance are, however, often conflicting with each other. Hence, the company's goals must be coordinated so that the best overall compromise would be achieved. To succeed with that, the Lucent Technologies Microelectronics plant management has decided to use simulation technology.

Simulation was chosen as a tool for modeling the real world system for several reasons. Simulation modeling allows the representation of complex systems consisting of hundreds of deterministic and stochastic elements, which are elaborately related with each other. After its creation the simulation model can be used for experimentation. Simulation modeling also helps perform a sensitivity analysis on the system and answer "What-if " type of questions concerning the influence of input parameters on the output measures of performance. The result of this effort is of concern, however, if the variables identified in the Response Surface Analysis depend upon the technique used. Also, the end result can be very different depending upon the significance level chosen at Stage I of the group screening.

Simulation modeling by itself lacks optimization capability. The simulation model is often referred to as a "black box", because explicit relationships existing between input and output parameters are typically unknown. Thus, 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 is introduced by Watson (1961) and further developed for multiple-stage designs by Patel (1962) and Li (1962). Mauro and Smith have made significant contributions to the group-screening design method in numerous papers on the robustness and effectiveness of the method (Mauro and Smith 1982, Mauro 1984, Mauro and Smith 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. Kleijnen (1979) introduces regression metamodel concepts to simulation. A long-term advocate for the implementation of multiple response regression metamodels as part of simulation output analysis is Friedman (Friedman, 1984, Friedman, 1987, Friedman, 1989).

Although group-screening combined with regression metamodel analysis is a technique well suited for use in large-scale semiconductor manufacturing simulation models, there are a limited number of papers dealing with this type of experimental design analysis of semiconductor manufacturing facilities. Comparison of group screening and the usual 2^{k-p} results have not been done. Further, a full discussion of the importance of the choice for the significance level has not taken place. A whole-line simulation model for the purpose of estimating the future Work-In-Process (WIP) levels, cycle times and throughputs for two basic semiconductor products in one of Lucent Technologies manufacturing facilities was built. Further, the most significant input factors for the production measures of performance are to be identified through the application of group-screening and response surface methodology applied to the simulation model. The purpose of the present study is to show that the results obtained are not necessarily the same and depend heavily on the choice of α at each stage of the process. In order to illustrate this, only one of the response variables was considered for only one product.

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. It also discusses the model validation process and the output analysis steps. Section Four shows the application of a 2-stage group-screening design to the simulation model metamodel analysis. Section Five contains the analysis not using group screening; and finally, Section Six summarizes the results from the present study.

2 THEORETICAL BACKGROUND

2.1 Two-Stage Group Screening Design

The objective of a factor screening strategy is to detect as many important factors as possible in as few runs, as possible. One of the most efficient experimental design techniques satisfying these objectives is the group-screening experimental design. Watson (1961) suggests that the k input factors in the 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 level. The lower level of a group-factor is determined by setting all individual factors at their lower levels. This way of setting the group-factor levels and the assumption that there are no interactions present (the k individual factors have main effects only) ensures that no cancellation of effects could occur, i.e. a group-factor gives a non-zero effect. 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.

Several rules of thumb should be considered when using a group-screening experimental design technique in order to avoid cancellation of factors and to detect as many of the effective factors as possible:

(1) A factor with an unknown direction of effect should be placed alone in a group.
(2) Factors with assumed important positive effects should be placed in one group.
(3) Factors with assumed small effects and the same direction should be placed in a group.
(4) Factors with possible effects and the same direction should be placed in a group.
(5) Resolution IV design should be used to calculate main effects unbiased by possible second-order interactions.

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. Regression metamodels can also be built relating each response to the most significant input factors. These metamodels are referred to as model of the model (Kleijnen 1979). The simplest metamodel is the additive first-order (linear) model. Linear regression equations are built after each experimental stage and hypothesis testing is done to determine significant groups/variables. The input factors determined to be "significant" through t tests are used as individual or group factors in the second stage of the group-screening design. At the end of the second experimental stage similar tests are performed to determine the significant input factors and so forth. Group-screening designs can have more than 2 stages and the last stage is the one in which no groups are present and all factors are studied individually.

3 THE WHOLE-LINE SIMULATION MODEL

3.1 Model Assumptions and Definitions

The wafer fab simulation model consists of the following elements: (1) equipment, (2) labor (operators), (3) products, (4) processes for each product, and (5) operating rules that control the interactions between the elements. Two basic process flows and two products are included in the simulation model, a "linear" process and a "digital" process, respectively. The one analyzed for comparison of the two statistical modeling techniques was the "linear" process. The ManSim/X simulator, developed by Tyecin Systems Inc., is used to build the simulation model. ManSim/X has been specifically designed for capacity analysis and production planning of semiconductor manufacturing facilities. The whole-line simulation presents 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 simulation. Different operational rules are used to control the interactions among the model elements.

3.2 Output Analysis Results

At the beginning of the output analysis, a procedure was followed for determining the length of the transient (warm-up) period in a steady-state type of simulation. Five 360 days long simulation replicates were run using different random number seeds. 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 the four system measures of performance, namely cycle time for product "linear", cycle time for product "digital", throughput for product "linear" and throughput for product "digital". Thus, by using Bonferroni's inequality, the individual confidence level for each response was set at 0.95. Initial tests showed that 5 replication runs were enough to achieve the desired error of +/- 5% of the mean for each response variable.

Queue size analysis identified the three most critical production facility groups, namely, implanters, steppers and etchers. The comparatively large queue sizes at these workstations, which formed even in the case of stable WIP output time-series, remind of the danger that the workstations could easily become a fab "bottleneck" in certain situations. Therefore, there is 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 - Stage I

The objective was to determine the importance of input factors. The response considered herein is throughput (number of wafers out) for product "linear". Throughput was calculated to assess the capability to start a certain number of wafers a week and to obtain the same number of wafers out of the line after a certain period of time.

Seventeen input factors were selected at the beginning of the experiment. Fifteen of these are related to the three most critical wafer fab facility groups. As already mentioned, these three facility groups are implanters, steppers and etchers. Two additional input factors related to the overall fab performance were taken into consideration. Following is a list of the input factors at the beginning of the 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 ($MTTR_s$)
X5 = MTTR for implanters ($MTTR_i$)
X6 = MTTR for etchers ($MTTR_e$)
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 (how many operators are available at each machine)
X14 = Operator/machine Ratio for implanters
X15 = Operator/machine Ratio for etchers
X16 = Lot Release Rule (rule which organizes the lot release into production)
X17 = Hot Lots percentage for both products.

Our first approach was to test the seventeen input factors 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.

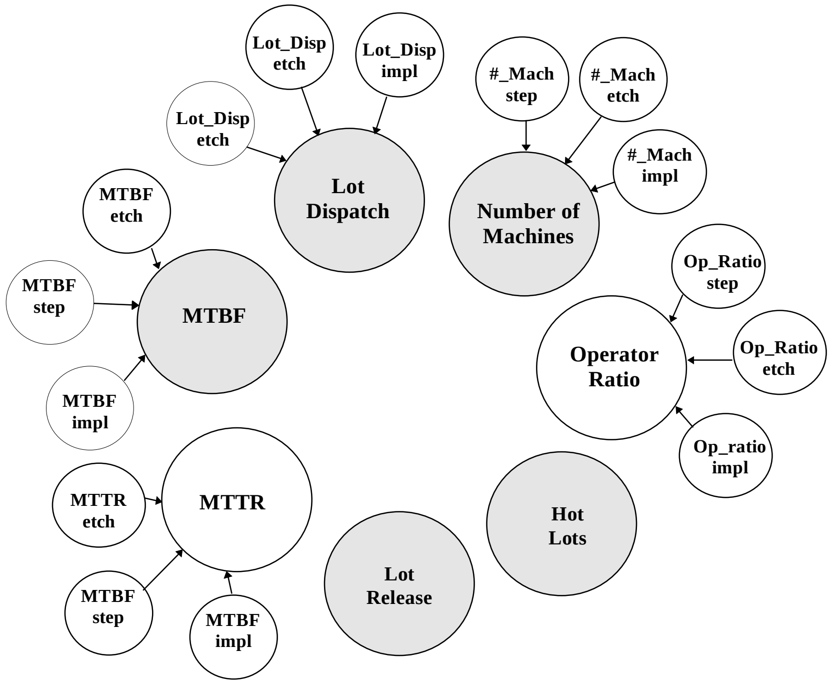

Figure 1: Group-Screening Design - Stage I

A two-level fractional factorial (2^{7-3}_{IV}) design with 16 runs and 5 replicates for each design run was planned. The design generators for this resolution IV design are:

$I = ABCE = BCDF = ACDG.$

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 levels. Trial runs were performed to make sure that the model was stable under the low factor level setting. The high factor levels were set as an improvement over the base level for each factor. 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 the experimental design runs (Hood and Welch 1992).

The next step involves building a regression metamodel based on the Stage I experimental results and determining the significant group-factors. The JMP software, a product of SAS Institute Inc., was used for this purpose. It is assumed that the simulation model output results can be generalized in a linear regression metamodel with no interaction 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 are the published guidelines for maintaining the linearity assumption.

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 of 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%

The response values for the 32 design runs are tested for normality by using histograms, probability plots and the Shapiro-Wilkinson test. At a level of significance of .05, none of the group screening variables is significant. This might lead one to believe that the variables modeled in the simulation and measured in the experiment do not affect the throughput of product linear. However, because this is an early stage of screening, one may decide to be more liberal in choosing the level of significance and decide to use .10 or even .15. Obviously this choice is quite important because it affects not only the results of this step but also the final model.

Using .10 as the significance level, the only group-screening variable that is significant is the Operator Machine Ratio variable. On the other hand, if one uses .15 as the significance level, two additional group screening ariables become significant, Mean Time to Repair and Lot Release Rule.

In conclusion, at the end of Stage I of the group screening design, one could have 0, 1, or 3 group-screening variables to further analyze depending upon the choice of level of significance. These group factors were further investigated in Stage II of the experimental design. The other group variables were dropped from analysis in the next stage of the experiment.

4.2 Group-Screening Design - Stage II

In Stage II, the significant group-factors were separated into individual factors and a second 2^{k-p} design was run. Using a level of significance of .10 in Stage I, the variables considered were:

H= Operator to Machine Ratio for Steppers
I= Operator to Machine Ratio for Implanters
J= Operator to Machine Ratio for Etchers

Of these variables, variable H was determined to be insignificant using α=.05.

Using an α=.15 in the group screening portion results in adding to the above three variables

K= Mean Time to Repair for Steppers
L= Mean Time to Repair for Implanters
M= Mean Time to Repair for Etchers
N= Lot Release Rule

A 32 run experiment was conducted and the model fit allowed for the interactions between the Operator Machine Ratio, Mean Time to Repair, and the Lot Release Rule variables. In this case, using α= .05, the significant variables were Operator to Machine Ratio for Implanters, Operator to Machine Ratio for Etchers, and Lot Release Rule as well as three interaction terms. The interaction terms were Mean Time to Repair for Implanters with Lot Release Rule, Mean Time to Repair for Etchers with Operator to Machine Ratio for Implanters, and Mean Time to Repair for Etchers with Operator to Machine Ratio for Etchers.

5 ANALYSIS WITHOUT GROUP SCREENING

A 64 run, 2^{k-p} design was run using all 17 of the original variables identified in Section 4.1 as well as some of the two factor interactions between the variables. Because of the size of the design, not all of the interactions could be fit. The experimenters used best subset regression in MINITAB to determine the set of variables that were "most" significant. The best model contained five or six terms. The five variable model was derived using α=.05 and contained Operator to Machine Ratio for Implanters, Lot Release Rule, and Hot Lots plus the following interactions: Mean time to Repair for Etchers by Operator to Machine Ratio for Implanters and Mean Time to Repair for Steppers by Lot Release Rule. The six variable model was derived using α=.10 and contained all of the previous variables plus the interaction of Lot Release Rule by Hot Lots. It is interesting to note that the best models using Group Screening and not using group screening each had six variables, three of which were the same and three of which were different. Note that using α=.15 does not add any variables to the equation for the 2^{k-p} method.

Table 2 provides a comparison of the two methods as well as illustrating the role of the significance level. Besides the obvious fact that the models differ when using Group Screening and 2^{k-p} methods and that the choice of variables depends on the choice of the significance level, it is very important to note the last two columns of the table. Using α=.10 for stage 1 of group screening (and .05 for stage 2), results in an R^2 of 6.8. Using .15 for the stage 1 significance level (and .05 for stage 2) dramatically improves R^2 to 26. However, not using group screening results in R^2 of 35.3 using a .05 significance level and an R^2 of 37.9 using α=.10. These values are quite a bit better. Looking at the $(MSE)^{1/2}$ values, we don't see as dramatic a difference as with R^2 but one can observe that the values are uniformly larger when using group screening.

6 CONCLUSIONS

A whole-line simulation model of an ASIC wafer fab was built and validated. A 2-stage group-screening experiment and a 2^{k-p} were designed to study the efficacy of the two procedures and to investigate the proper choice of α. In order to make the comparisons fair, the same total number (64) of experimental runs were used for each of the procedures. Although the group screening design is perhaps the only possible approach to some problems having large numbers of input factors, one should be aware

Table 2: Results of Using Versus Not Using Group Screening

		o/m_i	o/m_e	Lot Release	$MTTR_i \times$ Lot Release	$MTTR_e \times o/m_i$	$MTTR_e \times o/m_e$	Hot Lots	$MTTR_s \times$ Lot Release	Lot Release $\times$ Hot Lots	R^2	MSE
$\alpha =$ 0.05	Group	NONE										
	2^{k-p}	X		X		X		X	X		35.3	24.94
$\alpha =$ 0.10	Group	X	X								6.8	29.18
	2^{k-p}	X		X		X		X	X	X	37.9	24.64
$\alpha =$ 0.15	Group	X	X	X	X	X	X				26.0	26.90
	2^{k-p}	X		X		X		X	X	X	37.9	24.64

that the result may not be the same as if one had not used group screening. Additionally, it appears that one should carefully consider the appropriate level of significance to use, particularly at the Stage I Group Screening portion of the analysis. Because one usually would like to be conservative and not delete possibly significant variables, and based on the results of this example, we would recommend that experimenters use a significance level of .15 in the Stage I Group Screening stages of the experiment. It seems appropriate to switch to .05 (or .10) in the latter stages when enough group variables have been eliminated so that one can look at individual variables and possibly their interactions.

REFERENCES

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

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

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

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

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

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

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

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

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

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

Mauro, Carl A. and Dennis E. Smith, "Factor Screening in Simulation: Evaluation of Two Strategies based on Random Balance Sampling", Management Science, Vol. 30, No. 2, February 1984, 209 - 221.

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

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

Watson, G. S., "A Study of the Group Screening Method", Technometrics, Aug. 1961, Vol. 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.

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. 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 textbook.

MANSOOREH MOLLAGHASEMI is an associate professor of Industrial Engineering and Management Systems at the University of Central Florida. She 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 same University. Her research interests involve simulation modeling and analysis, multiple response simulation optimization, and neural networks. She is a member of IIE and INFORMS.