ESTIMATING CLEARING FUNCTIONS FROM SIMULATION DATA

N. Baris Kacar
Reha Uzsoy

Edward P. Fitts Department of Industrial and Systems Engineering
Campus Box 7906
North Carolina State University
Raleigh, NC 27695-7906, USA

ABSTRACT

We examine the problem of estimating clearing functions that relate the expected output of a production resource as a function of its expected workload, from empirical data. We use a simulation model of a scaled-down wafer fabrication facility to generate the data and evaluate the performance of the resulting clearing functions. We compare several different regression approaches and report computational results.

1   INTRODUCTION

The use of optimization models for production planning has a long history going back to the seminal work of Modigliani and Hohn(1956) and Holt et al(1960). A critical issue in these models is how to represent the relationship between the pattern of work releases into the system over time and the resulting output. Understanding this relationship lies at the heart of any production planning model, since we wish to manage the release of orders and raw material into the system in order to coordinate supply and demand in an optimal manner. A critical component of this relationship is the cycle time, the time between work being released into the system and its emerging as finished product.

However, the modeling of this relationship has proved quite challenging. It is widely recognized that the behavior of most capacitated production resources is governed by queuing, as work arrives randomly over time at resources whose processing capabilities are subject to varying degrees of uncertainty. Queuing models (Hopp and Spearman 2001) have shown a nonlinear relationship between resource utilization and cycle time. Resource utilization, in turn, is determined by the work release decisions made by the planning models. We are thus faced with a circularity, where a planning model needs to be aware of cycle times in order to plan effectively, but the cycle times are determined by the decisions made by the planning decisions itself.

The majority of the production planning literature has taken one of two basic approaches to addressing this issue. By far the most common is to treat cycle times as an exogenous parameter that is unaffected by resource utilization. This approach is taken in the widely used Material Requirements Planning (MRP) procedure (Vollmann, Berry et al. 2005) and most linear and integer programming models (Johnson and Montgomery 1974; Voss and Woodruff 2003). The alternative approach is the use of either detailed scheduling models (Dauzère-Péres and Lasserre 1994) or simulation models in an iterative framework with a mathematical programming model. The optimization model will begin with initial estimates of either cycle times or a related parameter relating the releases to the production system in a given period to future output. The schedule of work releases obtained from the optimization model is then simulated in order to estimate the cycle times or output patterns that would be realized from the given release schedule. The parameters used in the optimization model are then updated, and the process repeated until some form of convergence is obtained. Methods of this type have been proposed, with varying de-

grees of success, by Byrne and Bakir(1999), Byrne and Hosain(2005), Hung and Leachman(1996), Hung and Cheng(2002) and Kim and Kim(2001), among others. Irdem et al.(forthcoming) present a computational comparison of two of these methods, those of Hung and Leachman(1996) and Kim and Kim(2001).

An alternative approach that has shown considerable promise in recent years(Missbauer 2002; Asmundsson, Rardin et al. 2006; Pahl, Voss et al. 2007; Asmundsson, Rardin et al. 2009; Missbauer and Uzsoy forthcoming) is the use of nonlinear clearing functions that represent the relationship between the expected output of a production resource and some measure of the work in process inventory (WIP) at the resource during the planning period. The advantage of these models is that they can capture the nonlinear relationship between resource utilization and cycle times without requiring time-consuming simulation runs as part of the planning process. However, the use of this approach requires effective methods of estimating the clearing functions themselves. Previous work (Asmundsson, Rardin et al. 2009) has shown that the estimation of clearing functions from empirical data is a nontrivial exercise for which a firm theoretical foundation still remains to be established. In this paper we use a simulation model as a surrogate for a production system to generate data from which clearing functions can be estimated. We explore the use of several different linear regression models for estimating clearing functions and compare the performance of the production plans obtained in order to obtain insight into the structure of the clearing functions and the degree to which multiple regression can be used for this purpose.

2   PREVIOUS RELATED WORK

A comprehensive review of the literature on production planning in general, and clearing functions in particular is beyond the scope of this paper. An extensive discussion of the state of the art in this area is given by Missbauer and Uzsoy(forthcoming) and Pahl et al.(2005; 2007).
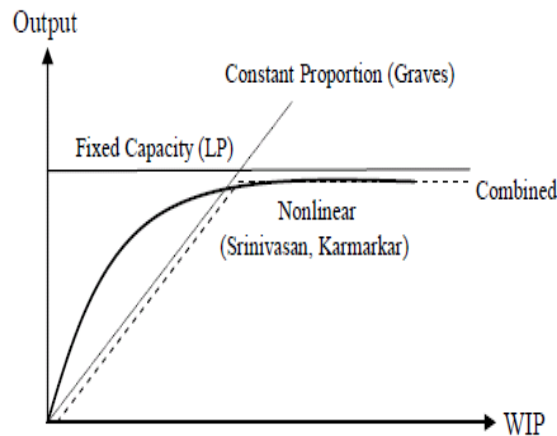

Figure 1: Examples of Clearing Functions (Karmarkar 1989)

Several examples of clearing functions discussed in the literature are depicted in Figure 1. The "Fixed Capacity" function places a fixed upper bound on production, assuming instantaneous production regardless of WIP level. Graves(1986) proposes a clearing function in the form of $X_t = \alpha\, W_t$ where output X_t is a linear function of WIP and W_t denotes some reasonable measure of WIP. This "constant proportion" function assumes the system can be operated to maintain a fixed lead time at all reasonable utilization levels, and may allows unlimited production. This function may yield infeasible levels of output at high WIP levels unless limited by a fixed upper bound on output, as shown in the "combined" clearing function. Karmarkar(1989) proposes a nonlinear clearing function where output increases as a concave non-decreasing function of W_t, reaching an asymptotic maximum. Srinivasan et al.(1988) propose another clearing function that is a concave, non-decreasing function of WIP. Figure 1 illustrates these latter functions as the "effective" clearing function.

Clearing functions can be derived analytically using steady-state or transient queuing models, or estimated empirically from empirical data. Different authors have implemented somewhat different approaches. Asmundsson et al.(2009) formulate the clearing function as a relationship between the expected throughput of a resource in a planning period as a function of the time-average WIP level at the resource during the period. Other authors, such as Karmarkar(1989) and Missbauer(2002) assume the clearing function depends on the expected workload, which they define as the sum of the work in process available at the start of the period and the material released during the period. In addition, two different functional forms have been suggested to estimate clearing functions from empirical data. Karmarkar(1989) proposes the functional form

$$f(\hat{W}) = \frac{K_1 \hat{W}}{K_2 + \hat{W}}$$

where $\hat{W}$ denotes a measure of workload, while Srinivasan et al. (1988) suggest the form

$$f(\hat{W}) = K_1(1 - e^{-K_2 \hat{W}})$$

In both forms, K_1 represents the maximum possible output in a period and K_2 is a user-specified parameter controlling the curvature of the clearing function.

While a number of authors have proposed the use of queuing models to determine clearing functions analytically (Srinivasan, Carey et al. 1988; Selçuk , Fransoo et al. 2007; Asmundsson, Rardin et al. 2009), we shall focus in this paper on empirical estimation from simulation data, which has been the most prevalent approach in the literature. In this approach, the production system under study is simulated and data collected from the simulation model in each planning period on the quantities of interest, which depend on the independent variables postulated for the clearing function. Once the data have been obtained, they are fit empirically using some form of regression analysis to a predefined functional form.

The estimation of clearing functions in a multiproduct environment poses a number of challenges due to the fact that since the potential output of the resource in a planning period is determined by the WIP level of all products. Asmundsson et al.(2009) show that this can lead to anomalous behavior where WIP of one product is held in order to process another product rapidly. These authors suggest an allocated clearing function formulation to resolve this difficulty, under which an aggregate clearing function for the resource is constructed in units of time and the resulting output capacity is allocated to the different products. In this paper, we take a different approach where we estimate a clearing function for each individual product based on the release quantities and WIP levels of this product and other products in the current and immediately preceding periods. The values of all possible variables of interest are regressed against the output of each product. In this way we hope to obtain insight into which variables are important to include in the clearing functions and which can safely be omitted.

In the next section we give a brief description of the simulation model of the production facility used as a testbed in this research.

3 SIMULATION MODEL

Our simulation model of a re-entrant bottleneck system was built with attributes of a real-world semiconductor wafer fabrication facility (fab) environment previously studied by our research group (Kayton, Teyner et al. 1997). The major characteristics of wafer fabrication, including re-entrant product routings, unreliable machines, batch processing machines and multiple products with varying process routings and number of operations are included in the model. There is a distinct re-entrant bottleneck station representing the photolithography process. The processing times for all other stations were scaled so that no non-bottleneck station would have utilization approaching that of the bottleneck. The model has batching stations (Stations 1 and 2) early in the process, representing furnaces that perform diffusion and oxidation processes. These machines can process between two and four lots, which may be of different product types, simultaneously as a batch. The remaining stations process one lot at a time.
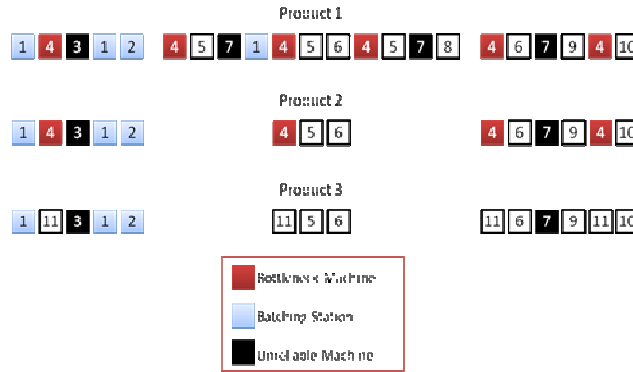
Figure 2: Re-entrant Bottleneck Model Process Chart for Products

There are 11 machines (stations) and 3 products in this model. The number of operations for product 1, product 2 and product 3 are 22, 14 and 14 respectively. Machine 4 is the bottleneck station shown in red in

Figure **2**, and has two servers. Products 1 and 2 visit the bottleneck station 6 and 4 times respectively. Product 3 does not use the bottleneck machine. This product visits Station 11 which is the only station that is allowed to exceed 80% utilization, but it is not allowed to exceed the bottleneck utilization. Machines 3 and 7, shown in black in Figure 2, are unreliable machines whose failures cause starvation at the bottleneck. The system is required to produce a product mix that is 3:1:1 of Products 1, 2, and 3 respectively. Details are given in Irdem, Kacar et al.(2008), and are available from the authors upon request.

All processing times follow a lognormal distribution. The processing times are in minutes. The standard deviations of process times are less than or equal to 10 percent of the mean. The processing times for all individual products in a machine are same. Also, the processing time of a machine for different operations of products is same. The mean time to failure (MTTF) and mean time to repair (MTTR) distributions follow gamma distributions. Machines 3 and 7 are the unreliable machines can produce a product in a very short time but can starve the bottleneck due to poor availability.

In the next section we will briefly describe the production planning model based on clearing functions that is used to examine the performance of the different methods of estimating clearing functions.

4   OPTIMIZATION MODEL FOR PRODUCTION PLANNING

For all the planning models presented in this paper, we shall define an operation l to be the processing of a product at a specific machine k. We define $L(k)$ to be the set of all operations l taking place on a specific machine k. We also note that all operations for all products that take place on a specific machine require same processing time.

We define the following notation:

Indices:

g : product index.

k : machine index.

l : operation index.

Decision Variables:

Y_{gpl} : output quantity of product g from operation l in period p.

$Y_{gp}(k)$: output quantity of product g from machine k in period p; $Y_{gp}(k) = \sum_{l \in L(k)} Y_{glp}$

X_{gpl} : release quantity of product g in period p at operation l.

$X_{gp}(k)$: release quantity of product g in period p at machine k; $X_{gp}(k) = \sum_{l \in L(k)} X_{gpl}$

Y_{gp} : output quantity of product type g in period p.

X_{gp} : quantity of product g released into the first station in the line in period p.

W_{gpl} : WIP of product g at operation l at the end of period p.

$W_{gp}(k)$: WIP of product g during period p at machine k; $W_{gp}(k) = \sum_{l \in L(k)} W_{gpl}$.

Z_{gp}^k : Fraction of capacity at machine k allocated to product type g in period p.

I_{gp} : units of product g in finished goods inventory at the end of period p.

B_{gp} : units of product g backlogged at the end of period p.

Parameters:

h_{gp} : Unit inventory holding cost for product g in period p.

b_{gp} : Unit backlogging cost for product g in period p.

ω_{gp} : Unit WIP cost for product g in period p.

D_{gp} : demand for product g during period p.

$M(k,g)$: set of indices denoting the variables included in the Regression Model at machine k.

μ_{kg} : Intercept of the clearing function for product g at machine k.

β_{kg}^n : Coefficient of variable n in the clearing function for product g at machine k.

C_k : Maximum capacity of machine k in units of products. Since in our simulation model all operations of all products at the same machine require the same amount of time, this is a valid representation.

The formulation can now be given as follows:

Objective function:

$$min \sum_{g \in G} \sum_{p=1}^{P} \sum_{k=1}^{K} \omega_{gpk} W_{gp}(k) + + \sum_{g \in G} \sum_{p=1}^{P} h_{gp} I_{gp} + \sum_{g \in G} \sum_{p=1}^{P} b_{gp} B_{gp} \qquad (1)$$

Subject to:

WIP Balance:

$$W_{gpl} = W_{g,p-1,l} + X_{gpl} - Y_{gpl} \quad g \in G, p = 1, ..., P, l \in L \qquad (2)$$

Inventory Balance:

$$Y_{gp} + I_{g,p-1} - B_{g,p-1} - I_{gp} + B_{gp} = d_{gp} \; g \in G, p = 1, ..., P \qquad (3)$$

Capacity Constraints:

$$Y_{gp}(k) \leq \mu_{kg} + \sum_{n \in M(kg)} \beta_{kg}^n MRV_g^n, \quad g \in G, \; p = 1, ..., P, k \in K, \qquad (4)$$

$$\sum_{g \in G} Y_{gp}(k) \leq C_k, p = 1, ..., P, k \in K \qquad (5)$$

Variable Nonnegativity:

$$W_{gpl}, Y_{gpl}, I_{gp}, B_{gp}, X_{gpl} \geq 0 \; g \in G, p = 1, ..., P, k \in K \qquad (6)$$

where MRV_g^n denotes the appropriate variables identified in the regression models discussed in the next section. In this formulation, we aim to minimize the sum of finished goods inventory holding, WIP holding and backordering costs across the planning horizon. The first constraint sets (2) and (3) represent WIP and finished inventory balance for each product at each operation. The capacity constraints (4) represent the clearing functions estimated for each product using linear regression as described in the following section. Finally, the total output of all products at each machine is constrained not to exceed the overall capacity of the machine in the planning period given in (5).

5   DATA COLLECTION AND MULTIPLE LINEAR REGRESSION MODELS

We simulate the performance of the system over a year with release schedules corresponding to bottleneck utilization of approximately 49%, 60%, 70%, 77%, 87%, 94% and 99% with fifteen replications, collecting the release, WIP and output values observed for each product in each period for individual products. Based on exploratory experiments, we focused on the three linear regression models below. For brevity we shall state the models for Product 1 only; those for other products are specified in an analogous manner.

Model 1: Under this model the output of each product at machine k in period p is determined by the releases of all products in that period and the WIP level of each product at the start of period p.

$$Y_1(k,p) = \mu_{k,1} + \beta_{k,1}^1 X_{1,p}(k) + \beta_{k,1}^2 W_{1,p-1}(k) + \beta_{k,1}^3 X_{2,p}(k) + \beta_{k,1}^4 W_{2,p-1}(k) + \beta_{k,1}^5 X_{3,p}(k) + \beta_{k,1}^6 W_{3,p-1}(k)$$

Model 2: This model considers the same set of variables as Model 1, with the addition of the releases of Product 1 in the immediately preceding period.

$$Y_1(k,p) = \mu_{k,1} + \beta_{k,1}^1 X_{1,p}(k) + \beta_{k,1}^2 W_{1,p-1}(k) + \beta_{k,1}^3 X_{2,p}(k) + \beta_{k,1}^4 W_{2,p-1}(k) + \beta_{k,1}^5 X_{3,p}(k) + \beta_{k,1}^6 W_{3,p-1}(k) + \beta_{k,1}^7 X_{1,p-1}(k)$$

Model 3: This model considers the variables in Model 2 with the addition of the releases of all products in the immediately preceding period.

$$Y_1(k,p) = \mu_{k,1} + \beta_{k,1}^1 X_{1,p}(k) + \beta_{k,1}^2 W_{1,p-1}(k) + \beta_{k,1}^3 X_{2,p}(k) + \beta_{k,1}^4 W_{2,p-1}(k) + \beta_{k,1}^5 X_{g=3,p}(k) + \beta_{k,1}^6 W_{3,p-1}(k) + \beta_{k,1}^7 X_{1,p-1}(k) + \beta_{k,1}^8 X_{2,p-1}(k) + \beta_{k,1}^9 X_{3,p-1}(k)$$

The specific forms of these models are likely to be somewhat specific to our experimental environment. For example, we very seldom had cycle times of more than two periods at any station; if longer cycle times were incurred, additional release variables would be likely to enter the regression equations. For each of these models, we applied three different regression criteria:

1)   Adjusted R square Selection (ARSQ):
   The Adjusted *R*-Square method finds the subset of independent variables that best predict a dependent variable by linear regression in the given sample, selecting the subset of dependent variables yielding the highest adjusted R square value.

2)   Forward Selection (FORWARD):
   For each independent variable, this method calculates *F*-statistics that reflect the variable's contribution to the model if it is included. The *p*-values for these *F* statistics are compared to a threshold significance level. If no *F*-statistic has a significance level greater than threshold level, the procedure stops. Otherwise, the method adds the variable with the largest *F*-statistic to the model. The method then recalculates *F*-statistics for the variables not yet included, and the evaluation process is repeated. Thus variables are added to the model one by one until no as yet unincluded variable produces a significant *F*-statistic.

3)   Backward Elimination (BACKWARD)
   This selection technique begins by calculating *F*-statistics for a model, including all independent variables. Variables are then deleted from the model one by one until all the variables remaining in the model produce *F*-statistics significant at the threshold significance level At each step, the variable showing the smallest contribution to the model is deleted.

6   EXPERIMENTAL DESIGN

Our experiments are designed to examine the effects of two different factors on the performance of the different production planning models: the bottleneck utilization at different demand patterns, and the severity of the machine failures.

1) Bottleneck Utilization with Different Demand Patterns:

Queuing theory suggests that the nonlinear relationship between resource utilization and flow times becomes more severe at high utilization levels. Hence we experiment with two bottleneck utilization values of 0.7 and 0.9. The utilization level is achieved by varying the demand of all products while maintaining the 3:1:1 product mix. We also consider two different demand patterns. Under the first demand stays constant throughout the planning horizon, representing what ought to be favorable conditions any planning model. Under the second case the demand changes from period to period. In this latter case the first and last three periods demands are set constant in order to minimize beginning and ending effects. Again we have two varying demand patterns that will give an average bottleneck utilization of 0.7 and 0.9 over the planning horizon as shown in Figure 10 below for the case of 90% utilization.
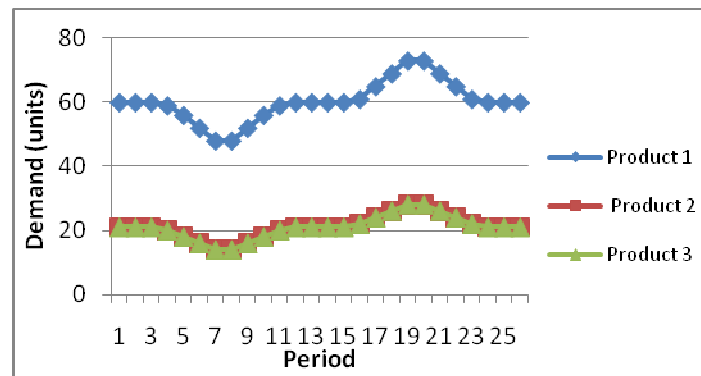

Figure 3: Varying Demand Pattern corresponding 90% Utilization

2) Severity of Machine Breakdowns at Unreliable Machines:

We would expect the performance of planning models to deteriorate under higher machine failures, since these increase the variability in the system. For the Short Failure case, we use MTTF = 7200 minutes and MTTR = 1800 minutes. Our period length is 10080 minutes, meaning that on average the machines will fail once in every period. In the Long Failure case, we double both the MTTF and MTTR values, thus keeping the average availability constant while increasing its variability. Under the Long Failure case, we estimate new clearing functions, as the nature of the underlying production system is substantially altered by the change in machine failure distributions.

We compare the different methods for estimating clearing functions as follows:

Step 1: Solve the LP using the clearing functions estimated by the different regression models to obtain a planned release schedule.

Step 2: Simulate the operation of the fab under this release plan for fifteen replications recording the realized output.

Step 3: Compare the output planned in the LP model for each planning period to the average output in that period over the fifteen replications.

We will also compare the objective function values realized by the simulated execution of the plans suggested by the different models. To accomplish this we take the releases from LP models and give them as input to the simulation model to obtain the realized average inventory, backlog and output levels. This

allows us to calculate the realized profit by multiplying these values by their corresponding costs based on the average of 15 replications. We use values of 50 for backlog cost, 35 for WIP cost, 3 for material cost, 15 for inventory cost and 60 for revenue. We assume that WIP cost is high due to restrictions of space in the factory layout and use the same cost values for all products.

For brevity we use a triple to denote each scenario. In the first position, S stands for short failures and L for long failures. In the second position, the numbers 70 and 90 correspond to 70% and 90% bottleneck utilization, respectively. In the final position the letters C and V stand for constant and varying demand, respectively. V1A stands for Model 1 with Adjusted *R*-square selection. F and B denote Forward and Backward selection, respectively, while V2 and V3 denote Models 2 and 3 respectively.

7   EXPERIMENTAL RESULTS

Our first comparison is based on the realized profit values shown in Figure 4. To further probe the relationship between the performance of the different models, Tukey's Studentized Range (HSD) Test is used to statistically compare the mean of profit values of all algorithms. This test compares the mean of every treatment to the mean of every other treatment; that is, it applies simultaneously to the set of all pairwise comparisons and identifies where the difference between two means is greater than the standard error would be expected to allow. The confidence coefficient for the set, when all sample sizes are equal, is exactly $1 - \alpha$. In our analysis, we use $\alpha = 0.05$.
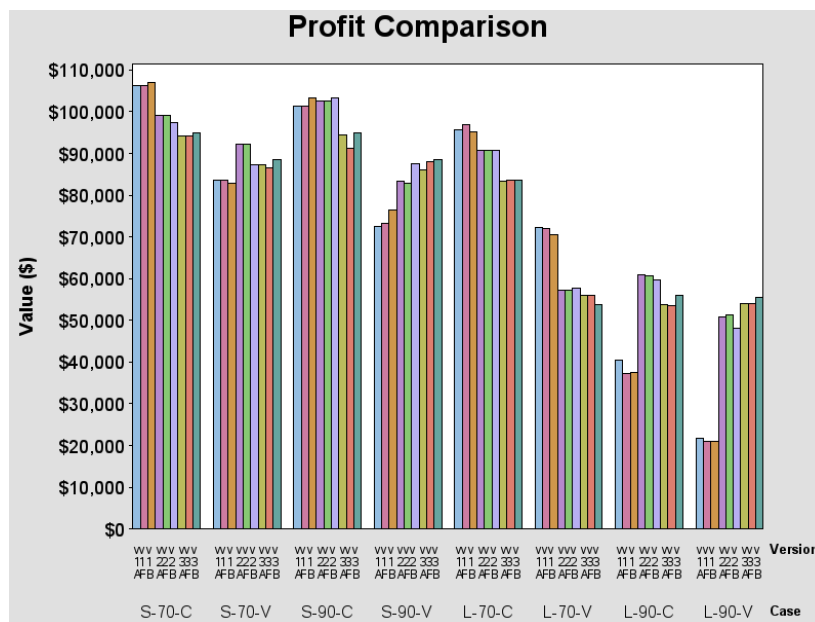

Figure 4: Profit Comparison of Models

The results from this test are shown in Table 1. Identical letters in the Solution (S) column represent mean values that are not statistically different from each other at 95% confidence level. These results suggest that there is no single dominant model or selection method that consistently has higher profit values than the others. However, in general, the same models with different selection procedures tend not to be statistically different at a 95% confidence level. This is clearly observed for the S-70-C case where modes are sorted by the version numbers and only one of the selection procedures is statistically different. For all 70% utilization cases for any demand or length of failure, Version 1 performs best with any selection procedure except in the S-70-V case where Model 2 takes the lead. This suggests that, under low utilization, the models including only variables corresponding to that period (V1) perform better than models including variables addition related to the immediately preceding period (V3). This is intuitive; at low

utilization the production system is likely to be able to process all work released to it in the same period, without carrying large amounts of WIP into future periods and maintaining low cycle times.

At high utilization, there is no consistent selection procedure that is dominating for low utilization, such as in L-70-V, top three models in profit are version 1s, Adjusted and forward selection yielding no statistical difference, whereas in L-70-C, the top three models are all Model 1 but all selection procedures are statistically different.

Table 1: Tukey Multi Comparison Analysis

S-70- C			S-70-V			S-90-C			S-90-V		
S	Mean	Model	S	Mean	Model	S	Mean	Model	S	Mean	Model
A	106954	V1B	A	92133	V2A	A	103225	V1B	A	88653	V3B
B	106258	V1F	A	92132	V2F	A	103203	V2B	A	88132	V3F
B	106257	V1A	B	88514	V3B	B	102563	V2F	B	87621	V2B
C	99030	V2A	C	87360	V2B	B	102563	V2A	B	86022	V3A
C	99029	V2F	C	87222	V3A	B	101411	V1A	C	83300	V2A
D	97361	V2B	C	86650	V3F	B	101347	V1F	C	82933	V2F
E	94939	V3B	D	83609	V1F	C	94977	V3B	D	76536	V1B
F	94238	V3A	D	83608	V1A	C	94336	V3A	E	73283	V1F
F	94232	V3F	D	82818	V1B	D	91327	V3F	E	72600	V1A
L-70- C			L-70-V			L-90-C			L-90-V		
S	Mean	Model	S	Mean	Model	S	Mean	Model	S	Mean	Model
A	96895	V1F	A	72260	V1A	A	60816	V2A	A	55452	V3B
B	95702	V1A	A	72061	V1F	A	60796	V2F	A	54082	V3F
C	95179	V1B	B	70636	V1B	A	59756	V2B	A	54082	V3A
D	90746	V2A	C	57790	V2B	B	56116	V3B	B	51221	V2F
D	90743	V2F	C	57330	V2A	C	53824	V3A	B	50926	V2A
D	90668	V2B	C	57326	V2F	C	53426	V3F	C	48155	V2B
E	83532	V3F	D	56025	V3A	D	40360	V1A	D	21621	V1A
E	83480	V3B	D	55876	V3F	E	37600	V1B	D	21009	V1F
E	83412	V3A	E	53898	V3B	E	37235	V1F	D	20991	V1B

In the 90% utilization cases, we start to see the emergence of Model 3 as the leading regression model especially in the varying demand cases. These models include the releases of all products in the immediate prior period in the regression model. This can be expected since for high utilization, outputs of products are more like to be affected by the releases of preceding periods; cycle time is long, and all the product released in the prior period may not have exited the resource. Hence output of a given product in a period is significantly affected both by the releases of that product in the immediate prior period (Model 2) and also feel the effects of other products (Model 3).

For the constant demand cases at 90% utilization, Model 3 is outperformed by Model 2 for both long and short failure cases. In the L-90-C case Model 2s are the top three with no statistical difference between selection procedures. In the S-90-C case, Model 2 with backward selection (V2B) shares the lead with Model 1 and backward selection (V1B) with no statistical difference between them. This suggests that under constant demand, releases and WIP of the products are quite similar from period to period, and releases and WIP to be somewhat correlated over time, limiting the difference between Models 2 and 3.

It is interesting to examine the cost breakdowns shown in Figure 5 to see where the differences in profit between models arise. In the vast majority of cases, as would be expected, the total revenue generated by all models is almost equal; the differences arise from the different ways in which the models choose to incur costs to satisfy demand. Material costs are almost constant across all models, since the all generate approximately the same total output. In the long failure cases, the performance of all models deteriorates due to the higher level of variability created by the less frequent but longer machine failures. The cost structure of each model under different selection procedures are quite similar, consistent with the findings from the Tukey analysis. In most cases, WIP cost is almost as much as all other costs combined, and is higher for some. This reflects that the release schedule given by the LP models tend to hold excess WIP in the system, and the difference between regression models is determined by which one holds less WIP than the others. The backlog and inventory cost trade off against each other; when there is more in-

ventory, it means less backlog cost and vice versa. Still the leading cost in the models is WIP cost except in the S-70-C case which is the most favorable case for the models.
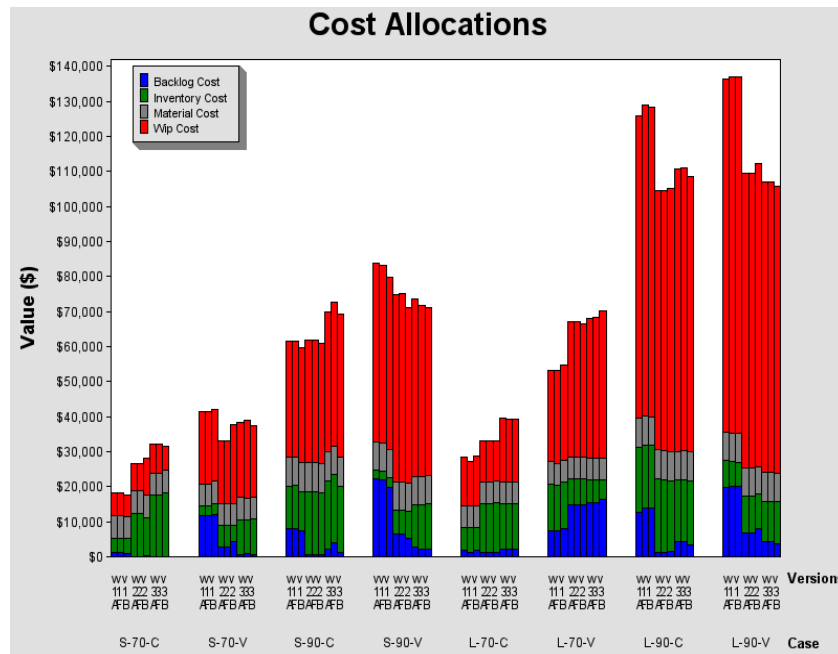

Figure 5: Cost Allocations of Models

In order to further explore the differences between the models we calculated the R^2 values for all models on all machines. While these results are not presented in detail for brevity, it was notable that in almost all cases very high R^2 values were obtained, despite the quite substantial differences in profit yielded by the different models. Different selection procedures result in different numbers and choices of variables but we show that in the profit comparison, for most of the cases, for same version different selection methods are not statistically different from each other.

8 SUMMARY AND FUTURE DIRECTIONS

Taken together, these experimental results suggest that the task of estimating clearing functions from simulation data is far from being a straightforward exercise in linear regression. As suggested by Missbauer(2009), the output of a production resource is highly state dependent, with the variables impacting the clearing function changing based on system state. The clearing functions employed by most researchers to date represent an average relation over a wide range of operating states, but may be quite inaccurate for a given sample path of system evolution. It is also interesting that all models give very high R^2 values despite the significant differences in costs they induce when used in a production planning model. Thus the usual goodness of fit criteria are not sufficient to determine the quality of fit of a clearing function.

The most important future direction emerging from this work is the need to construct a mathematical model from first principles upon which procedures for the estimation of clearing functions, or of other, more suitable constructs, can be placed on a firm foundation. A number of efforts (Selçuk , Fransoo et al. 2007; Missbauer 2009) in this direction have tried to determine clearing functions based on transient system states. The queueing models of Riano(2003) are also an important step in this regard, and suggest a direction forward.

ACKNOWLEDGMENTS

REFERENCES

Asmundsson, J. M., R. L. Rardin, R. Uzsoy, C.H. Turkseven. 2009. "Production Planning Models with Resources Subject to Congestion." *Naval Research Logistics* 56: 142-157.

Asmundsson, J. M., R. L. Rardin, R. Uzsoy. 2006. "Tractable Nonlinear Production Planning Models for Semiconductor Wafer Fabrication Facilities." *IEEE Transactions on Semiconductor Manufacturing* 19: 95-111.

Byrne, M. D. and M. A. Bakir 1999. "Production Planning Using a Hybrid Simulation-Analytical Approach." *International Journal of Production Economics* 59: 305-311.

Byrne, M. D. and M. M. Hossain 2005. "Production Planning: An Improved Hybrid Approach." *International Journal of Production* Economics 93-94: 225-229.

Dauzère-Péres, S. and J.-B. Lasserre 1994. *An Integrated Approach In Production Planning And Scheduling.* Berlin ; New York, Springer-Verlag.

Graves, S. C. 1986. "A Tactical Planning Model for a Job Shop." *Operations Research* 34: 552-533.

Holt, C. C., F. Modigliani, J. F. Muth. 1956. "Derivation of a Linear Rule for Production and Employment." *Management Science* 2(2): 159-177.

Holt, C. C., F. Modigliani, J. F. Muth, H. A. Simon. 1960. *Planning Production, Inventories and Work Force*. Englewood Cliffs, NJ, Prentice Hall.

Hopp, W. J. and M. L. Spearman 2001. *Factory Physics : Foundations of Manufacturing Management*. Boston, Irwin/McGraw-Hill.

Hung, Y. F. and G. J. Cheng 2002. "Hybrid Capacity Modelling for Alternative Machine Types in Linear Programming Production Planning." *IIE Transactions* 34: 157-165.

Hung, Y. F. and R. C. Leachman 1996. "A Production Planning Methodology for Semiconductor Manufacturing Based on Iterative Simulation and Linear Programming Calculations." *IEEE Transactions on Semiconductor Manufacturing* 9(2): 257-269.

Irdem, D. F., N. B. Kacar, R. Uzsoy. 2008. An Experimental Study of an Iterative Simulation-Optimization Algorithm for Production Planning. *2008 Winter Simulation Conference*. S. J. Mason, R. Hill, L. Moench and O. Rose. Miami, FL.

Irdem, D. F., N. B. Kacar, R. Uzsoy. 2010. "An Exploratory Analysis of Two Iterative Linear Programming-Simulation Approaches for Production Planning." *IEEE Transactions on Semiconductor Manufacturing*.23:442-455.

Johnson, L. A. and D. C. Montgomery 1974. *Operations Research in Production Planning, Scheduling and Inventory Control*. New York, John Wiley.

Karmarkar, U. S. 1989. "Capacity Loading and Release Planning with Work-in-Progress (WIP) and Lead-times." *Journal of Manufacturing and Operations Management* 2(105-123).

Kayton, D., T. Teyner, C Shwartz, R. Uzsoy. 1997. "Focusing Maintenance Improvement Efforts in a Wafer Fabrication Facility Operating Under Theory of Constraints." *Production and Inventory Management*(Fourth Quarter): 51-57.

Kim, B. and S. Kim 2001. "Extended Model for a Hybrid Production Planning Approach." *International Journal of Production Economics* 73: 165-173.

Missbauer, H. 2002. "Aggregate Order Release Planning for Time-Varying Demand." *International Journal of Production Research* 40: 688-718.

Missbauer, H. 2009. "Models of the transient behaviour of production units to optimize the aggregate material flow." *International Journal of Production Economics* 118:387-397.

Missbauer, H. and R. Uzsoy (forthcoming). Optimization Models for Production Planning. *Planning Production and Inventories in the Extended Enterprise: A State of the Art Handbook*. K. G. Kempf, P. Keskinocak and R. Uzsoy. Norwell, MA, Springer.

Pahl, J., S. Voss, D. L. Woodruff.. 2005. "Production Planning with Load Dependent Lead Times." *4OR: A Quarterly Journal of Operations* Research 3: 257-302.

Pahl, J., S. Voss, D. L. Woodruff. 2007. "Production Planning with Load Dependent Lead Times: An Update of Research." *Annals of Operations Research* 153: 297-345.

Riaño, G. (2003). *Transient Behavior of Stochastic Networks: Application to Production Planning with Load-Dependent Lead Times*. School of Industrial and Systems Engineering. Atlanta, GA, Georgia Institute of Technology.

Selçuk , B., J. C. Fransoo, A. G. De Gok. 2007. "Work in Process Clearing in Supply Chain Operations Planning." *IIE Transactions* 40: 206-220.

Srinivasan, A., M. Carey,T. E. Morton. 1988. *Resource Pricing and Aggregate Scheduling in Manufacturing Systems.* Graduate School of Industrial Administration, Carnegie-Mellon University. Pittsburgh, PA.

Vollmann, T. E., W. L. Berry, Whybark D.C, F. R. Jacobs. 2005. *Manufacturing Planning and Control for Supply Chain Management. * . New York, McGraw-Hill.

Voss, S. and D. L. Woodruff. 2003. *Introduction to Computational Optimization Models for Production Planning in a Supply Chain.* Berlin ; New York, Springer.

AUTHOR BIOGRAPHIES

NECIP BARIS KACAR is currently working as Graduate Industrial Trainee in SAS Institute. He is also Research Assistant and holds an M.S degree with Minor in Operations Research in Industrial Engineering in the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University .He is currently pursuing Ph.D. degree at the same university. He received a BS degree in Mechanical Engineering from Bogazici University, Istanbul, Turkey. His research interests are in simulation based optimization, production planning and logistics. He can be reached via email at <nbkacar@ncsu.edu>.

REHA UZSOY is Clifton A. Anderson Distinguished Professor in the Edward P. Fitts Department of Industrial and Systems Engineering at North Carolina State University. He holds BS degrees in Industrial Engineering and Mathematics and an MS in Industrial Engineering from Bogazici University, Istanbul, Turkey. He received his Ph.D in Industrial and Systems Engineering in 1990 from the University of Florida. His teaching and research interests are in production planning, scheduling, and supply chain management. His research has been supported by the National Science Foundation, Intel Corporation, Hitachi Semiconductor, Harris Corporation, Kimberly Clark, Union Pacific, Ascension Health and General Motors. He was named a Fellow of the Institute of Industrial Engineers in 2005 and Outstanding Young Industrial Engineer in Education in 1997, and has received awards for both undergraduate and graduate teaching. His email address is <ruzsoy@ncsu.edu>.