



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 6N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-56316-8

Canada

Dispatch Rule and Batch Size Effects
on the Performance of Flow Shops
Characterized by Multiple Revisit Workstations

B. Alexander T. Edwards

Thesis submitted to
the School of Graduate Studies and Research
in partial fulfillment of the requirements for the
degree of
Master of Applied Science
in
System's Science



B. Alexander T. Edwards, Ottawa, Canada, 1989



UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

ABSTRACT

This thesis examines a set of service strategies for flow shops characterized by high utilization levels and multiple revisit workstations. The service strategies encompass a set of machine batch sizes applicable to machines capable of processing more than one job simultaneously, and a set of dispatch rules for workstation queues. The machine resources at these high utilization and multiple revisit workstations are referred to as 'hubs' and their operation significantly affects work flow in the entire system.

The integrated circuit (IC) industry provides good examples of manufacturing processes characterized by 'hubs'. A data model provided by Northern Telecom of an IC fabrication process serves as a basis for the flow shop models in this thesis. Discrete simulation models at three levels of complexity are examined. They are: (1) models of the major 'hubs' considered in isolation, (2) a network model of the major 'hubs', and (3) a full process model with explicit representation of all workstations in the IC fabrication process. Trends and effects observed on the less complex models are extended and generalized to the more complex models.

The results show that there are performance differences between the service strategies and that selection of a particular strategy for an individual system must satisfy sometimes competing factory performance objectives. As well, job arrival distributions and machine unreliability are shown to have a significant effect on the performance of the service strategies.

ACKNOWLEDGEMENTS

Special thanks to my family.

I would like to thank Professors Dan Lane and Jeff Sidney at the University of Ottawa and Bob Poulsen at Northern Telecom for all their time, effort, counsel, and keen interest. These were greatly appreciated.

To all my friends, whose support and friendship are so important.

CONTENTS

<u>Section</u>	<u>Page</u>
1. INTRODUCTION.....	1
1.1 Integrated Circuit Manufacturing.....	2
1.2 Research Approach.....	4
2. LITERATURE REVIEW.....	7
2.1 Job Dispatching in Flow and Job Shops.....	7
2.2 Modelling IC Manufacturing.....	11
2.3 Sampling Methods in Discrete Simulation.....	15
3. GENERAL MODEL.....	20
3.1 Modelling Approach.....	20
3.2 Sampling Considerations.....	31
4. ISOLATED HUB MODELS.....	34
4.1 Furnace Batch Level Experiments.....	34
4.2 Stepper Workstation Experiments.....	43
4.3 Dispatch Rule Comparison.....	45
4.4 Summary.....	47
5. HUB NETWORK.....	49
5.1 Hub Network Experiments.....	49
5.2 Dispatch Rule Combinations.....	55
5.3 Summary.....	58
6. FULL PROCESS MODEL.....	60
6.1 Standard Factorial Experiments.....	60
6.2 Job Arrival Distribution.....	67
6.3 Machine Failure Analysis.....	75
6.4 Summary.....	78
7. SUMMARY OF RESULTS.....	79
7.1 Furnace Batch Sizes.....	79
7.2 Comparison of Dispatch Rules.....	81
7.3 System Bottleneck Utilization Level.....	83
7.4 Extensions to the Full Process Model.....	83
8. CONCLUSIONS & EXTENSIONS.....	86
8.1 Conclusions.....	86
8.2 Extensions.....	88

APPENDIX A - SIMULATION DATA MODELS.....	9 1
APPENDIX B - EXPERIMENTAL RESULTS.....	9 9
APPENDIX C - QUEUEING EXAMPLE.....	1 1 6
BIBLIOGRAPHY.....	1 1 9

List of Tables

<u>Table</u>		<u>Page</u>
4.1	Constant Job Arrival Rates	36
4.2	Job Arrival Rates	39
4.3	Constant Workstation Utilization Levels	40
4.4	Queue Time versus Throughput.....	42
4.5	Stepper Workstation Results.....	44
4.6	Dispatch Rule Comparison.....	45
5.1	Hub Network Results	51
5.2	Queue Time versus Throughput.....	53
5.3	Dispatch Rule Comparison.....	54
5.4	Utilization Level Effect.....	55
6.1	Full Process Model Results	62
6.2	Queue Time versus Throughput.....	64
6.3	Dispatch Rule Comparison.....	65
6.4	Utilization Level Effect.....	66
6.5	Stochastic Job Arrival Results.....	68
6.6	Dispatch Rule Comparison (Stochastic Job Arrival).....	69
6.7	Gated Release Policy Results.....	73
6.8	Dispatch Rule Comparison (Gated Release Policy).....	74
6.9	Machine Failure Results.....	76
6.10	Dispatch Rule Comparison (Machine Failure).....	77

List of Figures

<u>Figure</u>		<u>Page</u>
3.1	Simulation System.....	20
3.2	Process Time Breakdown.....	23
3.3	Furnace B Model.....	24
3.4	Hub Network Model.....	25
3.5	Furnace Workstation.....	29
3.6	Inventory Trace.....	33
3.7	Cycle-Time Trace.....	33
4.1	Factorial Furnace Experiment (Constant Arrival Rates).....	35
4.2	Factorial Furnace Experiment (Constant Utilization).....	39
4.3	Queue Time Range.....	47
5.1	Factorial Hub Network Experiment.....	50
5.2	Dispatch Rule Combination Experiment.....	56
5.3	Dispatch Rule Combinations.....	57
6.1	Factorial Experiment.....	61
6.2	Inventory Trace.....	70
6.3	Cycle-Time Trace.....	71

1. INTRODUCTION

Prioritization techniques for ordering job queues at workstations in flow shops have received much attention in scheduling literature. An important and relatively unexplored class of problems occurs in flow shops with looping. Looping occurs in a flow shop when the process sequence requires the same workstation at two or more noncontiguous steps. The machine resources at these workstations are often significant bottlenecks, affecting flow in the entire system. Workstations satisfying this criterion are referred to as 'hubs' and frequently serve several process sequences and job types in the system. Analysis and study of queueing techniques in such flow shops will lead to more efficient utilization of these critical system components.

The integrated circuit (IC) industry provides good examples of manufacturing processes that are characterized by looping. IC manufacturing involves complex process flows in a demanding production environment with looping involving multiple processes and product types. With many machines costing well over a million dollars, it is vital that scheduling methodologies be well understood to ensure efficiency with respect to production goals.

This research is aimed at furthering our understanding of the effects of scheduling and batching rules on the performance of manufacturing systems characterized by looping. In particular the thesis focuses on the hub elements in the system. The thesis provides a methodology for the evaluation of scheduling and batching rules for hubs with selected stresses being applied to the system to enhance the completeness and robustness of the results.

The methodology is applied to a complex IC manufacturing system to evaluate its performance with respect to scheduling operations. Models will be developed to reflect three increasingly complex views of the system.

These views are:

1. Single hub system with multiple revisits
2. Multiple hub network
3. Complex queueing network including hub and off-hub machines

The Northern Telecom Electronics Advanced Technology Centre in Nepean, Ontario, is the basis for this thesis. This IC facility is used for both research and development and production purposes; and provides a 'real life' environment for the models developed in this thesis.

1.1 Integrated Circuit Manufacturing

Integrated circuits are manufactured on wafers of polished silicon of perhaps one half millimetre thickness and nine to fifteen centimetres in diameter. Each silicon wafer may contain up to a few hundred identical individual circuits. These wafers are generally processed in lots of twenty to twenty-five wafers during the wafer fabrication stage. After completion of work in the 'wafer-fab' area the wafers enter various stages of testing and die cutting. Here, they are cut into individual circuits, called dies, several millimetres square. Next the dies enter the bonding and packaging stages during which the I/O wires are attached followed by mounting into protective plastic or ceramic packages.

A wafer fabrication process may consist of several hundred individual steps. These steps are performed on several dozen to one hundred machine types. Many machines are each used for several steps in the process. A typical process involves a number of cyclical loops in which the basic formations of the individual circuit elements, typically transistors and capacitors, are manufactured on the surface of the wafer. These basic structures typically include the elemental transistor building blocks: the gate, source, and drain, and the various levels of insulating material and interconnections between

circuit elements. The three-dimensional topology of these circuits requires that the cyclical nature of the process slowly build up the circuit elements. The nature of the process is such that these cycles involve very similar steps so that the same machines and equipment types can be used at each cycle. Some set-up is usually involved to configure a machine appropriately for the current cycle. In addition to these cycles numerous inspections and tests are performed to monitor the circuit's compliance with process specifications. Additional steps may be added to the basic process to allow for circuit options.

The extreme susceptibility of the final product's quality to variations in the manufacturing process requires that much time, expense, and research be put into minimizing this variability. Sources of this variability range from the standard concern of too much 'work-in-process' standing idle to controlling sub-micron dust particles and thermal stresses in processing equipment. If idle time can be reduced then occurrences of wafer quality and yield problems are generally reduced. Reduced product cycle-times yield large benefits in the area of quality, yield, predictability, and bottom line cost.

The machines involved in this process may operate on individual die or several lots of wafers at a time. Where it is possible to operate on more than one lot of wafers at a time the lots require the same set of processing conditions in that machine. These conditions are referred to as a 'recipe'. During the manufacturing cycle, when one of the multi-lot machines becomes available compatible, i.e. equal recipe, lots must be loaded into the machine and processed. The situation often arises when there are too few lots available to process as a complete batch. The decision at this point is whether to process the few available lots or wait until a complete batch of lots is available. Determining the right batch size is crucial with respect to the factory capacity and product cycle-time. In the manufacturing model studied in this research the multi-lot machines are furnaces with a capacity of four lots.

At any machine in the manufacturing sequence, circumstances will always arise when more than one lot is waiting to be processed. The decision on which lot should be processed first is referred to as the dispatching decision. Distinction is made between classes of dispatch rules based on the type and location of information used in the decision criteria. When the information required can be found in the local queue the rule is referred to as a 'local dispatch rule'. A 'global rule' requires information found in the system wide set of machine queues and process flows. As well, some planning and scheduling systems take into account external management concerns and priorities. Many approaches to studying a wide number of rules have been explored in research dating back thirty years. This history is examined in the literature review following this section, and includes recent work in the area of integrated circuit manufacturing.

1.2 Research Approach

A series of discrete simulation models built on data obtained from Northern Telecom are used as a vehicle for this research. Three critical hub workstations have been selected from the manufacturing process for extensive study. The machines in these workstations all perform fundamental processing steps in the fabrication of wafers. Two of these hub workstations consist of furnaces each capable of processing up to four lots at a time, while the third hub workstation contains stepper machines. Furnaces are used for several tasks including growth of insulating oxides on the wafer surface and annealing of the wafers to alter their electronic profiles. Steppers are used to imprint the horizontal circuit layout on the wafers. For a detailed look at the physics and fabrication processes of integrated circuits see Sze (1983).

Many methods have been used to characterize and measure factory performance. Product cycle-times, factory capacities, yield, quality measures, and product costs have all been used for this purpose. Dispatch and batching methods are used to reduce the

amount of time that a product spends waiting in workstation queues. Cumulative waiting time in queues has direct and indirect effects on all of the performance measures listed above and is the primary method of evaluation used in this thesis.

The simulation models studied in this thesis range in complexity from individual work stations to a large queueing network of a complete wafer fabrication area. The focus in each model is the looping structure of the associated hubs. The aim is to understand and demonstrate the relative characteristics of dispatch and batching methods on simple models and then to extend these to more realistic and complex models.

One model is developed for each of the three critical hubs selected for study. In each of these, the hub is modelled as an isolated workstation, with the work between successive visits to the hub being modelled as a single time delay representing the aggregation of several tasks. The effect of repeat visits to the individual workstations is thereby isolated and studied with respect to various dispatching and batching strategies and evaluated in terms of job queue times in the system.

An intermediate model is then developed which incorporates the three critical hubs into one queueing network. Again, the off-hub steps are represented as a series of time delays. The ideas and experience gained from the experiments on the isolated hub models are used to refine and focus the experimentation on this intermediate model.

The experience gained from the initial and intermediate models is used to provide direction to design and perform a series of experiments on a manufacturing model incorporating all the workstations in a wafer fabrication line. In addition to narrowing the focus from the intermediate level of experimentation a series of experiments incorporating machine failure and variable demand

distributions is performed to demonstrate the ability of these models to explore more realistic scenarios.

2. LITERATURE REVIEW

An overview of experimental research into job shop scheduling was needed to establish a framework for further study. Specifically, the areas of research reviewed include:

- 1) - the evaluation of job dispatching methods in flow and job shops.
- 2) - simulation based scheduling, planning, and modelling of integrated circuit manufacturing facilities.
- 3) - aspects of sampling techniques in simulation modelling.

2.1 Job Dispatching in Flow and Job Shops

Early work using discrete simulation models to study job shops first appeared in the late 1950's. Some important early papers include the work of Jackson (1957) and Rowe (1958). Since then countless papers have been published in the field. As computational performance grew, so did the scope and complexity of the studies.

Nelson (1967) expanded on previous work by considering both labour and machines as scarce resources in the manufacturing system model. Three machine assignment rules, 'first in system first out', (FISFO), 'first come first serve', (FCFS), and 'shortest processing time', (SPT), were combined with five labour assignment rules to define the sequence of job service in the system. His factory model, implemented in the simulation language SIMSCRIPT, consisted of a series of machines and labourers together with a set of jobs whose process flows were determined by a Markov process. Nelson examined the effect of experimental factory configurations on the mean and variance of the average time spent in the factory per job. His decision to use a dual resource constrained model as a vehicle for

evaluating dispatch policies was a milestone, yet, there remains insufficient work in this area.

Elvers (1974) explored another area that had until then received little attention. The majority of simulation studies had assumed a certain type of arrival distribution for jobs entering the work place. Elvers studied the relative effectiveness of various dispatch rules and their sensitivity to different arrival distributions. He performed 160 simulations using ten dispatch rules and sixteen stochastic arrival distributions in a job shop consisting of eight machines with job routings and processing times being assigned probabilistically. Using job tardiness as the evaluation criteria, Elvers concluded that the arrival rate distribution did not impact on the relative effectiveness of dispatch rules.

Panwalker and Iskander (1977) published a paper defining more than one hundred priority dispatch rules drawn from two decades of sequencing and scheduling research. The rules were grouped into classes. The first class consisted of simple priority rules, or a combination of simple priority rules and weighted priority indexes. The next class, called heuristic scheduling rules, was defined as those rules which take into account more complex information such as anticipated machine loading or alternate job routing. Finally, a miscellaneous category was created for rules which did not fit the above definitions. They cross-indexed the rules with literature references summarizing the references on, (1) the type of study performed, (2) the performance measures used, and (3) the rules covered in the study. Their conclusions contained a recommendation for more research based on real problems. The recommendation was based on their concern about the hypothetical nature of the manufacturing models used in most studies.

Blackstone et al (1982) surveyed results obtained in the job scheduling literature and tried to assess the 'state of the art' in the study of dispatch rules and secondly to identify, if possible, 'good' dispatch rules. They found that flow time, lateness and tardiness

were the performance measures most often used in the studies. However, these criteria ignored the total cost associated with manufacturing, and the authors claimed that an accurate evaluation of dispatch rules would be delayed until a measure of total cost could be produced. They continued the assessment by pointing out the problems associated with analytical methods of studying queueing systems. The second part of the study summarized many of the results in the literature to that date. They grouped the dispatch rules into four classes:

- (1) - rules involving processing times.
- (2) - rules involving due dates.
- (3) - simple rules using neither due dates nor processing times
- (4) - rules mixing two or more of the first three classes.

They found the SPT rule to be superior in most cases to the other top performers, which included 'earliest due date', FCFS and 'critical ratio'.

Studies since Blackstone et al (1982) that have surveyed the available literature include Kiran and Smith (1984a), Sen and Gupta (1984), and Gupta et al (1988). Kiran and Smith surveyed the shop performance criteria used by researchers in simulation studies of job shop scheduling. They classified the performance criteria based on, (1) use of job completion times, (2) use of job due dates, and (3) use of costing techniques. Sen and Gupta examined scheduling methods whose performance criteria involved job due dates. Gupta et al concentrated on scheduling rules used in flexible manufacturing systems (FMS). Their paper is discussed in more detail later in this chapter.

Dar-El and Wysk (1982) presented a well planned methodology for evaluating job shop dispatch rules. They considered six dispatch rules and several confounding factors in examining the relative performance of these rules. Work flow characteristics, processing time distributions, due date distribution, and shop load levels were

examined. An important aspect of the paper is the statistical rigour used in the design and analysis of the experiments. Pairing of results was performed on every run and ANOVA calculations were performed to determine relative significance of factors and results. Dar-EI and Wysk used tardiness measures to compare the performance of the dispatch rules. Here, the SPT rule performed exceptionally well with the 'work in next queue', (WINQ), FCFS, and 'earliest due date' rules vying for second spot. The paper evaluates the dispatch rules in a systematic way and is a model for further work in this area.

Smith and Seidmann (1983) examined due date selection procedures. Since many dispatch rules are based on some form of job due date, the method by which these due dates are set has an impact on the performance of these rules. They concluded that due date assignment procedures do play a role in the ranking of dispatch rules and that simulation studies rarely, if ever, use due date assignment methods comparable to those used in industry. Several studies since 1983, including Baker (1984), Sen and Gupta (1984) and Bookbinder and Noor (1985), focussed on dispatch rules involving due dates.

Huang et al (1984) studied machine and labour limited models and the effect of shop load conditions on the relative performance of dispatch rules. The study included four dispatch rules, three labour assignment rules, and three shop load levels, and used mean job flow time as the primary performance measure. Their general conclusions were that machine assignment rules have a larger impact on performance measures than do labour assignment rules. Among the dispatch rules they found that the SPT rule delivered the best performance, although at lower shop load levels the margin of difference was diminished.

Neelamkavil and Thomson (1986) introduced a manufacturing planning system that incorporated a scheduling system and modelling interface. As a result, a framework for evaluating dispatch rules with respect to their differing strengths and management

performance objectives was established. The framework includes factors affecting dispatch rules, a classification of rules, an outline of the bias in studies involving dispatch rules, and finally, a rule selection strategy. An overview of the planning system, developed jointly by ICAM Technologies Corporation and the National Research Council of Canada, outlines the three key modules of the system. First, the scheduling module is used to assess production plans and to produce detailed shop schedules. Second, the modeller performs discrete simulations based on shop schedules, 'what if' scenarios, and planning requirements. Lastly, the shop monitor is used for processing shop floor data and updating the manufacturing database.

Gupta et al (1988) reviewed the current list of dispatch rules and their role with respect to scheduling flexible manufacturing systems (FMS), versus more traditional manufacturing operations. They discuss the objectives of FMS scheduling and its relatively unique constraints. The remainder of the study includes a large list of dispatch rules, an outline of related results found in the job dispatching literature, and a cross indexing of references versus dispatch rules. They concluded that dispatching and scheduling strategies are highly system dependent and that rules should be adopted and modified as required by the individual systems.

2.2 Modelling IC Manufacturing

Discrete simulation analysis of integrated circuit manufacturing first appeared in Atherton and Dayhoff (1984) and was the first of a series of articles authored by them. They referenced three earlier sources, Fishman (1978), Kleinrock (1976), and Till and Luxton (1982), none of which are focussed on discrete simulation analysis of IC manufacturing.

In their 1984 paper Dayhoff and Atherton introduced their simulation model and analytical approaches to measuring IC factory performance with the statement:

" Complex process flows, large numbers of different processes, and rapid change in processes makes the VLSI foundry an excellent candidate for analysis with simulation."

They used illustrative examples to demonstrate the potential of their model for analysis of inventory levels, cycle times, and factory capacities. The models and ideas presented in their 1984 paper provided a base for the subsequent articles published by them .

Dayhoff and Atherton (1986) described a method for comparison of dispatch rules using the models and ideas developed in their 1984 paper. Three rules, 'lowest-step-first', 'highest-step-first', and 'round-robin', were evaluated against several performance measures across a set of increasing factory demands and equipment failure scenarios. The 'highest-step-first' rule performed better than the other two, especially, for high factory demand rates where the emphasis on 'pulling' from the end of the system ensured a steady stream of lots leaving the factory. Atherton and Dayhoff described the use of 'signature analysis', a technique of plotting a performance measure against a set of simulations run at successively increasing factory demand rates, for studying inventory, cycle time, and throughput trade-offs.

Burman et al (1986), at AT&T, reviewed three modelling approaches used for performance analysis of IC manufacturing lines. These approaches included discrete event simulation, deterministic capacity analysis, and more complex queueing network analysis. According to Burman, the main advantage of simulation over the other two approaches was:

" that it allows us to address a wide variety of complex problems that would otherwise defy analysis."

Some drawbacks of the other methods included a very limited scope of issues which can be addressed by deterministic capacity models and the extensive analysis of a manufacturing system needed for the queueing network models to produce reasonable results. Although Burman favoured simulation as the model of choice, he recognized the need for the short development time of small deterministic capacity models and the fast performance of queueing network models.

Edwards et al (1987) and Poulsen and Sutcliffe (1987) presented, in two papers, a modular simulation and scheduling system and a set of performance analysis experiments based on a Northern Telecom IC fabrication line in Nepean, Ontario. The first paper described the simulation and scheduling system and included a set of results demonstrating the ability to study the effects of equipment failure, high priority lots, shift and personnel policies along with dispatch rules and factory demand rates. As with Dayhoff and Atherton (1984), Burman et al (1986), and O'Grady and Harrison (1985), the system was built around a generic model that can be used across multiple manufacturing lines, with system specific information such as process flows and equipment statistics being supplied by input data files. This methodology permits application to other factories with a much smaller start-up time.

In the second paper, Poulsen and Sutcliffe used designed experiment techniques to characterize furnace batch size effects. Results for multiple performance measures were analyzed using Taguchi and 'best compromise' methods, (See references in Poulsen and Sutcliffe (1987)), to determine the 'best' load sizes for the twelve furnaces then employed in the factory plan. The load size of a furnace varied between one and four lots and had significant effects on the performance measures of cycle time, capacity, and operator utilization, which were included in the study. Taguchi analysis was used to determine the relative effect of the different load sizes on the performance measures while the 'best compromise' analysis was used to determine the various load size settings which delivered the

'best' compromise among the three performance measures. The term 'best compromise' refers to the problem of competing manufacturing goals and the determination of a factory policy that delivers acceptable performance of the measures that characterize these goals. A mixture of batch sizes delivered optimal results.

Glasse and Resende (1988) used a discrete simulation model to study the performance of a job release algorithm for IC manufacturing. Most studies have used either fixed or exponential inter-arrival times for jobs entering the simulation model. Glasse introduced a method for job release he terms 'starvation avoidance'. His method assumed that there is a single capacity limiting bottleneck in the system and that control of the queue at the bottleneck is key to an efficient job release policy. The idea is to release new lots into the system so that they arrive at the bottleneck machine just in time to prevent it from standing idle. The goal of this policy is to maintain high utilization of the bottleneck machine while maintaining a low 'work in process' inventory. The trade-off between inventory and utilization is determined by the re-order point in the bottleneck queue. Too high a re-order point results in excess queueing while too small a re-order point will result in a higher probability of idling the bottleneck machine. His method outperformed the three other policies he examined, especially when the system was close to capacity and the bottleneck was 'near' the start of the system.

Wein (1988) studied the impact that scheduling may have on IC manufacturing. His study included seven dispatch rules, four job release mechanisms, and three wafer fab layouts. His results showed that job release mechanisms have a greater effect on average cycle time than lot dispatch rules. Among release strategies he found that deterministic mechanisms delivered significantly lower average cycle-times than the stochastic mechanisms. He found a mixture of results among the dispatch rules although, the difference in performance among the rules was greater for the random than the deterministic release mechanism. In summary, the deterministic

release policy clearly outperformed the other policies, and dispatch rule advantages were dependent on the release mechanism and to a lesser degree the wafer fab layout.

2.3 Sampling Methods in Discrete Simulation

Research of sampling techniques for discrete simulation has centred on the problem of determining an accurate estimate of the mean of the performance measure of interest. The choice of sample size, the required number of samples, and the initialization and stopping criteria for a simulation have all been studied.

Generally the mean of the sample :

$$\hat{x} = \frac{1}{n} \sum_{i=1}^n x_i ;$$

and the variance of the sample :

$$s^2 = \frac{\sum_{i=1}^n (x_i - \hat{x})^2}{n - 1}$$

n = sample size
 x_i = sample element
 \hat{x} = sample mean
 s^2 = variance of the sample

provide a satisfactory estimate of the true mean when the population elements are uncorrelated. However, when the population elements are correlated the variance is a much harder statistic to estimate accurately. This is important since the variance is helpful in determining how well the sample mean estimates the true mean. Unfortunately, simulation results tend to have significant positive auto-correlation. (See Kleijnen (1984)). This is easy to visualize in IC

models as the lots are all waiting in the same queues and competing for the same scarce resources.

A simulation of a flow shop generally begins with the system model in an empty state and as time progresses performance measures such as inventory and job throughput reach levels which tend to fluctuate around their long term averages. Researchers are usually concerned with obtaining accurate estimates of the long term averages of these performance measures. The initialization period in a flow shop simulation is an interval of time at the start of a simulation run during which the system is allowed to reach a 'steady-state' before the system performance is characterized. Statistics which include results collected before the system has reached a steady state will be biased towards the early data. Thus, it is important that the initialization period be sufficiently long.

A general data collection method begins after an initialization period and continues until the termination of the run. The problem with this method is that jobs remaining in the system at the end of the simulation run are not included in the statistics. This problem is known as censoring data and may have a real effect on the results. (See Blackstone et al (1982)).

Fishman (1971) presents a methodology for determining the appropriate sample size needed to estimate the population mean within a specified degree of precision. This method does not require interaction by the researcher during the simulation and permits adjustment for initialization bias. A flowchart in the paper outlining the procedure is summarized here :

1. Collect observations from an initial simulation run
2. Calculate the variance criteria for the observations
3. Determine number of additional observations required to obtain the desired accuracy

4. Collect additional observations by continuing the simulation and return to the second step repeating the iteration sequence until the desired accuracy is obtained.

The method can be implemented as a computer algorithm for use during simulation runs, but, determining the initial sample size is a problem noted by Fishman. See also Blackstone et al (1982) for a summary of the method.

Another method of calculating an estimate of the population mean involves grouping the simulation into a series of sub-runs and using the sample averages from the sub-runs as an estimate of the population mean. This method, known as the 'batch means method', requires the number of samples and the number of observations per sample to be determined consistent with the requirement of independent batch means.

Fishman (1978) presented a version of the batch means method using the von Neumann ratio for testing the independence of the batches. The sample size is determined by an iterative procedure using von Neumann's theory to test for acceptance, at a predetermined confidence level, at each iteration. When the sample size has been determined, the number of batches is calculated as N/M where N is the number of observations and M the sample size. Fishman chose this method for its simplicity and ease of use over three others, independent replications, spectrum analysis, and autoregression.

Kleijnen et al (1982) reviewed the batch means method based on the von Neumann test. He was interested in the effectiveness of the von Neumann test in determining the independence of the sub-runs required by the batch means method. He concluded that the power of the test was significantly reduced when the number of sub-runs was less than one hundred. Consequently, the possibility of a researcher

accepting a false verdict of independence increases as the number of sub-runs drops to less than one hundred.

Adlakha and Fishman (1982) studied starting and stopping criteria for simulations. In the paper he presents two rules; one for starting collection of data and the other for stopping collection of data. Both rules are based on selecting the point during the simulation where the difference between the theoretical activity level, defined as :

$$\frac{\text{job arrival rate}}{(\# \text{ of servers})(\text{service rate})}$$

and an estimated activity level, defined as :

$$\left(\sum_{i=1}^n S_i \right) / \left(c \sum_{i=1}^n A_i \right)$$

c = number of servers

S_i = service time of i'th arrival

A_i = elapsed time between arrival of the (i-1)st and i'th jobs.

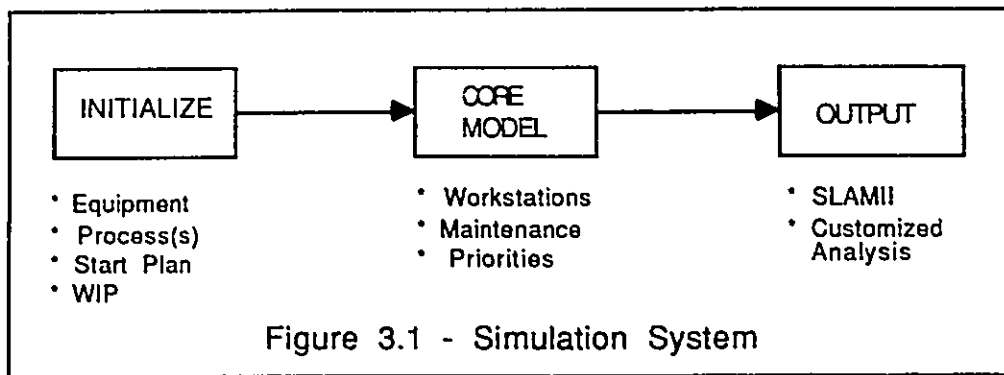
falls below a predefined level, subject to certain preconditions which differentiate the rules. The initialization rule will trigger data collection when the system reaches a steady state of congestion. Fishman claimed that when the second rule is used in conjunction with the first, that the rules put to advantage the initial conditions and skewness.

Kleijnen (1984) reviewed the progress of statistical analysis in the area of steady-state simulations. He covered the topics of autocorrelation in queueing data, replicated runs, initialization phases, and batching with sub-runs. It is a good review of much of the work performed in this area and provides an introduction to the issues and methods being studied.

The collection of papers reviewed in this chapter provide the technical background and introduction to the research presented in this thesis.

3. GENERAL MODEL

A simulation system for IC manufacturing developed at Northern Telecom is used to perform the simulation experiments in this thesis. The system is implemented in the SLAM II simulation language with a modular design which accepts as input manufacturing data including process flows, machine models, work in process, and product start plans. Figure 3.1 is an illustration of the model's organization.



The core of the system includes a description of the workstations and the priority schemes including the dispatch and batching rules. The system's output consists of standard SLAM II (Pritsker 1986) reports as well as a detailed log of the complete simulation which is used for customized analysis. The modularity of the system permits easy updating and execution of simulation experiments. Edwards et al (1987) provides a detailed description of the system. A data model provided by Northern Telecom of an IC process provides a real factory as a basis for the studies in this thesis.

3.1 Modelling Approach

The full data model includes a single process sequence of 129 steps which requires 53 types of processing machines. Each step is described in three ways: (1) its place in the process sequence, (2) the machine type required at that step, and (3) the amount of processing

time required on that machine. Each workstation consists of a single machine except for the furnace and stepper hubs, described in Chapter 1, which have multiple units operating in parallel. The furnaces, labelled types A and B, include four type A's and two type B's. There are two stepper machines in parallel.

3.1.1 Workstation Model The simulation model maintains a single queue for each workstation. A workstation model contains machines of the same type and each machine in a workstation can service only a single lot at a time. The exceptions are the furnaces which can process batches of several compatible lots. See Section 3.1.4 for a discussion of the furnace workstations. The state of a workstation during a simulation is described in two ways: (1) whether it is busy or idle, and (2) the size of the queue. Two events can change the state of a machine workstation:

- 1) Arrival of a lot at the workstation.
 - a) - a request is made for service on the machine type.
 - b) - if a machine is free the lot is serviced, otherwise, it is placed in the queue for that machine type.
- 2) Departure of a lot from the workstation.
 - a) - a lot finishes service at the workstation by freeing the machine and leaving for the next process step.
 - b) - the machine queue is searched for the highest priority lot and servicing begins.
 - c) - if the queue is empty the machine awaits the arrival of the next lot.

The service sequence, or priority, of lots in the queue is determined by the dispatch rule applied to the queue.

3.1.2 System Models The Northern Telecom facility includes hundreds of processing steps in several evolving processes. Multiple products, stochastic machine failures, machine maintenance, labour policies, and operating procedures all contribute to a complex model. To gain basic insight into the effects on factory performance of dispatching and batching methods at the major hubs, many of the above factors have been removed in a set of less complex models developed for this study. These models contain only the essential information required to preserve the congested, multiple visit nature of the hubs. The major assumptions of the models include:

- single process and product type
- 24 hour days and 7 day weeks
- 100% operator availability
- 100% machine reliability except in Section 6.3

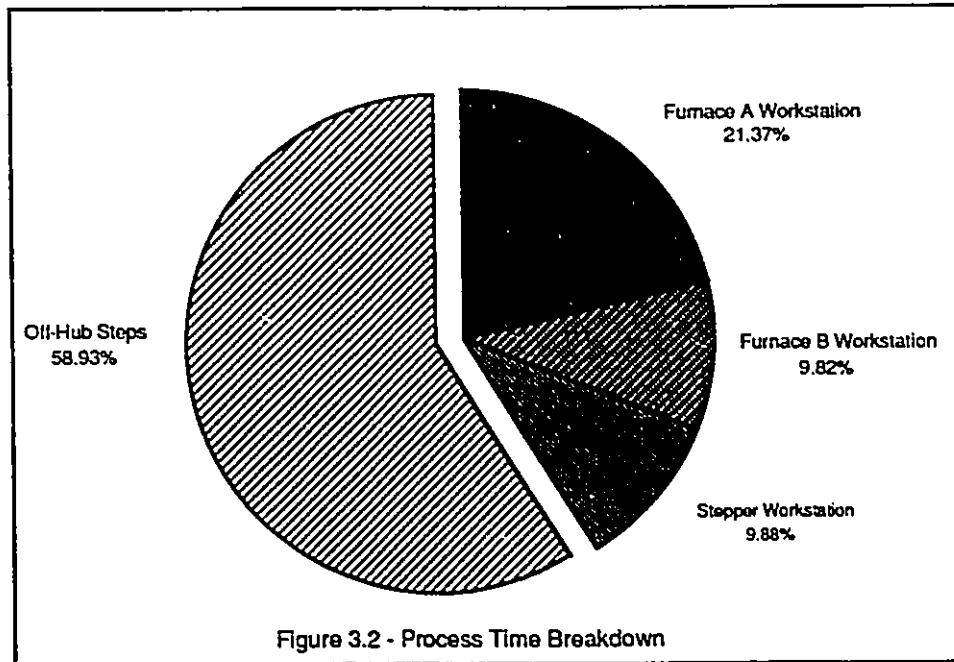
Except for a set of simulations dealing with random job arrivals the models in this thesis are completely deterministic. Processing times in IC manufacturing are essentially deterministic with the variability in the system tending to come from machine failure, operator availability, and organizational procedures. (See also Burman (1986)).

Three levels of model complexity have been studied. These levels are:

- First - the 3 major hubs considered in isolation
- Second - the 3 major hubs in a single network model
- Third - a full process model with all workstations explicitly represented

The three workstation hubs selected for study were determined with consideration of the number of visits and the amount of processing time required by a lot at the workstations. There is a total of 180.2 hours of processing time required throughout the complete sequence of 129 steps. The breakdown of the total

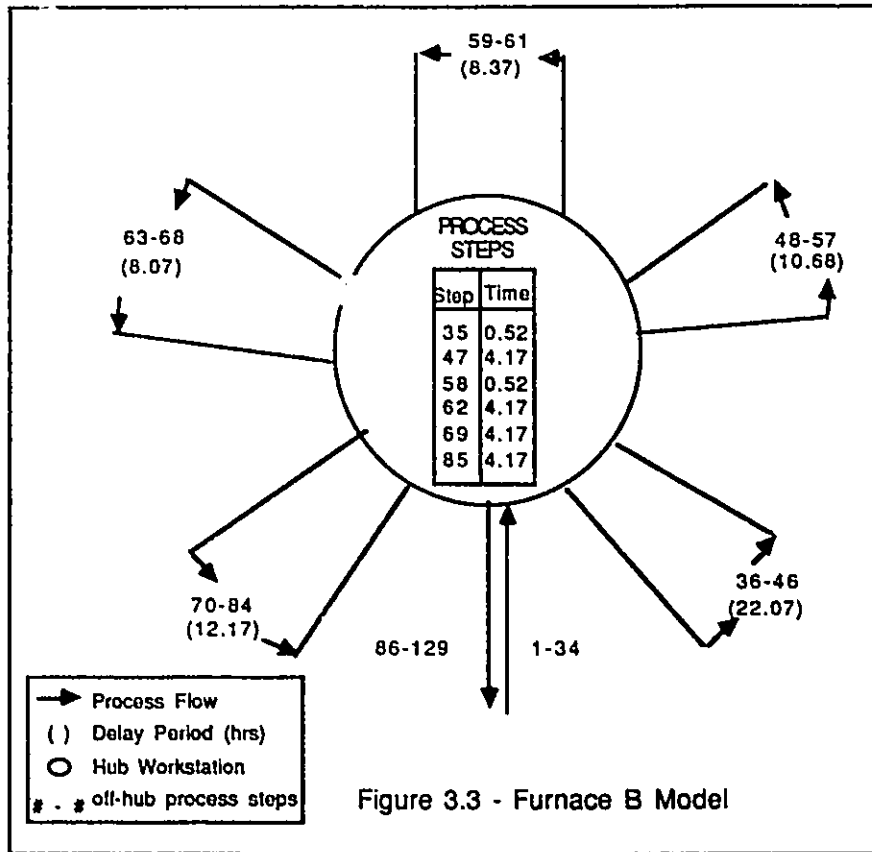
processing time required of the hub and 'off-hub' workstations is illustrated in Figure 3.2. The 'off-hub' steps are those process steps falling between visits to the hub machines under study.



Isolated Hub Models At the first level of modelling the three hubs are isolated in separate queueing models. These models do not include those process steps which come before the first visit or after the last visit to the hub workstation. Accordingly, on the furnace B model illustrated in Figure 3.3 steps 1-34 and steps 86-129 are not represented in the model. The off-hub steps are modelled by a delay period representing the time that would normally be required by the lot to complete the off-hub steps. The off-hub delay periods were calculated with a simulation of the full factory using the first-come-first-serve dispatch rule, a batch size of one lot at the furnace hubs, and a job arrival rate equivalent to the 95% utilization level of the system bottleneck (the furnace A workstation). Their values were defined as the average time between hub revisits of 500 lots monitored during the simulation. The purpose of the delay periods is

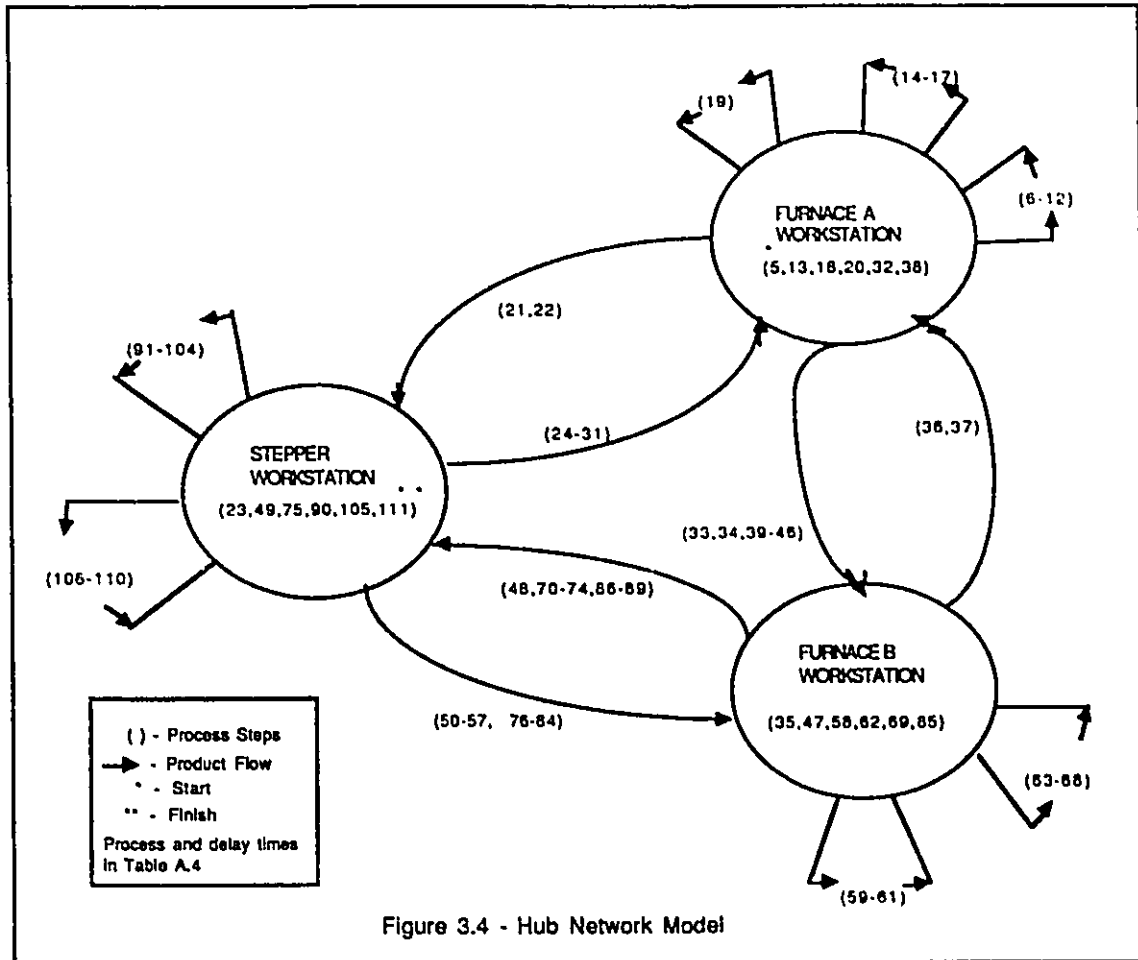
to shift the focus of the experiments onto the hub in question and the set of dispatch and batching rules employed there.

The process sequence of a lot in an isolated hub model consists of alternating periods of service at the hub workstation, and delay in the off-hub steps. The model of the Furnace B workstation is illustrated in Figure 3.3; the models for the stepper and furnace A workstations are given in Appendix A.1. In Figure 3.3 the furnace workstation is represented by the central circle and the process times for each visit to the workstation are listed. The directed line segments represent the 'off-hub' process flow. The set of process steps and delay periods are given for each group of 'off-hub' steps.



Hub Network Model A similar schema is used for the second modelling level. The three hubs are combined in a single network model, illustrated in Figure 3.4, in which the off-hub steps are

represented in the same manner as in the isolated hub models. Here, the added effect of interaction among the hubs on the performance of various dispatching and batching strategies is of interest. The process steps and process times are tabled in Appendix A.2. Process steps 1-4 and steps 112-129 are not included in the model since they fall before the first visit and after the last visit, respectively, to any of the three hubs.



Full Process Model The final phase of this thesis is the study of the complete 129 step model. The model includes explicit representation of all the steps in the process. The experimental framework of the previous two stages is extended to this model and the performance of the dispatching and batching strategies at the hubs are examined.

The impact of machine failure and job arrival distribution are also explored.

3.1.3 Dispatch Rules There are many frameworks for classifying dispatch rules. (See Gupta (1988)). The dispatch rules studied in this thesis cover several classification categories and serve as a starting point for understanding the types of rules which tend to perform well on systems characterized by hubs. These rules and their definitions are:

1. First Come First Serve (FCFS) - the earliest lot to have entered the workstation queue is the next lot to get serviced.
2. DUE DATE - the lot with the earliest due date in the workstation queue is the next lot to get serviced.

Note: (1) In this thesis a lot's due date is assigned as the date of entry of the lot into the system plus a fixed period of time which is equal for all lots. (2) In this thesis DUE DATE is equivalent to a 'highest step first' policy since a single process and a uniform due date assignment method are used.

3. Shortest Processing Time (SPT) - the lot with the shortest processing time requirement at the current workstation is the next lot to get serviced.
4. Lowest Step First (LOWSTEP) - the lot in the workstation queue at the earliest process step is the next lot to get serviced.
5. Critical Ratio (CR) - the lot with the lowest value of the following ratio:

$$CR = \frac{d_i - t}{s_i}$$

d_i = due date of i 'th lot

t = current time

s_i = sum of remaining processing time of i 'th lot

is the next lot to get serviced.

Note: an alternative definition of the CR rule includes both the remaining processing time and an estimate of the remaining queue time in the denominator. (See Baker (1984) and Blackstone (1982)).

6. Minimum Slack (MINSLK) - the lot with the minimum slack time:

$$\text{slack} = d_i - t - s_i$$

d_i = due date of i 'th lot

t = current time

s_i = sum of actual remaining processing time of i 'th lot

is the next lot to get serviced.

7. Longest Processing Time (LPT) - the lot with the longest processing time requirement at the current workstation is the next lot to get serviced.

Several of the rules have a secondary tie-breaking rule. The tie-breaking rule for FCFS is 'highest step first'. The tie-breaking rule for SPT, LOWSTEP, and LPT is FCFS. In this study each lot has a unique due date and, hence, DUE DATE, CR, and MINSLK do not require a tie-breaking rule.

These rules can be grouped into one or more, sometimes overlapping, classes including: (1) those rules involving due dates, (2) push and pull rules, (3) localized rules, and (4) dynamic and static rules.

(1) Rules involving due dates have been the subject of study by several people including Smith and Seidman (1983), Baker (1984), Sen and Gupta (1984), and Bookbinder and Noor (1985). These rules assume that jobs are assigned expected completion dates either as finished products or for some internal subset of the production process. The rules determine the priority of servicing jobs wholly or

partially on the value of the due date. Three rules defined above, DUE DATE, CR, and MINSLK fall into this category, with CR and MINSLK having only a partial reliance on the due date. The simulations in this thesis assign the due date for each lot as a fixed period of time added to the date of entry of the lot into the system. Since there is only a single product type and a single process sequence in this study the period of time allocated does not have an effect on the relative priority of lots as long it is applied uniformly to all lots.

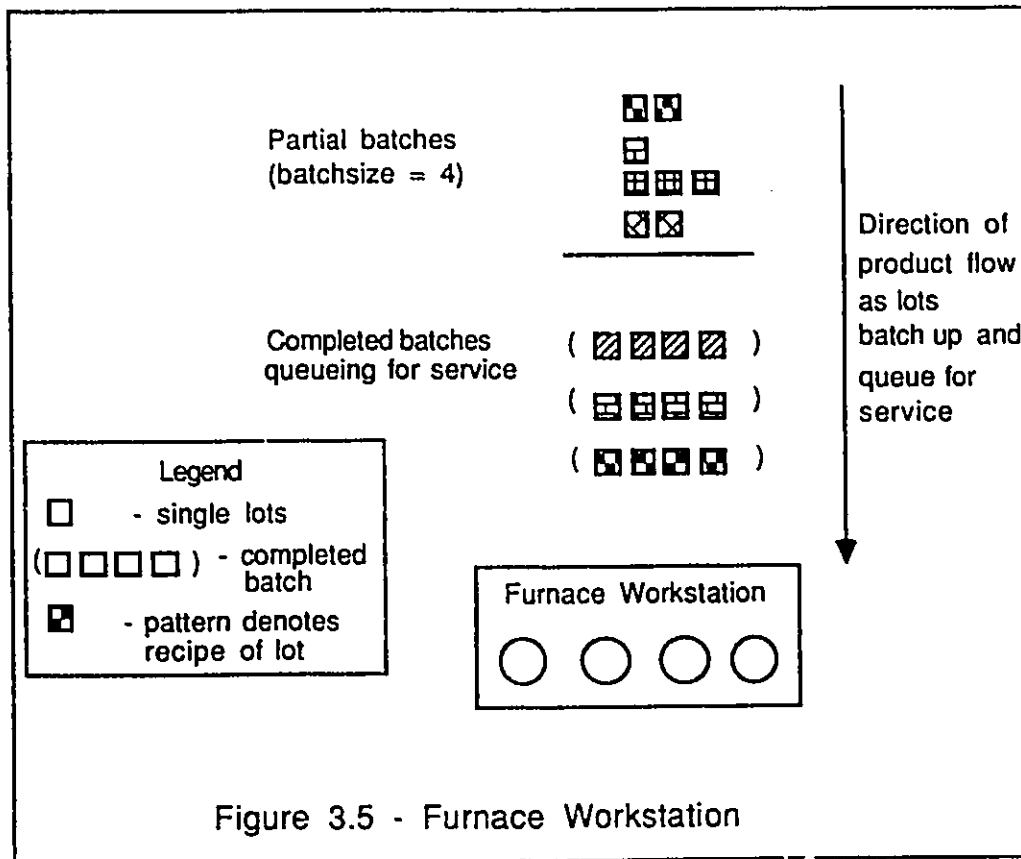
(2) Pull rules are characterized by pulling jobs out the back-end of the production line. Pull rules emphasize finishing jobs by giving priority to lots which may be late, have little slack, or are near completion. DUE DATE, CR, and MINSLK rules fall into this grouping.

Push rules emphasize aspects that do not focus on finishing jobs. These aspects might be the equality of jobs, as with the FCFS rule, or the completion of the most jobs by machine, as with SPT. The LOWSTEP rule, with its emphasis on moving product into the second half of the process, also falls into this category as does LPT.

(3) FCFS, SPT, and LPT fall into a classification that uses only extremely localized information. These rules consider no information outside of the local queue. The attractiveness of these rules is their simplicity and ease of use, and in the case of SPT the delivery of reasonable performance levels in many studies. (See Kiran and Smith (1984)).

(4) Dynamic rules may change a lot's relative priority in a queue as time passes, even when the composition of the queue may be static. Conversely, static rules do not change the relative ranking of lots once the lot has entered the queue. The CR rule is a dynamic rule, while the MINSLK rule which uses the same information is a static rule.

3.1.4. Furnace Batch Sizes Furnaces can process several lots at a time. Generally, compatible lots are said to have the same recipe at a process step when their machine operating requirements such as temperature, gas flows, and machine time required are the same. In this study since only a single process is modelled compatible lots are defined to share the same step in the process sequence. The batching decision at a furnace determines the size and service priority of the batches. This thesis considers batch sizes of one, two, three, and four lots, in conjunction with the seven dispatch rules defined earlier. The batching and service mechanism at a furnace works as follows: (1) a lot arriving at a furnace is placed in a compatible batch if one exists, otherwise a new batch is created, (2) when a batch is completed it is placed in the furnace queue. The service priority of a completed batch is determined by the lot with the highest priority in the batch according to the dispatch rule in effect. See Figure 3.5 for a pictorial view of the mechanism.



3.1.5. Performance Evaluation The impact of different batching and dispatching strategies on queues in a manufacturing system varies, so a method of evaluating the performance of these strategies is required. Many objectives are considered to be important in the operation of a manufacturing facility. A mandate of the Northern Telecom facility is to introduce new integrated circuits into Northern Telecom's many telecommunications products. There may be several new IC designs entering the production line every week and the pressure to deliver these new products in as short and predictable a time as possible is substantial. This pressure results from the competitive advantage that Northern Telecom can receive from early introduction of new product designs into the marketplace. Chen et al (1988) also addresses the importance of cycle-time in IC manufacturing.

The delivery date of a product is significantly affected by the cycle-time of the wafer-fab. The resulting pressure to reduce the mean and variability of the wafer-fab cycle-time increases the importance of understanding how dispatch and batching rules can help achieve these goals. The role of dispatch and batching rules can be examined by looking at how they affect cycle-time. Cycle-time can be defined as :

$$\text{Cycle-time} = \text{Machine-processing-time} + \text{Queueing-time}$$

For the purposes of this study, the processing time will include the actual machine processing time and the delay periods, (See Chapter 3.1.2), as they are a fixed part of each model. The dispatch and batching rules cannot reduce the processing time but they do affect the queueing time. The evaluation of simulation results will use the mean and standard deviation of the queueing time as the primary performance measures, as they provide a direct measure of how dispatch and batching rules affect cycle-time.

3.2 Sampling Considerations

There are many methods of sampling data in discrete simulation, several of which were discussed in the literature review, including Kleijnen (1984). The sampling procedure used in this thesis attempts to reduce the bias and variance of results using methods summarized in Blackstone (1982).

Blackstone was concerned with sampling procedures which censor data, and whose effects may include favouring certain types of dispatch rules. Simulations, in general, are terminated after a set length of time, at which point statistics including the sample mean and variance are calculated. Jobs remaining in the system at the end of the simulation run are not included in these statistics and the remaining results are said to be censored. For example, a dispatch rule with a large cycle-time variance will tend to leave proportionately more 'late' than 'early' jobs in the system at the end of the simulation. The late lots will not be included in the statistics, and so the sample mean will be biased lower than the population mean.

Another source of bias found in simulation results is due to the collection of statistics during the initialization of the system. Until the system has reached a steady state, jobs will tend to have shorter cycle-times and more predictable performance, as they have smaller queues to contend with. Research into this problem is discussed in Chapter 2.3.

Blackstone (1982) described a method for eliminating the censoring problem. The method requires sampling a set of numbered lots from the 'centre' of a simulation run. After an initialization period a sample of N consecutive lots are numbered as they enter the system. These N lots are screened out as they leave the system and used for calculation of system statistics.

In this study, the sample size, N , and the initialization time are based on the results of a simulation run which monitored the system inventory level and cycle-time. The simulation run was performed on the 129 step full process model at 95% of its theoretical capacity with a furnace batch size of one lot. This is a rate of one lot every 10.1 hours or 16.6 lots per week. The capacity and start rate of the model are relative to the assumptions of the model, which include 24 hour a day operation, 100% product yield, 100% machine reliability, and 100% operator availability. The goal of the simulation was to determine the point at which the system reached a steady state and the size of the sample required to reasonably cover any variability in the results.

Figure 3.6 is a plot of the inventory in the system versus simulated time for the scenario described above. After 220 hours the system has reached an inventory of 20 lots. There is little variation in the inventory thereafter with the majority of movement between twenty and twenty-two lots. Figure 3.7 is a plot of the cycle-time for each lot against their exit time from the system. Results of these inventory and cycle-time measures are evidence that an initialization time of 1667 hours, (100,000 minutes), and a sample size of 500 lots provide a confident estimate for reporting steady state results. The sample size of 500 lots represents 5060 hours or approximately 2.5 years of 40 hour weeks of factory operation. All queue time results for the subsequent simulations, (Chapters 4-6), were calculated as the mean queue time of 500 lots sampled after a 1667 hour initialization period.

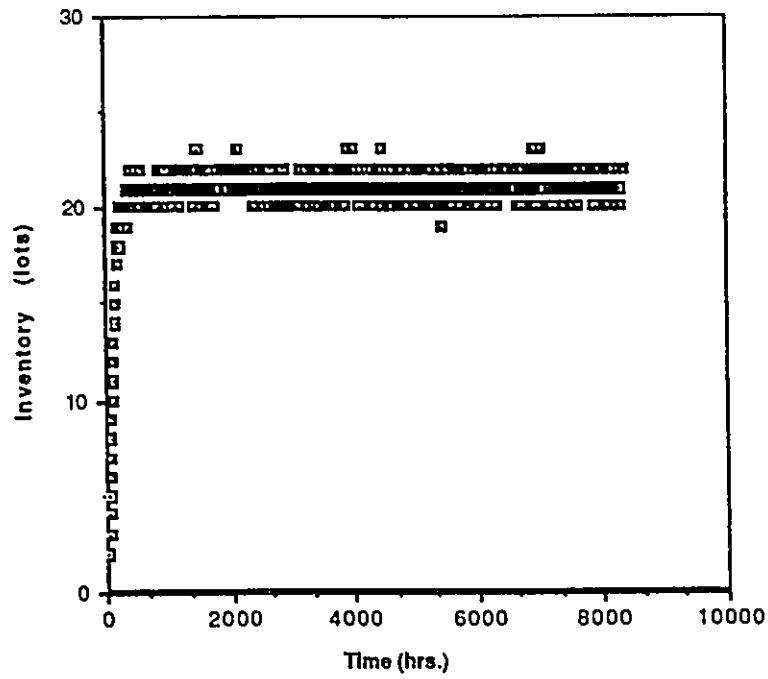


Figure 3.6 - Inventory Trace

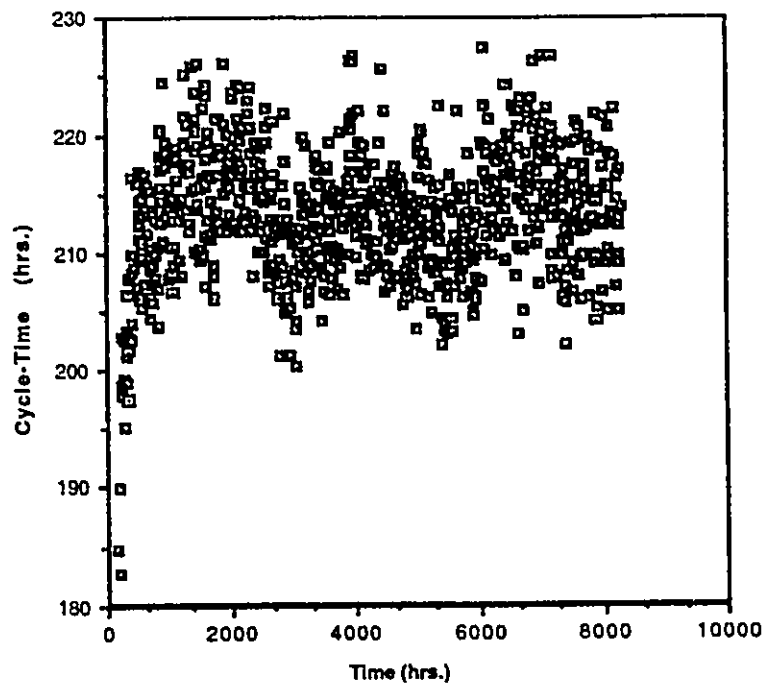


Figure 3.7 - Cycle-Time Trace

4. ISOLATED HUB MODELS

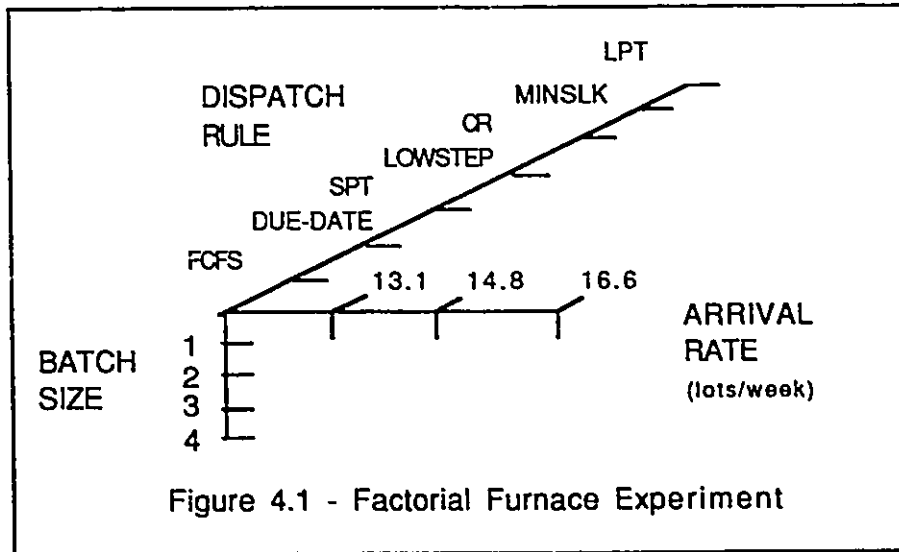
The objective of this chapter is to establish for the isolated hub models the effectiveness of various service strategies encompassing the dispatch rules and furnace batch sizes defined in the previous chapter. The chapter is divided into three sections. The first is an examination of the effect of furnace batch sizes on the two furnace workstation models. The second section examines the stepper workstation model and the third section is an evaluation and comparison of the performance of the dispatch rules on the workstation models.

4.1 Furnace Batch Level Experiments

The main experimental factors for the furnace A and B workstation models are the seven dispatch rules and four batching levels described in Chapter 3. Each dispatch rule can be used with any of the four batch sizes, resulting in 28 possible service strategies for each of the two furnace workstations. The experimental design also incorporates different workstation utilization levels (or equivalently job arrival rates) so that the service strategies may be evaluated at various factory loading levels. Three sets of experiments are performed on each of the furnace models. They are: (1) varied batch sizes at constant job arrival rates, (2) varied batch sizes at constant workstation utilization levels, and (3) comparison of queue time versus throughput for batch sizes of one and two lots.

4.1.1 Constant Job Arrival Rates In this experiment the 28 service strategies are examined on the two furnace models at three job arrival rates. These arrival rates correspond to 75%, 85%, and 95% of the 'theoretical' capacity of the system determined with a batch size of one lot at the furnace A workstation. The theoretical system capacity is the job arrival rate which will occupy the system bottleneck, (the workstation in the system with the lowest capacity), 100% of the time. The system bottleneck when the furnace batch size is one lot is the furnace A workstation with a capacity of 17.45 lots

per week (l.p.w.). Hence, the three job arrival rates in the experiment are 16.6, 14.8, and 13.1 l.p.w. . The job arrival rates are based on a seven day, 24 hour a day work week. The factorial experiment is illustrated in Figure 4.1. Seven dispatch rules are tested at four batching levels and three job arrival rates for a total of 84(=7x4x3) scenarios to be simulated on both the furnace A and B models.



Increasing the furnace batch size has the direct effect of increasing the capacity of the furnace workstation and, thus, lowering the utilization level of the workstation when the job arrival rate is held constant. For example, altering the batch size from one to two lots on the furnace A workstation doubles its theoretical capacity from 17.45 to 34.9 lots per week, while halving its utilization level from 95% to 47.5% based on a job arrival rate of 16.6 lots per week. The results from this experiment should be considered in this context as the furnace batch size is changed while maintaining constant job arrival rates.

Results are presented in Table 4.1 for both furnace models run with a job arrival rate of 16.6 l.p.w. Each table element consists of the mean queue time per lot, as defined in Section 3.1.5, and standard deviation (in brackets). Complete results for the

experiment illustrated in Figure 4.1 are listed in Appendices B.1.1 and B.1.2.

Table 4.1 - Furnace Batching Results*

Furnace Batch Size	Furnace A Model				Furnace B Model			
	1	2	3	4	1	2	3	4
FCFS	2.53 (0.7)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)	8.98 (4.1)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
DUEDATE	5.82 (2.3)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)	17.5 (4.6)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
SPT	2.53 (0.7)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)	5.98 (2.1)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
LOWSTEP	2.28 (0.8)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)	5.93 (3.8)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
CR	2.28 (0.8)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)	13.57 (4.4)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
MINSLK	5.47 (2.2)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)	12.78 (3.1)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
LPT	51.17 (14.5)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)	15.55 (4.8)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)

- * - queue times measured in hours
- job arrival rate = 16.6 l.p.w.
- sampling procedure per Section 3.2

Several interesting observations can be made about the queue times observed at the higher batch sizes.

The first observation is that on either model at a given batch size greater than one the dispatch rules give identical results. The low utilization levels of the furnaces at batch sizes greater than one result in virtually no competition for the furnace resource, and hence the dispatch rule chosen has little or no effect upon the dynamics of the system. (See also Huang et al (1984)). It was noted in the paragraph preceding Table 4.1 that increasing the batch size while maintaining a constant job arrival rate will reduce the utilization

level of the furnace workstation. For example, the results presented for the furnace A model at batch sizes one, two, three, and four lots correspond to utilization levels of 95%, 47.5%, 31.7%, and 23.75% of the furnace A workstation. Similarly, the corresponding utilization levels of the furnace B workstation are 87.4%, 43.7%, 29.1%, and 21.9%.

The second observation is that results for batch sizes two, three, and four on the furnace A model are virtually equal to the batch size two, three, and four results on the furnace B model, respectively. Both furnace models include six revisits for each lot, see Figure 3.3 and in Appendix A Figure A.1, and noting that it is the queue time which is measured and not the cycle-time, the two models appear very similar in structure. This similarity in structure and the argument in the previous paragraph helps explain the equal results for the two furnace models at batch sizes of 2-4 lots.

At least two batch size effects impact the results when the batch size is changed. The first effect (I) is the increased workstation capacity at higher batch sizes mentioned previously. The magnitude of this effect is dependent on the amount of queueing at the lower batch size. i.e. the more queueing there is at the lower batch size the more the higher batch size can help reduce that queueing. The second effect (II) is the increased queue time incurred when a furnace is left idle with incomplete batches of lots present. (See Section 3.1.4). The magnitude of this effect is dependent on the amount of furnace idle time, of which the amount of queueing at the lower furnace batch size is an indication.

The batch size scenarios in Table 4.1 can be divided into two groups: (1) scenarios which caused an increase in queue time when the batch size was increased by one lot, and (2) scenarios which caused a decrease in queue time when the batch size was increased by one lot. The group (1) scenarios include the FCFS, SPT, LOWSTEP, and CR rules changing from batch size one to batch size two. These scenarios had the shortest queue times at batch size one, and, in

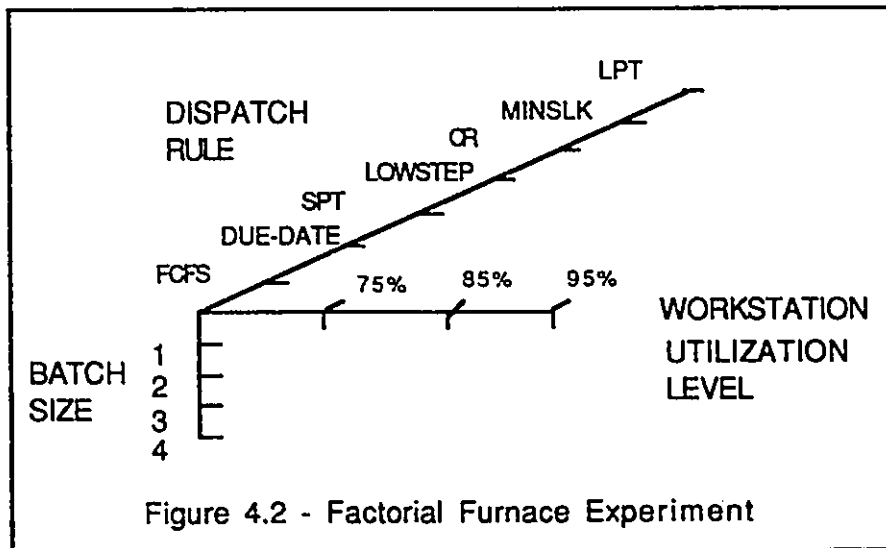
accordance with the argument in the previous paragraph, the capacity effect (I) was evidently small and the queue time effect (II) large. The interaction of the dispatch rules and batch size changes in the group (2) scenarios caused an increase in queue time at the larger batch size evidently due to the capacity effect (I).

4.1.2 Constant Workstation Utilization Rates; This experiment is designed to examine the gain in furnace workstation capacity which is achieved with increased batch sizes. Each combination of the seven dispatch rules and four batch sizes is tested at job arrival rates which correspond to constant workstation utilization levels of 95%, 85%, and 75% of the furnace A workstation at batch size one for a total of $84(=7 \times 4 \times 3)$ scenarios. The job arrival rate must be shifted in conjunction with the furnace batch size to maintain consistent workstation utilization levels. These equivalent job arrival rates and utilization levels are listed in Table 4.2 for the furnace A workstation. The furnace B model is examined at the same job arrival rates as the furnace A workstation.

Table 4.2 - Job Arrival Rates

Utilization Level of Furnace A Workstation	Batch Size (in lots)	Equivalent Job Arrival Rate (lots per week)
95%	1	16.6
	2	33.2
	3	49.8
	4	66.4
85%	1	14.8
	2	29.6
	3	44.4
	4	59.2
75%	1	13.1
	2	26.2
	3	39.3
	4	52.4

The factorial experimental design is illustrated in Figure 4.2 and is performed on both the furnace A and furnace B models.



Results are presented in Table 4.3 for both furnace models at the 95% utilization level of the furnace A workstation. Each column

corresponds to a particular furnace batch size and job arrival rate used to maintain the furnace A workstation at a 95% utilization level. (See Table 4.2). The mean queue time and standard deviation (in brackets) are reported. Complete results to the experiment illustrated in Figure 4.2 are listed in Appendices B.1.4 and B.1.5.

Table 4.3 - Furnace Batching Results*

Furnace Batch Size	Furnace A Model				Furnace B Model			
	1	2	3	4	1	2	3	4
Arrival Rate(l.p.w.)	16.6	33.2	49.8	66.4	16.6	33.2	49.8	66.4
FCFS	2.53 (0.7)	5.03 (2.7)	5.83 (2.8)	6.32 (2.9)	8.98 (4.1)	10.32 (4.2)	9.77 (3.7)	11.75 (4.5)
DUE DATE	5.82 (2.3)	8.50 (3.2)	9.75 (3.9)	9.92 (3.4)	17.5 (4.6)	20.50 (6.2)	21.97 (6.5)	22.85 (6.4)
SPT	2.53 (0.7)	5.03 (2.6)	5.83 (2.8)	6.32 (2.9)	5.98 (2.1)	8.52 (3.3)	9.45 (3.6)	10.05 (3.4)
LOWSTEP	2.28 (0.8)	4.82 (2.7)	5.73 (2.8)	6.07 (3.0)	5.93 (3.8)	8.98 (3.8)	9.95 (4.5)	10.25 (4.0)
CR	2.28 (0.8)	5.20 (2.7)	5.90 (2.8)	7.20 (3.2)	13.57 (4.4)	15.98 (5.4)	18.25 (5.4)	18.42 (6.3)
MINSLK	5.47 (2.2)	5.03 (2.6)	5.88 (2.8)	6.28 (2.9)	12.78 (3.1)	15.02 (4.3)	16.27 (4.5)	16.23 (4.2)
LPT	51.17 (14.5)	42.87 (12.9)	40.43 (11.3)	37.05 (11.7)	15.55 (4.8)	17.03 (5.4)	19.45 (5.4)	18.35 (5.2)

- * - queue times measured in hours
- arrival rate = 95% utilization level of furnace A workstation.
- sampling procedure per Section 3.2

Except for four instances in the furnace A results and two instances in the furnace B results, increasing the batch size caused an increase in queue time. Four of the six exceptions occurred with the LPT rule. However, LPT caused extremely large queue times relative

to the other rules and is accordingly not of further interest as a practical dispatch rule. Over a finite period of time LPT minimizes a workstation's throughput rate by always choosing the longest job available in the workstation queue. This is one cause of the long queue times recorded by LPT.

The increases in queue time must be considered in conjunction with the corresponding increase in system throughput as measured by the job arrival rate. Two of the factory's prime objectives are to reduce the cycle-time and increase the throughput. With the direct link between cycle-time and queue time, a trade-off between cycle-time and throughput is apparent in these results. At batch size one SPT and LOWSTEP cause some of the lowest queue times on both furnace models and as reported their queue times increase as the batch size is increased. An interesting aspect of SPT and LOWSTEP's results is that the queue time versus batch size curve for both rules is concave, or in other words the increase in queue time is largest during the change from batch size one to batch size two and is smallest during the change from batch size three to batch size four. The other rules, ignoring LPT, exhibit behavior of this kind but not consistently on both models or at all batch sizes. The goal for a particular system is to find the combination of utilization level, dispatch rule, and batch size that delivers the cycle-time and throughput combination which best meets the factory's needs.

4.1.3 Stepper as the System Bottleneck In the full process model when the furnace workstations are operating with batch sizes greater than one the stepper workstation replaces the furnace A workstation as the system bottleneck. This experiment is designed to examine the increased capacity of the isolated furnace models at batch size two while recognizing the capacity constraint placed on the system as a whole by the stepper workstation. This experiment tests each of the seven dispatch rules on both furnace models with a batch size of two lots and at an arrival rate of 17.9 l.p.w., equivalent to the theoretical 95% utilization level of the stepper workstation. There are fourteen simulations performed.

The results for both furnace models are presented in Table 4.4 next to results from Table 4.1 for batch size one and the arrival rate of 16.6 l.p.w. (95% utilization of the system bottleneck at batch size one).

Table 4.4 - Furnace Batching Results*

Furnace Batch Size	Furnace A		Furnace B	
	1	2	1	2
Arrival Rate(l.p.w.)	16.6	17.9	16.6	17.9
FCFS	2.53 (0.7)	4.68 (4.7)	8.98 (4.1)	5.83 (4.2)
DUE DATE	5.82 (2.3)	4.68 (4.7)	17.52 (4.6)	5.83 (4.2)
SPT	2.53 (0.7)	4.68 (4.7)	5.98 (2.1)	5.83 (4.2)
LOWSTEP	2.28 (0.8)	4.68 (4.7)	5.93 (3.8)	5.83 (4.2)
CR	2.28 (0.8)	4.68 (4.7)	13.57 (4.4)	5.83 (4.2)
MINSLK	5.47 (2.2)	4.68 (4.7)	12.78 (3.1)	5.83 (4.2)
LPT	51.17 (14.5)	4.68 (4.7)	15.55 (4.8)	5.83 (4.2)

- * - queue times measured in hours
- batch size 1 = 95% utilization furnace A
- batch size 2 = 95% utilization stepper
- sampling procedure per Section 3.2

On both furnace models the dispatch rules delivered identical results at batch size two due to the low utilization levels of the furnace workstations, as explained in Section 4.1.1. The utilization levels of the furnace workstations are then about 50%, and at these levels there is virtually no competition for the furnace resources and the dispatch rules have little effect on the dynamics of the system.

The change from batch size one to batch size two should be examined with respect to the 7.8% increase in system throughput from 16.6 l.p.w. to 17.9 l.p.w. . The 'best' rules at batch size one on the furnace A model, FCFS, SPT, LOWSTEP, and CR, caused significant increases in queue time during the change to batch size two. However, the two 'best' rules at batch size one on the furnace B model, SPT and LOWSTEP, caused slight decreases in queue time in the change to batch size two. The difficulty of evaluating the service strategies is illustrated here by the mixed results obtained for these two 'simple' models.

4.2 Stepper Workstation Experiments

The experimental design for the stepper model tests each of the seven dispatch rules at four job arrival rates. Three of these arrival rates correspond to the 75%, 85%, and 95% utilization levels of the furnace A workstation when the batch size is one. (See Table 4.2). The fourth arrival rate is 17.9 lots per week and corresponds with the 95% utilization level of the stepper workstation. The fourth arrival rate models the situation when the stepper workstation replaces the furnace A workstation as the system bottleneck when batch sizes greater than one lot are used. The results from these 28 simulations are listed in Table 4.5

Table 4.5 - Stepper Workstation Results

Arrival Rate(i.p.w.)	17.9	16.6	14.8	13.1
Utilization Level	95%	88%	78%	69%
FCFS	7.17 (2.2)	9.12 (2.1)	2.93 (1.3)	5.18 (2.0)
DUE DATE	13.1 (5.4)	6.87 (2.8)	2.53 (1.8)	4.03 (2.2)
SPT	6.98 (2.1)	9.12 (2.1)	2.93 (1.3)	5.17 (5.2)
LOWSTEP	6.55 (1.3)	10.35 (2.7)	3.05 (1.3)	4.60 (2.8)
CR	6.45 (1.3)	8.03 (2.7)	3.05 (1.3)	5.17 (2.0)
MINSLK	7.95 (1.9)	7.53 (2.6)	3.50 (2.0)	3.98 (2.2)
LPT	6.93 (2.1)	9.12 (2.1)	2.93 (1.3)	5.17 (2.0)

- * - queue times measured in hours
- sampling procedure per Section 3.2

One observation is the counter intuitive behavior exhibited by the queue times in Table 4.5. Except in two instances for DUE DATE and MINSLK, all the results follow a similar pattern. The queue times increase in the change from the 95% to 88% utilization level; they decrease in the change from the 88% to 78% utilization level, and they increase in the change from the 78% to 69% utilization level. Intuition suggests that lower utilization levels should result in lower queue times due to the decrease in competition for the machine resources. The 'zig-zag' pattern seen in Table 4.5 has been found in other systems such as the two workstation model (Lane 1988) illustrated in Appendix C. In that example the effect is due to the

phasing period between the time of entry of the new lot into the system and the status of the machine at the first workstation.

4.3 Dispatch Rule Comparison

A central issue in the isolated hub experiments above is the relative performance of the dispatch rules vis-a-vis queue time. The main interest is focussed at high levels of system congestion at which many factories tend to run. Accordingly, a ranking of the performance of the dispatch rules for the experiments described in Sections 4.1 and 4.2 is presented in Table 4.6. Results for all three hub models, with furnace batch size one, and arrival rates of 16.6 and 14.8 lots per week are used.

Table 4.6 - Dispatching Rule Comparison*

Arrival Rate (l.p.w.)	Furnace A		Furnace B		Stepper		Sum Rank	Rank
	16.6	14.8	16.6	14.8	16.6	14.8		
FCFS	2	4	3	2	4	2	17	2
DUE DATE	4	1	7	6	1	1	20	3
SPT	2	3	2	3	4	2	16	1
LOWSTEP	1	5	1	1	5	3	16	1
CR	1	5	5	4	3	3	21	4
MINSLK	3	2	4	5	2	4	20	3
LPT	5	2	6	7	4	2	26	5

* - furnace batch size = 1

Each dispatch rule is ranked for each column in order of lowest queue time. The two columns on the right denote the sum of a dispatch rule's column rankings, and its overall ranking, based on the sum of the column rankings. The excellent performance of the SPT rule is not unexpected; and is in agreement with results in the literature noted by Blackstone et al (1982). The LOWSTEP rule shares

first place ranking with SPT. Its performance is surprising given its characteristic of ignoring jobs near completion in favour of jobs near the beginning of the process. FCFS ranked a close second to these two rules, with CR, MINSLK, and DUEDATE clustered in the middle, and LPT placing a distant last.

The utilization level of the workstation models in Sections 4.1 and 4.2 had significant effects on the relative performance of the dispatch rules. There was a general increase in the range of queue times observed for the dispatch rules as the workstation utilization level increased. The queue time range for a scenario (denoted by a column of results in the results tables) is defined as the maximum queue time less the minimum queue time in each column. Figure 4.3 is a graph of the range of queue times observed versus the corresponding utilization level of the furnace A workstation model for five sets of results. Four sets of results correspond to both furnace models run at batch sizes of one and two lots. (See Appendix B Tables B.1.7 - B.1.9, and B.1.10 - B.1.12, respectively). The fifth set of results corresponds to the stepper workstation. (See Table 4.5).

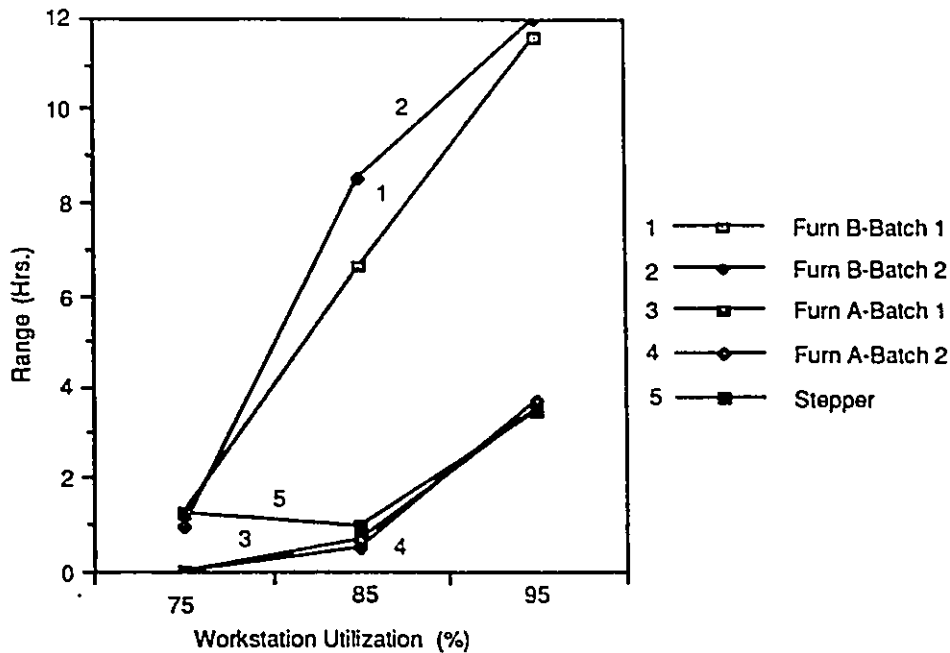


FIGURE 4.3 - QUEUE TIME RANGE

Two observations for the LPT rule, on the furnace A model, at 85% and 95% utilization levels were not included in the calculations because their values of 42.8 and 51.2 hours are out of line with the other results and would have far too significant an effect on the resulting ranges.

The steady increase in range with the utilization level in Figure 4.3 illustrates the importance of choosing the dispatch strategy carefully when the system is operating at high congestion levels.

4.4 Summary

A main issue of the isolated hub analyses was the effect of varied batch sizes on the factory objectives of capacity and cycle-time improvements. Batch sizes of one and two lots delivered competing levels of queue times and throughput levels for the 'best' dispatch rules as measured by their queue times. Batch size one gave lower queue time and batch size two gave higher throughput.

Batch sizes of three and four lots generally resulted in significant increases in queue time compared to smaller batch sizes, especially when the job arrival rates were constant. Since the increased furnace workstation capacity with batch sizes of three and four lots cannot be taken advantage of, due to the lower capacity of the stepper workstation, these two batch sizes will not be included in further experiments in subsequent chapters.

The dispatch rules produced a wide variety of results. A ranking of the rules showed SPT and LOWSTEP to have generally lower queue times, although FCFS ranked a close second. The LPT rule had a significant number of large queue times in comparison with the other rules and was ranked last. CR, MINSLK, and DUEDATE completed the middle rankings. The utilization level had significant effects on the range of queue times for each scenario. This effect, shown in Figure 4.3, illustrates the need for careful selection of dispatch rules at high levels of system congestion. The sometimes very large queue times recorded for LPT and its general low queue time ranking prompt us not to include this rule in examination in subsequent chapters.

5. HUB NETWORK

The one stepper and two furnace workstations are combined in a single network model to continue the experimentation completed on the isolated hub models. The hub network model is outlined in Chapter 3 and adheres to the modelling concepts developed for the isolated hub models presented in Chapters 3 and 4.

Two objectives are pursued in the experiments on this model. The first is to examine the effect of interaction between the hubs on the performance of the dispatch and batching rules, vis-a-vis the experience and knowledge gained from the isolated hub models. The second objective is to explore the possibility that a combination of different dispatch rules on the three hub workstations may result in improved performance compared with the application of a single dispatch rule on all hubs.

The results presented in Chapter 4 illustrated the poor queue time performance of furnace batch sizes of three and four lots, and the inability to utilize the larger furnace capacities due to the constraining capacity of the stepper workstation. As well, the LPT dispatch rule performed very poorly compared with the other rules. For these reasons batch sizes of three and four lots and the LPT dispatch rule were not included in the experimental design for the hub network model.

5.1 Hub Network Experiments

The first experiment on the hub network model tests six dispatch rules at two furnace batch sizes and four job arrival rates. The furnace batch sizes are one and two lots. Three of the job arrival rates, 16.6, 14.8, and 13.1 lots per week (l.p.w.), correspond to 95%, 85%, and 75% utilization levels of the furnace A workstation at a furnace batch size of one lot. The fourth job arrival rate of 17.9 l.p.w. is equivalent to a 95% utilization level of the stepper workstation.

This scenario corresponds to a furnace batch of two when the stepper workstation replaces the furnace A workstation as the system bottleneck. Since the arrival rate of 17.9 l.p.w. is greater than the system capacity of 17.5 l.p.w. at batch size one, no simulations are performed with that combination. The factorial experiment design is illustrated in Figure 5.1.

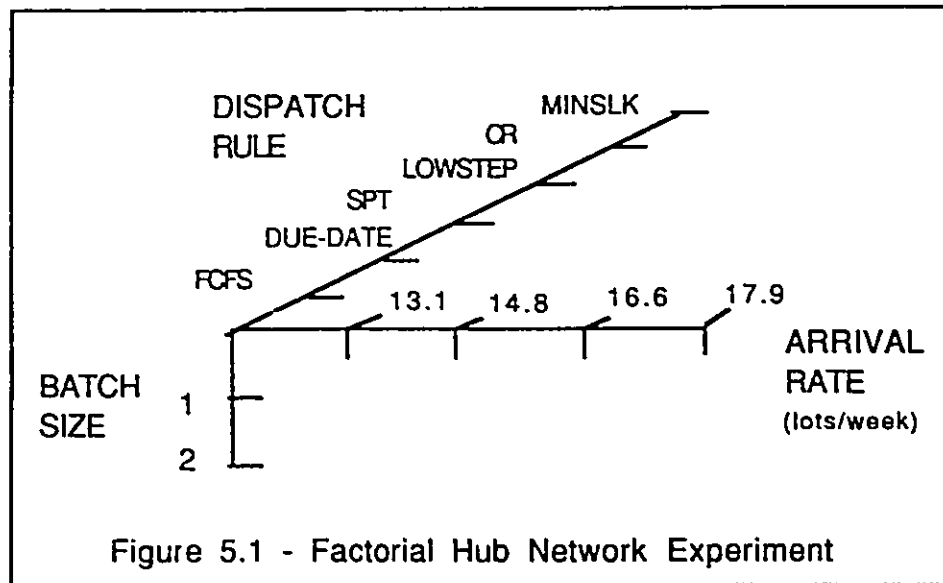


Figure 5.1 - Factorial Hub Network Experiment

Three observational issues are examined: (1) the effect of furnace batch sizes, (2) the relative performance of the dispatch rules and, (3) the effect of the system bottleneck utilization level.

5.1.1 Furnace Batch Sizes The six dispatch rules and four job arrival rates are examined at furnace batch sizes of one and two lots as illustrated in Figure 5.1. Results for the three job arrival rates of 16.6, 14.8, and 13.1 l.p.w. are listed in Table 5.1. Each element in the table consists of the mean total queue time and standard deviation (in brackets).

Table 5.1 - Hub Networks Results

Arrival Rate (l.p.w.)	16.6		14.8		13.1	
Furnace Batch Size	1	2	1	2	1	2
FCFS	21.0 (4.6)	21.5 (6.9)	12.8 (4.0)	15.0 (6.2)	9.1 (3.7)	18.5 (7.7)
DUEDATE	65.3 (9.4)	19.0 (7.0)	16.0 (4.7)	10.1 (6.0)	7.3 (3.7)	18.0 (8.8)
SPT	18.0 (4.4)	22.5 (6.4)	12.1 (3.6)	15.0 (6.2)	8.0 (3.4)	20.8 (7.1)
LOWSTEP	16.3 (5.1)	26.5 (8.5)	12.3 (4.1)	14.3 (6.7)	7.6 (3.4)	16.7 (7.4)
CR	36.5 (6.6)	22.5 (4.4)	13.5 (3.5)	28.3 (5.9)	9.6 (3.6)	20.1 (7.6)
MINSLK	25.6 (4.1)	21.3 (5.0)	13.8 (3.5)	21.8 (8.4)	8.0 (3.9)	20.1 (5.9)

- * - queue times measured in hours
- batch size 1 => furnace A is system bottleneck
- batch size 2 => stepper is system bottleneck
- sampling procedure per Section 3.2

There are eighteen changes from batch size one to batch size two recorded in Table 5.1. Fourteen of those changes resulted in an increase in queue time and four changes resulted in a decrease in queue time. The relative increases in queue time were generally larger for the lower arrival rates. These effects were also observed on the isolated hub models in Chapter 4 and may be a result of the two batch size effects described in Section 4.1.1. Interaction between the dispatch rules and batch size change caused the increases in queue time for the first fourteen changes evidently due to the queue time effect (II). Conversely, interaction between the dispatch rules and the batch size change for the four remaining batch size scenarios caused the decreases in queue time evidently due to the capacity effect (I).

In Table 5.1 larger increases in queue time are observed at lower arrival rates. The magnitude of the queue time effect (II) is dependent on the amount of furnace idle time. At batch size two lower arrival rates cause more furnace idle time since lots wait longer to complete batches and, accordingly, these arrival rates had the largest relative queue time increases.

The fourth job arrival rate of 17.9 l.p.w. is equivalent to the 95% utilization rate of the stepper workstation. Changing from an arrival rate of 16.6 l.p.w. with a batch size of one, to 17.9 l.p.w. with a batch size of two, is a 7.8% increase in system throughput as measured by the job arrival rate. These two scenarios correspond to 95% utilization levels of the system bottlenecks at furnace batch sizes of one and two lots, respectively. Results for these two scenarios are listed in Table 5.2 next to batch size one results listed previously in Table 5.1. The mean total queue time and standard deviation (in brackets) are reported.

Table 5.2 - Queue Time vs Throughput Trade-off*

Furnace Batch Size	1	2
Arrival Rate(l.p.w.)	16.6	17.9
FCFS	21.0 (4.6)	25.8 (7.6)
DUE DATE	65.3 (9.4)	29.0 (8.8)
SPT	18.0 (4.4)	27.5 (7.6)
LOWSTEP	16.3 (5.1)	26.3 (7.9)
CR	36.5 (6.6)	30.1 (7.2)
MINSLK	25.6 (4.1)	30.1 (6.5)

- * - queue times measured in hours
- batch size 1 => furnace A is system bottleneck
- batch size 2 => stepper is system bottleneck
- sampling procedure per Section 3.2

Only CR and DUE DATE cause decreases in queue time in the change to batch size two; and they also have the longest queue times at batch size one. The remaining four dispatch rules caused increases in queue time during the change to batch size two. This same type of effect appears in Table 5.1 as well and is discussed there. The potential gain in system throughput of 7.8%, noted previously, must be weighed against the potential increase in queue time, which in these results is realized on only the three hubs in the model. The trade-off between queue time and throughput observed in these results was also seen on the isolated hub models.

5.1.2 Dispatch Rule Performance The relative performance of the dispatch rules is of interest, particularly at high levels of system congestion. Table 5.3 contains rankings of the queue time results

found in Tables 5.1 and 5.2 for the six dispatch rules, two furnace batch sizes, and three job arrival rates of 14.8, 16.6, and 17.9 l.p.w. .

Table 5.3 - Dispatch Rule Comparison

Arrival Rate(i.p.w.)	16.6		14.8		17.9	Sum of Column Rankings	Overall Ranking
	1	2	1	2	2		
Furnace Batch Size	1	2	1	2	2		
FCFS	3	3	3	3	1	13	3
DUE DATE	6	1	6	1	4	18	4
SPT	2	4	1	3	3	13	2
LOWSTEP	1	5	2	2	2	12	1
CR	5	4	4	5	5	23	6
MINSLK	4	2	5	4	5	20	5

- batch size 1 => furnace A is system bottleneck
- batch size 2 => stepper is system bottleneck

The ranking for Tables 5.1 and 5.2 is based on queue time with the number one rank being given to the dispatch rule with the minimum queue time by column. (See Tables 5.1 and 5.2). The two columns on the right are the sum of the dispatch rule's column rankings and the overall rank based on the minimum sum of the column rankings.

The rankings place LOWSTEP, SPT, and FCFS as the best overall performing rules. Of interest is the complete about-face in performance of the DUE DATE rule between batch size one and batch size two. The relative performance of the six dispatch rules is consistent with that obtained in the previous chapter for the isolated hub models.

5.1.3 Utilization Level The third issue is the impact of the utilization level of the system bottleneck on the range (as defined in Section 4.3) of queue times. Range and median results for the columns of data in Table 5.1 are presented in Table 5.4. The first set of utilization levels reflect the batch size one results while the second set of utilization levels reflect the batch size two results when the furnace A workstation is replaced by the stepper workstation as the system bottleneck. The same job arrival rates are used in both tables.

Table 5.4 - Utilization Level Effect *

a) batch size one

Utilization Level (of furnace A)	95%	85%	75%
Median Queue Time	23.3	13.2	8.0
Queue Time Range	49.0	3.9	2.3

b) batch size two

Utilization Level (of the stepper)	88%	79%	70%
Median Queue Time	22.0	15.0	19.3
Queue Time Range	7.5	18.2	4.2

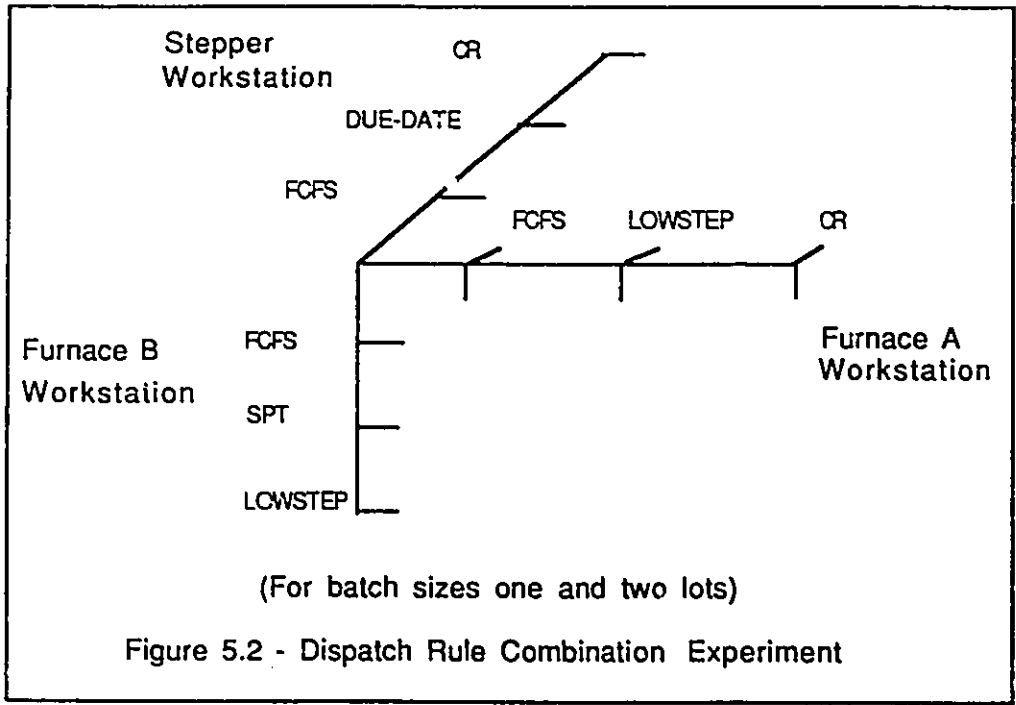
* - all medians and ranges are measured in hours

The batch size one results show a steady decline in both the range and median value of the queue times as the utilization rate of the system bottleneck drops. At batch size two the range and median do not follow a consistent pattern and may be a result of the batch size effects described in Section 5.1.1.

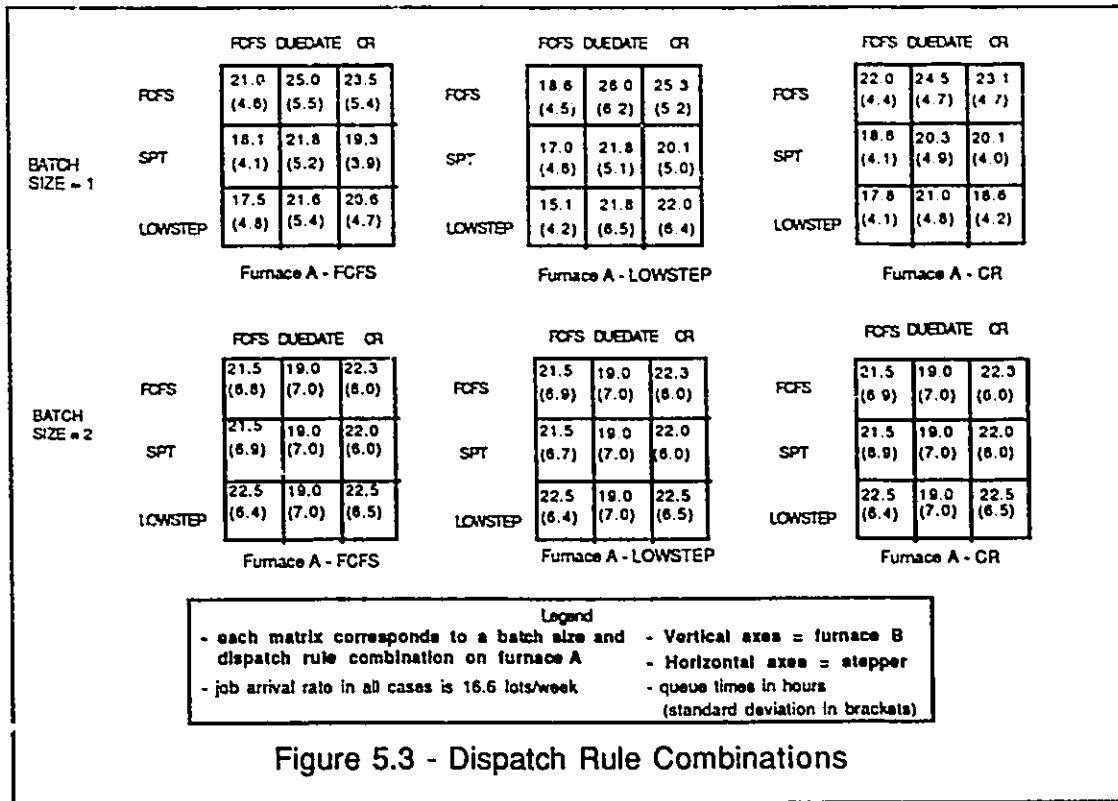
5.2 Dispatch Rule Combinations

The second experiment on the hub network model tests 27 combinations of dispatch rules on the hub workstations at furnace batch sizes of one and two lots and at a job arrival rate of 16.6 lots

per week. The selection of rule combinations was determined by the performance of the rules on the corresponding isolated hub models in Chapter 4. The objective of this experiment is to determine whether a judicious selection of dispatch rules can outperform a single rule applied at all workstations. The experiment is a 3x3x3 factorial design illustrated in Figure 5.2. The experiment examines three dispatch rules on each of the three hubs.



The results from this experiment are listed in Table 5.3 and can be compared to the 'single' rule results in the first two columns of Table 5.1.



Many of the results are comparable with the better queue times observed for the single rule scenarios listed in Table 5.1. However, as the goal was to find a superior combination of dispatch rules, the results are somewhat disappointing with only one combination improving on the 'lowest' single rule result. With batch size one the combination of LOWSTEP on both furnace workstations and FCFS on the stepper workstation outperformed the LOWSTEP result found in Table 5.1 by 1.2 hours or approximately 8 percent.

An interesting observation for the batch size one results is that, of eight queue times below nineteen hours, seven occurred when the stepper workstation operated with the FCFS rule. The batch size two results include nine queue times less than or equal to nineteen hours and all nine values occur when the stepper workstation is operated with the DUE DATE dispatch rule. This reliance on the dispatch rule used at the stepper emphasizes the importance of operating the stepper workstation efficiently. At batch size two the importance of

operating the stepper workstation efficiently is obvious since it is the system bottleneck. (See Section 5.1.1). However, DUEDATE's large effect at batch size one is less obvious as the furnace A workstation is the system bottleneck.

The identical results at batch size two for each of the stepper dispatch rules reflects the extremely low utilization levels of the furnace workstations. The stepper has a utilization of 88% versus less than 50% for the furnace workstations. Therefore, the furnaces have fewer completed batches waiting for service, and selection of a dispatch rule at the furnaces is not as critical.

5.3 Summary

The results from the hub network experiments reflected and strengthened the observations from the previous chapter. A furnace batch size of one delivered the lowest queue times for the majority of dispatch rules and utilization levels. Where individual rules showed an improvement at batch size two, these rules were always the 'poor' performing rules at batch size one. The trade-off between queue time and throughput was again confirmed and the gain in throughput at batch size two is limited to the ceiling imposed by the stepper which becomes the system bottleneck at batch size two.

As in the previous chapter, SPT, LOWSTEP, and FCFS gave the best overall performances among the dispatch rules. The relatively poor performances of CR and MINSLK are more surprising and disappointing. However, these rules are designed to take advantage of extra manufacturing information to enhance on-time delivery of lots, which may sometimes not be a compatible objective with queue time reduction. No significant improvements were observed for any combination of dispatch rules over the 'single' rule results although many competitive values were recorded. While this is disappointing, it is perhaps a blessing in disguise as the problems associated with implementing different rules on different workstations are not trivial.

In moving to the third and final level of modelling on the full process model it is desirable to focus the simulation effort on the positive and interesting aspects of the research to this point. By doing this the results can be refined and the recommendations will be easier to develop. As well, a base for exploration of further system factors such as machine failure and job arrival distribution will be easier to implement and understand.

The consistent and intuitively obvious reduction in queue time at lower workstation utilization levels, combined with greater interest in systems at high congestion levels, make the 75% utilization level less attractive to study. An obvious interest in the better performing rules warrants their inclusion in further experimentation. CR and MINSLK have been quite comparable in their overall performance so far. The dynamic nature of CR versus MINSLK makes it a more interesting subject for study and, given their use of identical information, MINSLK will be dropped from further investigation. Therefore, in the full process model five dispatch rules, FCFS, DUE DATE, SPT, LOWSTEP, and CR, furnace batch sizes of one and two lots, and job arrival rates of 14.8, 16.6, and 17.9 lots per week will be investigated.

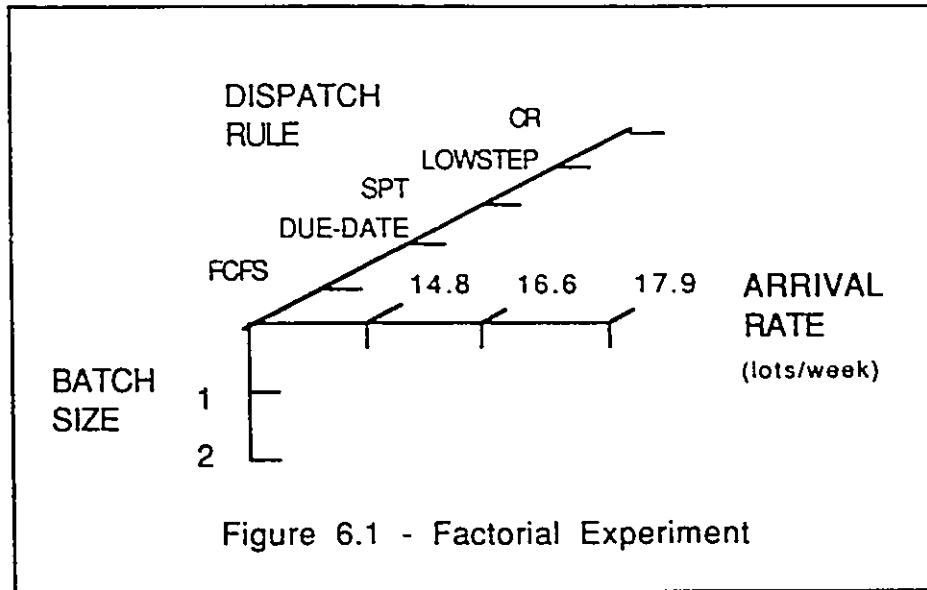
6. FULL PROCESS MODEL

The model of the full process is an extension of the hub network model of the previous chapter and includes all the off-hub steps explicitly as individual workstations. The model has a total of 180.2 hours of machine processing time on 129 processing steps which are divided between hub and off-hub steps as illustrated in Figure 3.2. Data for the full process model is tabled in Appendix A.3.

Five dispatch rules, FCFS, DUE DATE, SPT, LOWSTEP, and CR, and furnace batch sizes of one, and two lots are the main factors in three sets of experiments. The first set follows the design exercised in the previous two chapters and examines the dispatch rules and furnace batch sizes at three job arrival rates. The second and third experiments examine job arrival distributions and machine failure scenarios, respectively.

6.1 Standard Factorial Experiments

The first set of experiments tests the five dispatch rules and two furnace batch sizes at three job arrival rates. The job arrival rates correspond to 95% and 85% utilization levels of the system bottleneck at batch size one and 95% utilization level of the system bottleneck at batch size two. The system bottleneck at batch size one is the furnace A workstation. The job arrival rates corresponding to the 95% and 85% utilization levels are 16.6 and 14.8 l.p.w., respectively. The system bottleneck at batch size two is the stepper workstation and its 95% utilization level corresponds to a job arrival rate of 17.9 l.p.w. . Since at batch size one the arrival rate of 17.9 l.p.w. is greater than the system capacity of 17.5 l.p.w. no simulations are performed with that combination. The factorial experiment is illustrated in Figure 6.1.



Three main issues are examined as in the previous chapters: (1) the effect of furnace batch sizes, (2) the relative performance of the dispatch rules, and (3) the effect of the utilization level of the system bottleneck.

6.1.1 Furnace Batching Levels The results of the experiment illustrated in Figure 6.1 for job arrival rates of 16.6 and 14.8 l.p.w. are listed in Table 6.1. Each table element consists of the mean total queue time per lot and standard deviation (in brackets).

Table 6.1 - Full Process Results

Arrival Rate (l.p.w.)	16.6		14.8	
Furnace Batch Size	1	2	1	2
FCFS	33.1 (4.9)	66.3 (8.1)	22.1 (4.5)	64.6 (9.1)
DUE DATE	71.6 (9.4)	70.3 (7.2)	24.0 (5.0)	63.3 (6.9)
SPT	32.8 (5.0)	66.3 (7.6)	20.8 (4.5)	64.1 (8.5)
LOWSTEP	31.1 (6.8)	67.5 (9.9)	21.3 (4.9)	69.1 (8.5)
CR	31.3 (3.1)	71.0 (7.3)	21.1 (4.2)	63.3 (6.3)

- * - queue times measured in hours
- batch size 1 => furnace A is system bottleneck
- batch size 2 => stepper is system bottleneck
- sampling procedure per Section 3.2

Every change from batch size one to batch size two resulted in a significant increase in the queue time except for the DUE DATE results at 16.6 l.p.w. . The DUE DATE rule has a large queue time at batch size one and the increased furnace capacity available with a batch size of two lots was evidently able to reduce the queueing.

In Table 6.1 the queue times at 16.6 l.p.w. and batch size two were approximately double those at batch size one while, at 14.8 l.p.w., the queue times at batch size two were approximately treble those at batch size one. The difference observed in the magnitude of the batch size effect between job arrival rates may be explained by the queue time effect (II) first described in Section 4.1.1. The magnitude of the queue time effect (II) is dependent on the amount of furnace idle time. At batch size two lower arrival rates cause more furnace

idle time since lots wait longer to complete batches and, accordingly, these arrival rates had the largest relative queue time increases.

The queue times in Table 6.1 reflect queueing delay at each of the 57 workstations comprising 129 processing steps in the model. The queue times are for the most part significantly longer than their counterparts on the hub network model in Chapter 5 and the isolated hub models of Chapter 4. The increases are due to the increasing number of workstations represented explicitly in each successive level of modelling. Recall that in the isolated hub models only the particular hub workstation was represented explicitly with the off-hub steps being modelled by a set of delay periods. The hub network model contained only the three critical hubs with the remaining off-hub steps represented again as delay periods. The full process model being studied in this chapter has explicit representation for all workstations and as a result more complex system dynamics. Complete data models for all three sets of models are given in Appendix A.

The third job arrival rate in the experiment illustrated in Figure 6.1 is 17.9 l.p.w. and is equivalent to the 95% utilization level of the stepper workstation, which replaces the furnace A workstation as the system bottleneck at batch size two. The results obtained at this arrival rate are listed in Table 6.2 next to the batch size one and 16.6 l.p.w. results, for comparison of the two scenarios with respect to the queue time versus throughput trade-off.

Table 6.2 - Queue Time vs Throughput Trade-off

Arrival Rate (l.p.w.)	16.6	17.9
Furnace Batch Size	1	2
FCFS	33.1 (4.9)	75.0 (8.0)
DUEDATE	71.6 (9.4)	101.3 (13.4)
SPT	32.8 (5.0)	73.1 (7.5)
LOWSTEP	31.1 (6.8)	72.5 (10.2)
CR	31.3 (3.1)	101.3 (13.5)

- queue times measured in hours
- batch size 1 => furnace A is system bottleneck
- batch size 2 => stepper is system bottleneck
- sampling procedure per Section 3.2

The results in Table 6.2 must be considered with respect to the approximate 7.8% increase in system throughput as measured by the increase in the job arrival rate. The change to batch size two caused significant increases in queue time for all the dispatch rules; and the increase is larger than the increase observed when the job arrival rate was kept constant at 16.6 l.p.w. . (See Table 6.1). The significant trade-off between system throughput and queue time observed in the previous two chapters is confirmed by these results.

6.1.2 Dispatch Rule Comparison The results listed in Tables 6.1 and 6.2 for the five dispatch rules varied more across batch sizes and job arrival rates than dispatch rules. However, Table 6.3 is a ranking of the dispatch rules as they performed with respect to these results.

Table 6.3 - Dispatch Rule Comparison

Arrival Rate(l.p.w.)	16.6		14.8		17.9	Sum of Column Rankings	Overall Ranking
Furnace Batch Size	1	2	1	2	2		
FCFS	4	1	4	3	3	15	4
DUE DATE	5	3	5	1	4	18	5
SPT	3	1	1	2	2	9	1
LOWSTEP	1	2	3	4	1	11	2
CR	2	4	2	1	4	13	3

- batch size 1 => furnace A is system bottleneck
- batch size 2 => stepper is system bottleneck

The ranking of results from Tables 6.1 and 6.2 is based on queue time with the number one rank being given to the dispatch rule with minimum queue time by column. The two columns on the right are the sum of a dispatch rule's column ranking and the overall rank based on the minimum sum of the column rankings.

The ranking places SPT and LOWSTEP in the first and second spots, respectively. This is consistent with results obtained in the previous two chapters. However CR, which did not rank well in the two previous chapters, has displaced FCFS and DUE DATE to obtain the third-place ranking. The DUE DATE rule continued its earlier poor performance and ranked last.

6.1.3 Utilization Level The third issue is the effect of the utilization level of the system bottleneck on the range (defined in

Section 4.3) of queue times. Range and median results for the columns of data in Tables 6.1 and 6.2 are presented in Table 6.4. The first box contains the batch size one results from Table 6.1 in which the furnace A workstation is the system bottleneck. The second box contains the batch size two results from both tables in which the stepper workstation is the system bottleneck. The job arrival rates for the first box are 16.6 and 14.8 l.p.w. corresponding to the 95% and 85% utilization levels of the furnace A workstation, respectively. The arrival rates for the second box are 17.9, 16.6, and 14.8 l.p.w. corresponding to the 95%, 88%, and 79% utilization levels of the stepper workstation, respectively.

Table 6.4 - Utilization Level Effect

a) batch size one

Utilization Level (of furnace A)	95%	85%
Median Queue Time	32.8	21.3
Queue Time Range	40.5	3.2

b) batch size two

Utilization Level (of the stepper)	95%	88%	79%
Median Queue Time	75.0	67.5	64.1
Queue Time Range	28.8	4.7	5.8

* - all medians and ranges are measured in hours

At batch size one both the median and range of the queue times decrease after the change from 95% to 85% utilization. At batch size two the median decreases as the utilization level drops, but the range of queue times slightly increases after the 88% to 79% utilization level change. These results echo those in Table 5.4 from the hub network model. The results for batch size one are consistent in both tables, while the results for batch size two are mixed and may reflect more complex system dynamics imposed on the system by a batch size of two.

6.2 Job Arrival Distribution

Thus far, the simulation experiments in this thesis have all used constant inter-arrival times (time between subsequent job arrivals) to generate the required job arrival rate. In many systems the entry of jobs cannot be controlled tightly, so it is of interest to examine the effect of alternative arrival distributions on the service strategies studied in this thesis.

Two additional job arrival distributions were studied:

- (1) stochastic arrival of jobs
- (2) a gated release policy

The stochastic job arrival scenario models the situation of random and independent arrival of jobs. An exponential distribution is used to model the job inter-arrival times. The 'gated release policy' is representative of a typical strategy in integrated circuit manufacturing. Groups of lots are periodically released into the manufacturing line after being 'ok'd' by management. These release dates may occur once to several times a week.

These two scenarios are examined over the range of service strategies and job arrival rates indicated in the factorial experiment outlined in Figure 6.1. Hence, the results are paired and thus comparable with the results from the constant job inter-arrival time experiment. (See Tables 6.1 and 6.2).

6.2.1 Exponential Job Inter-Arrival Distribution The exponential distribution is used to model random and independent arrival of jobs into the system. The arrival rates specified in Figure 6.1 (14.8, 16.6 and 17.9 l.p.w.) dictate the means of the exponential distribution using the calculation :

$$\text{inter-arrival mean (hrs/lot)} = \frac{(168 = \# \text{ hours/wk.})}{(\text{arrival rate}(\text{lots/wk.}))}$$

Results for batch size one and arrival rates of 16.6 and 14.8 l.p.w. are presented in Table 6.5 next to the equivalent paired results from Table 6.1 (constant inter-arrival times) for comparison. Each table element consists of the mean queue time and standard deviation (in brackets).

Table 6.5 - Stochastic Job Arrival*

Arrival Rate (l.p.w.)	16.6		14.8	
Arrival Dist.	Constant	Stochastic	Constant	Stochastic
FCFS	33.1 (4.9)	208.0 (146.7)	22.1 (4.5)	98.0 (82.4)
DUE DATE	71.6 (9.4)	154.8 (97.5)	24.0 (5.0)	81.5 (58.6)
SPT	32.8 (5.0)	171.5 (117.8)	20.8 (4.5)	84.3 (69.9)
LOWSTEP	31.1 (6.8)	1108.5 (636.7)	21.3 (4.9)	399.8 (241.8)
CR	31.3 (3.1)	155.5 (99.5)	21.1 (4.2)	87.3 (61.5)

- * - queue times measured in hours
- sampling procedure per Section 3.2
- 16.6 l.p.w. = 95% theoretical utilization of furnace A workstation
- 14.8 l.p.w. = 85% theoretical utilization of furnace A workstation
- ** stochastic results recorded in an over capacity situation

Dramatically increased queue times were observed for the stochastic inter-arrival scenarios compared to the constant inter-arrival scenarios. The increase is much larger at 16.6 than at 14.8 l.p.w. and may be due to an increased sensitivity of the system to

temporary increases in inventory resulting from the stochastic entry of lots into the system. LOWSTEP is much more sensitive to the stochastic arrival distribution than the other rules and was unable to maintain the system under capacity at both arrival rates.

Table 6.6 is a ranking of the dispatch rules as they performed relative to each other in the stochastic arrival experiment outlined in Figure 6.1.

Table 6.6 - Dispatch Rule Comparison

Arrival Rate (lpw)	17.9	16.6		14.8		Sum	Rank
Batch Size	2	1	2	1	2		
FCFS	4	4	4	4	4	20	4
DUE DATE	2	1	1	1	1	6	1
SPT	3	3	3	2	3	14	3
LOWSTEP	5	5	5	5	5	25	5
CR	1	2	2	3	2	10	2

Interestingly, DUE DATE and CR which have not performed well in previous scenarios are ranked in the first and second spots, respectively. The very poor performance of LOWSTEP is illustrated by its consistent fifth place ranking.

An unusual observation in this experiment was not seen in the previous experiments. The throughput rates observed for the scenarios in Table 6.5 were in many cases between five and ten percent less than the corresponding arrival rates. Each previous experiment in this study has had differences of less than one percent over the finite period of the simulation. Over a reasonable period of time the rate of lots leaving the system (the throughput rate) must equal the rate of lots entering the system (the arrival rate) when the system is operating within its long term capacity.

A first thought was that the system was operating above its long term capacity. However, if this was the case the system inventory levels and lot cycle-times would be increasing over the duration of the simulations. This was not the case except for LOWSTEP, for which the capacity was found to be less than the start rates studied. The results for FCFS were typical of those for the other four rules. The inventory and cycle-time results for the FCFS simulation are plotted in Figures 6.2 and 6.3. The plots may be compared with the scenario detailed in Figures 3.5 and 3.6. The simulation was run with batch size one and 95% system bottleneck utilization (16.6 l.p.w. based on the mean job inter-arrival time).

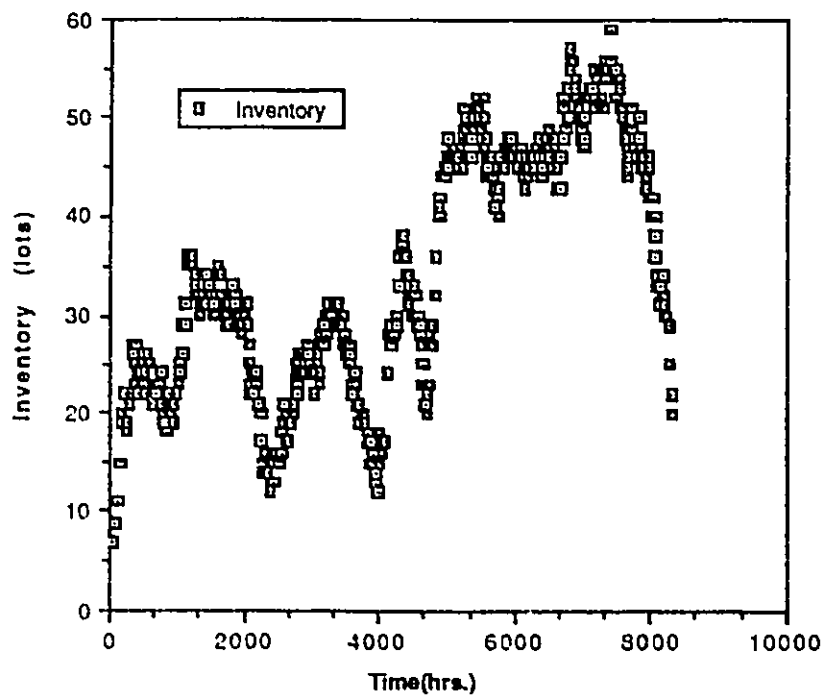


Figure 6.2 - Inventory Trace

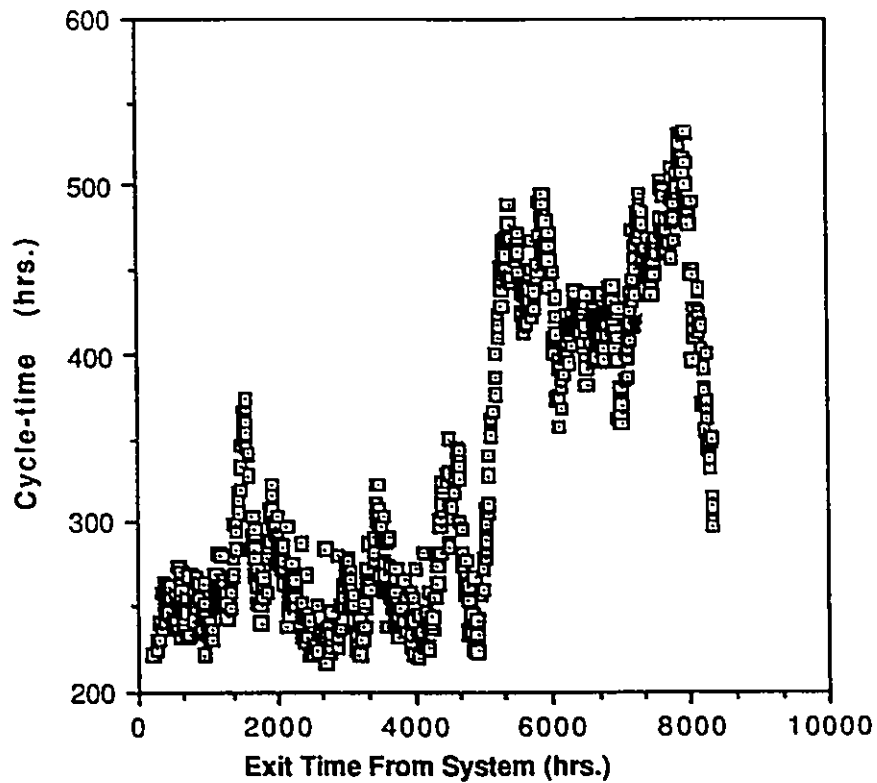


Figure 6.3 - Cycle Time Trace

In Figure 6.2 the inventory level, while cycling through many levels, is not steadily increasing after the initialization period. In Figure 6.3 the cycle-times vary but do not steadily increase as time progresses. These observations indicate that the system is not on average operating above capacity. Given the fluctuation in the inventory and cycle-time results recorded during the finite period of the simulation, the disparity between the arrival and throughput rates observed is not unexpected

6.2.2 Gated Release Policy The gated release policy applies to the job release strategy which releases lots into the system in groups as opposed to an orderly release of individual, evenly distributed lots (constant inter-arrival time). The gated release policy is intended to

model a policy sometimes employed for convenience in integrated circuit manufacturing.

The gated release policy works as follows: lots are released as in the constant inter-arrival method, but, entry to the system is controlled by a gate which opens only once every 42 hours or, equivalently, four times a week. When the gate opens all accumulated jobs proceed into the system for processing. For example, a gated job arrival rate of 20 l.p.w. would result, on average, in 5 lots being released into the system as a group every 42 hours. Note that the model uses 24 hour days and seven day weeks.

The experiment follows the factorial design given in Figure 6.1. At batch size one the furnace A workstation is the system bottleneck, and its 95% utilization level corresponds to an arrival rate of 16.6 l.p.w. or, equivalently, groups of 4 or 5 lots every 42 hours. At batch size two the stepper workstation is the system bottleneck, and the 95% utilization level corresponds to an arrival rate of 17.9 l.p.w. or, equivalently, groups of 4 or 5 lots every 42 hours. Results for 16.6 and 17.9 l.p.w. at batch sizes of one and two lots, respectively, are listed in Table 6.6. The complete set of results to the experiment is tabled in Appendix B.3.3. Results from Table 6.2 for the constant job inter-arrival experiment are listed for comparison. The table elements are mean queue times and standard deviations (in brackets).

Table 6.7 - Gated Job Arrival*

Arrival Rate (l.p.w.)	16.6		17.9	
Arrival Dist.	Constant	Gated	Constant	Gated
FCFS	33.1 (4.9)	81.6 (5.8)	75.0 (8.0)	116.1 (7.9)
DUE DATE	71.6 (9.4)	95.6 (8.3)	101.3 (13.4)	134.6 (10.0)
SPT	32.8 (5.0)	75.5 (16.7)	73.1 (7.5)	115.0 (18.0)
LOWSTEP	31.1 (6.8)	85.1 (18.3)	72.5 (10.2)	126.0 (21.0)
CR	31.3 (3.1)	78.5 (6.7)	101.3 (13.5)	139.3 (14.0)

- * - queue times measured in hours
- all scenarios at 95% system bottleneck utilization
- sampling procedure per Section 3.2

In Table 6.7 the queue times observed for the gated release policy are significantly larger than those for the constant inter-arrival scenarios. The large increases in the queue time are due to two effects. The first effect is the waiting time incurred by lots as they wait for management approval (the gate), and the second effect is the queue time impact of lots reaching the first bottleneck in groups. The large effect these groups may have is easy to imagine especially since the system bottlenecks in this study are the hub workstations used at many steps throughout the process.

Table 6.8 is a ranking of the dispatch rules with respect to their results from the gated release policy experiment outlined in Figure 6.1. Each column contains the column ranking of the dispatch rule where the number one rank is given to the rule with the shortest

queue time. The two columns on the right are the sum of a rules column ranks and its overall ranking based on the sum, respectively.

Table 6.8 - Dispatch Rule Comparison*

Arrival Rate(l.p.w.)	16.6		14.8		17.9	Sum of Column Rankings	Overall Ranking
Furnace Batch Size	1	2	1	2	2		
FCFS	3	3	4	3	2	15	3
DUE DATE	5	4	3	4	4	20	4
SPT	1	1	2	1	1	6	1
LOWSTEP	4	5	5	5	3	22	5
CR	2	2	1	2	5	12	2

- * - batch size 1 => furnace A is system bottleneck
- batch size 2 => stepper is system bottleneck

Two observations are worth noting. SPT ranked first overall by a significant margin, and LOWSTEP which has performed well in the experiments to Section 6.1, ranked last. CR ranked second overall as it did in the stochastic inter-arrival rate experiment.

The results from both the gated arrival and stochastic arrival scenarios indicate that how jobs enter the system can have significant effects on queue time. Although many flow shops may have to respond directly to customer orders and have a reduced ability to regulate the introduction of new jobs into the system, the size of the possible impact suggests an effort should be made to standardize job release policies.

6.3 Machine Failure Analysis

Unplanned machine failure can have very disrupting effects on the flow of product in an IC factory. When a machine fails at a system hub product flow can be affected throughout the entire factory. The machine failure experiment examines the five dispatch rules, two batching levels, and three job arrival rates as illustrated in the factorial design in Figure 6.1.

The three hub machines are subjected to failure scenarios, described by constant failure intervals and repair times provided by Northern Telecom. The failure interval and repair times of the furnaces are 336 hours and 24 hours, respectively. The failure interval and repair times of the stepper machines are 365 hours and 5 hours, respectively. The furnace down-time due to failure is approximately 7% while the stepper down-time is only 1.3%.

Results are listed in Table 6.8 for 95% system bottleneck utilization and batch sizes of one and two lots, along with equivalent results for the no-machine-failure scenario listed previously in Table 6.2. Complete results to the experiment outlined in Figure 6.1 are tabled in Appendix B.3.4.

Table 6.9 - Machine Failure

Arrival Rate (l.p.w.)	16.6		17.9	
Machine Dist.	No Failure	** Failure	No Failure	Failure
FCFS	33.1 (4.9)	220.0 (62.4)	75.0 (8.0)	80.1 (8.9)
DUE DATE	71.6 (9.4)	246.0 (35.0)	101.3 (13.4)	114.0 (14.2)
SPT	32.8 (5.0)	195.0 (53.7)	73.1 (7.5)	77.1 (8.2)
LOWSTEP	31.1 (6.8)	1141.6 (388.7)	72.5 (10.2)	77.3 (10.3)
CR	31.3 (3.1)	203.6 (36.6)	101.3 (13.5)	114.0 (10.9)

- * - queue times measured in hours
- arrival rate of 16.6 lpw => batch size one
- arrival rate of 17.9 lpw => batch size two
- sampling procedure per Section 3.2
- all scenarios at 95% system bottleneck utilization
- ** - observed in over capacity situation

The machine failure results at 16.6 l.p.w. (batch size one) were observed with the system operating over capacity. The arrival rate of 16.6 l.p.w. is equivalent to a 95% utilization level of the system bottleneck, the furnace A workstation. But, with the 7% furnace down-time the system was operating over capacity.

Of particular interest is the very poor performance of the LOWSTEP rule in the over capacity situation, since in previous scenarios it has performed quite well. When the system is over capacity there will be a steady increase in inventory and a steady supply of lots at early process steps at the furnace A hub. In the LOWSTEP scenario this ensures that lots near the end of the process

rarely get serviced at the hub and the queue times increase dramatically.

The queue time results for the machine failure scenario are large compared to those for the no-machine-failure case. The queue time increases at batch size two (17.9 l.p.w.) are much smaller than those at batch size one (16.6 l.p.w.), which is reasonable since the system is operating under capacity at batch size two. The increase in queue time at batch size two is due almost entirely to failure of the stepper machines since the furnaces are operating at less than 50% utilization. The increase in queue time at batch size one is due to the over capacity situation caused by the failure of the furnaces. The growing queue sizes in the over capacity situation cause the significant increases seen in queue time.

Table 6.10 is a ranking of the performance of the dispatch rules for the machine failure experiment. The results at 16.6 l.p.w. and batch size one observed in the over capacity situation are included in the ranking as the results are an indication of the ability of the dispatch rules in this situation.

Table 6.10 - Dispatch Rule Comparison

Arrival Rate (lpw)	17.9	16.6		14.8		Sum	Rank
Batch Size	2	1	2	1	2		
FCFS	3	3	3	2	4	15	2
DUE DATE	4	4	5	5	1	19	5
SPT	1	1	2	1	1	6	1
LOWSTEP	2	5	1	4	5	17	4
CR	4	2	4	3	3	16	3

SPT continued its outstanding performance recording the lowest queue times in four of five machine failure scenarios. FCFS, CR, and LOWSTEP placed in tight second, third, and fourth spots, respectively. DUE DATE again did quite poorly.

6.4 Summary

Many scenarios are examined in the simulations presented in this chapter. The first set of experiments strengthened many observations recorded in previous chapters. A furnace batch size of two lots in most instances resulted in higher queue times than batch size one. However, the above capacity situation in the machine failure experiment illustrated the ability of batch size two to handle a higher volume of jobs.

The results are very favourable for SPT, which continued the superior performance displayed in the earlier chapters. However, LOWSTEP faltered badly in the near-and-just-over-capacity situations found in the arrival distribution and machine failure experiments. CR which had not performed all that well on the isolated hub and hub network models was significantly improved on the full process model. The promise of the CR rule may lie such complex situations.

Both the stochastic and gated arrival scenarios caused significantly increased queue times with respect to the constant inter-arrival scenario. The machine failure experiment illustrated the large impact that just three critical machine types may have on the queue time.

7. SUMMARY OF RESULTS

Many simulation experiments have been performed on three levels of increasingly complex models: isolated workstation hubs, the hub network, and the full process model. Examination of the results from these models focused on three issues:

- 1) - the relative performance of the dispatch rules on hub machines.
- 2) - the effect of furnace batch sizes on queue time and system throughput.
- 3) - the effect of the utilization level of the system bottleneck.

In this chapter these issues are discussed as they relate to the three levels of models examined previously.

7.1 Furnace Batch Sizes

The effect of furnace batch size was looked at in two different contexts: (1) the comparison of system performance at different batch sizes with the job arrival rate maintained at a constant level, (2) the ability to increase system throughput using the larger furnace capacity provided by higher batch sizes.

7.1.1 Constant Arrival Rates The large majority of scenarios on the three levels of models caused significant increases in queue time when the batch size was increased. These increases were evidently due to the extra queue time incurred at the higher batch size by lots waiting to complete their batch. However, there were instances on all three model levels where the increase in batch size caused a decrease in queue time. These instances always had the longest queue times at the lower batch size and the increased furnace capacity of the higher batch size was able to reduce that queueing.

The significant differences in the queue times observed between batch size one and batch size two, and the changing conditions under which either batch size will yield lower queue times, illustrates the importance of understanding the nuances of the particular system under study.

7.1.2 Queue-time vs Throughput Trade-off In the models studied changing the furnace batch size from one to two lots doubled the capacity of the furnace workstations, and the stepper workstation replaced the furnace A workstation as the system bottleneck. Accordingly, the gain in system capacity was from 17.5 to 18.8 l.p.w. for an increase of approximately 7.5%.

The models were examined at 95% system utilization levels for batch sizes of one and two lots. The change to batch size two for the large majority of dispatch rules caused significant increases in queue time and the increases were larger on the more complex models. The larger effect is expected on more complex models since more workstations are explicitly included in these models and more queueing and jostling for service takes place.

The increase in queue time observed for batch size two must be considered in conjunction with the 7.8% increase in system throughput from 16.6 to 17.9 l.p.w., as measured by the job arrival rate. A clear example of the capacity and throughput increase at batch size two was illustrated in the machine failure experiment. At batch size one, the arrival rate of 16.6 l.p.w. was above the system capacity for all the dispatch rule scenarios, while at batch size two, the larger arrival rate of 17.9 l.p.w. was under the system capacity for all the dispatch rule scenarios. The trade-off between queue time and throughput which must be addressed in the batch size decision is of consequence, since in many environments throughput, capacity, and cycle-time improvement are major but partially conflicting objectives. Poulsen and Sutcliffe (1987) study the cycle-time versus throughput compromise with respect to the furnace batch size decision in IC manufacturing.

7.2 Comparison of Dispatch Rules

The method used for ranking the performance of the dispatch rules in this thesis was very simple. In each simulation scenario the dispatch rules were ranked in order of lowest queue time first. In each experiment the dispatch rules were ranked by summing the within scenario ranks and ordering the rules by lowest sum first. The method was an attempt to measure the general performance of the dispatch rules across many experiments. The consistency of the rankings from model to model demonstrates the effectiveness of this simple scheme.

SPT and LOWSTEP were consistently the top performers across all three levels of models. However, on the full process model LOWSTEP caused some very poor results in the stochastic arrival distribution and machine failure experiments. Since LOWSTEP gives priority to lots at the earliest process steps the large queues associated with high arrival rates at hub workstations will result in less time to service those lots at later steps in the process. The time available for processing lots in the last stages of the process will diminish in inverse proportion to the size of the workstation queue. Thus lowering the effective capacity of the system in these situations, and explaining LOWSTEP's extreme susceptibility to disruptions and irregularities so common in 'real' systems.

In the majority of experiments SPT ranked first with respect to queue time performance. The poor performance of LOWSTEP in several scenarios on the full process model clearly leaves SPT as the premiere rule in this study. However, in practise SPT may leave longer jobs idle for substantial periods of time when there is a steady supply of shorter jobs. A practical implementation of this rule would incorporate a clause to move to the front of the queue any lot waiting longer than a predetermined length of time. Additionally, SPT is only applicable to workstations with varying process times, so

forcing adoption of alternative policies at workstations with uniform processing times.

Queue times observed for CR in many but not all cases were competitive with those of SPT. CR performed best in the job arrival and machine failure experiments on the full process model. Certainly in this study it must be ranked lower than SPT. But, studies in this thesis used a single product type, a single process sequence, and a simple due date selection methodology, whereas, the full potential of CR may lie in more complex situations. Additionally, the CR rule places emphasis on achieving 'on time' delivery of product with respect to the job due date, a goal that is sometimes not compatible with a low cycle-time goal.

The performance of the DUEDATE rule was not very promising, and since the lot due date is included in the definition of the CR rule DUEDATE is not of much interest for further study. The performance of the FCFS rule was very consistent and received average to good rankings in the majority of experiments. Practical use of the FCFS rule is based on a combination of its simplicity of implementation and its average performance. However, SPT is equally simple to implement and can deliver much lower queue times in cases with varying process times.

Every factory or system has its own unique characteristics and objectives. The implementation of a particular queueing methodology will depend on these factors. In many circumstances a combination of dispatch rules and global system information may have to be applied to best achieve system objectives. The understanding gained in this thesis of how to characterize the effectiveness of dispatch rules will be very useful in developing solutions for these complex situations.

7.3 System Bottleneck Utilization Level

The effect of the utilization level of the system bottleneck on the queue times observed for the dispatch rules was measured by the median and range of queue times for each model scenario.

Queue times observed at batch size one on the three model levels had consistently higher medians and larger ranges for higher utilization levels of the system bottleneck. This trend is expected since higher workstation utilization levels result in more lots queuing for the same machine resources and, consequently, spending more time waiting for service. The dispatch rules then have more effect on the arrangement of the queues and, hence, on the performance of the system.

The results at batch size two were less consistent than those at batch size one. In general, higher utilization levels did result in higher medians and larger ranges of queue times but there were some inconsistent results. The additional dynamics imposed on the system by a batch size of two, combined with the small number of points, make conclusions at this batch size more difficult.

7.4 Extensions to the Full Process Model

Two job arrival distributions and a machine failure scenario were examined on the full process model to demonstrate the effect of system disturbances on the service strategies.

The stochastic lot arrival distribution had the largest effect on the system. Queue time means and standard deviations increased significantly, and in many cases the queue time means were five times larger than those for the constant arrival case. These results are consistent with Wein (1988) who examined a number of job arrival policies. Stochastic arrival distributions cause fluctuations of inventory levels in the system, as was illustrated in the corresponding experiment. Even though the same utilization level of

bottleneck workstations are maintained over a long period of time, the reduced performance observed may have been due to short term fluctuations in inventory levels at the workstations. At low levels of bottleneck congestion the bottleneck machines may stand idle resulting in lost throughput, while at high levels of bottleneck congestion lots spend additional time queueing for service resulting in increased cycle times.

The gated release policy caused similar, although less extreme, increases in both the mean and standard deviation of queue times. A gated arrival policy in an integrated circuit facility will tend to cluster lots together until they arrive at the first bottleneck in the system. There, the additional waiting time will affect product flow in the entire system since the main bottlenecks tend to be multiple revisit workstations. Therefore a constant release rate for lots is recommended in IC facilities to reduce possible queueing effects.

The machine failure experiment involved removing machines at the hub workstations from service according to failure interval and repair times provided by Northern Telecom. The 95% utilization level scenario of the system bottleneck resulted in the system operating in an over capacity situation. The furnaces in the bottleneck workstation were unavailable 7% of the time due to simulated failure, which clearly placed the workstation in a long term over capacity situation. The dramatically increased queue times that were recorded reflect the over capacity congestion levels in the system. In real systems there are many stochastic elements which make it difficult to determine when a system is in an over capacity situation. The magnitude of the effect emphasizes the need to be able to detect and predict over capacity situations so appropriate preventative and remedial action can be taken.

At batch size two and 17.9 l.p.w. the machine failure scenario caused an increase in queue times compared with the no-machine-failure case. The increases reflect the 1.3% 'down-time' average of the machines in the stepper workstation, which is the system

bottleneck at batch size two. The important point is that the increases in queue time observed are due to the failure of only eight of 57 machines in the full process model. Thus emphasizing the importance of maintaining high 'up-time' averages, especially, on hub machines.

bottleneck at batch size two. The important point is that the increases in queue time observed are due to the failure of only eight of 57 machines in the full process model. Thus emphasizing the importance of maintaining high 'up-time' averages, especially, on hub machines.

8. CONCLUSIONS & EXTENSIONS

8.1 Conclusions

8.1.1 Discrete Simulation The discrete simulation models developed in this thesis permitted examination of many scenarios involving complex queueing systems that would be difficult to study with analytic methods. The modularity of the general simulation system at Northern Telecom facilitated development of new models and workstation service strategies in an effective manner. With it the researcher can 'follow his/her nose' without a large expenditure of time and effort re-developing the model for each new experiment or set of simulations. Combining a well constructed simulation system in conjunction with the benefits of discrete simulation provides an efficient, powerful, and knowledgeable tool to study manufacturing operations.

8.1.2 Research Methodology The research methodology involved examining a set of dispatch rules and furnace batch sizes on a set of models of an integrated circuit facility. The models ranged in complexity from single workstation models to a network model of the critical hub workstations and, finally, a model of a full wafer fabrication process. The understanding and experience gained from the single workstation models was useful in focusing and shaping experimentation on the hub network model for further evaluation and determination of the promising service strategies that warranted examination on the full process model.

Two ideas confirmed in the study are introduced in support of the methodology:

(1) Trends and observations may reveal themselves more clearly on small, rather than more complex models. A basic understanding of the dynamics of service strategies on small models can aid understanding and identification of patterns and trends in more complex models where they may be more subtle and difficult

to identify. While complex models may provide more information, unless an expectation of results built upon experience is available, the interpretation of the results can be difficult and can lead to erroneous conclusions.

(2) The second idea is related to the first and pertains to the amount of intuition and experience that is involved in identifying major factors or issues of study and the subsequent design of effective and efficient experiments. If experimentation on scaled versions of a large system model can confirm the intuition and ideas of the researcher, then a refinement of the experiment can shorten the time and effort required to examine the application on the larger more realistic model and the probability of the experiments being successful is greatly improved. Certainly good judgement must be used in the scaling down process so that elements of the model essential to the problem are retained. The isolated hub models required approximately 2.5 minutes of computer time per simulation on a dedicated Hewlett Packard 9000 Unix system. The simulations on the full process model, on the other hand, each required approximately 30 minutes of computer time. Even the full process model is scaled down with respect to the 'real' factory with many complexities such as multiple processes and products, labour shifts, and preventative maintenance not being included. An expanded model incorporating these features, and possibly higher product volumes and stochastic distributions, sometimes requires several hours per simulation. The cost in terms of computer and people time to obtain feedback can easily become prohibitive if experiments are attempted blindly. The point is that preliminary simulation on scaled down models to refine the experimental design and develop intuition can easily pay off in terms of many days of computer and people time, not to mention the increased understanding and more accurate interpretation of results that will arise.

8.1.3 Experimental Results Seven dispatch rules and four furnace batch sizes were studied to gain a basic understanding of their performance on the hub elements of a IC fabrication line. These

hubs by definition are characterized by high utilization levels and multiple revisits during the fabrication process.

The results indicated the effectiveness of smaller batch sizes of one and two lots at delivering optimal factory performance, both in terms of cycle-time and system throughput. However, a trade-off between cycle-time and throughput clearly existed in this study.

The SPT dispatch rule was superior throughout this study and confirmed the previous results of many researchers. (See Dar-El and Wysk (1982)). The CR dispatch rule recorded average results throughout the study, but its good performance in the job arrival and machine failure experiments combined with its dynamic nature and use of additional manufacturing information suggest that its promise may not lie in the streamlined models used in this study but in more complex models and systems. An alternative definition of the CR rule includes in its denominator an estimate of total-queue-time-remaining in addition to the total-work-remaining and may provide enhanced performance characteristics. Queue time as it relates to cycle-time was used as the major performance measure in this thesis. 'On time' delivery of product as measured by the product due dates is another key performance measure in the Northern Telecom facility. The focus of CR on due date observance gives it a distinct advantage in this area. SPT while not emphasizing the due date has very good performance in this area (See Baker (1984)), but its applicability to IC environments may be limited as many machines have uniform processing times.

8.2 Extensions

This thesis examined a combination of dispatch rules and furnace batch sizes to gain a basic understanding of their performance on manufacturing systems characterized by hubs. There are many dispatch rules documented by researchers which involve combinations and variations of the rules examined here. (See Blackstone (1982)). As well, a number of heuristic rules are

available which take advantage of the timeliness and amount of the information available on the factory floor as a result of the expansion of information systems in the last several years. (See Panwalker (1977)). Certainly SPT and CR are starting points for further investigation but any study of dispatch rules should consider the individual factory and methods particularly suited to it. This thesis considered four furnace batch sizes applied uniformly across both furnace workstation hubs. Poulsen and Sutcliffe (1987) show that a mixture of batch sizes among furnace workstations and even among machines in particular workstations can yield more optimal performance mixes for a given IC facility.

A single process sequence and product type were assumed in the models in this thesis. In fact, the 'real' factory operates several processes sharing the same machine resources with several product types per process. This has many implications with respect to machine operation and, in particular, the compatibility issues of lots in furnace batches and setup sequences on the stepper machines. Further studies should include such additions to the process model.

The machine failure experiment on the full process model involved only the three critical hub workstations and used failure rate estimates provided by Northern Telecom. Further studies should include use of historical failure data to provide a more accurate model of the real factory, and analysis of the factory model to determine which machines the system is most sensitive to with respect to their failure characteristics. Two related issues are examination of recovery strategies from major machine failures which can result in a significant build-up of inventory and, similarly, determination of methods to recognize and recover from over capacity situations.

There are many complex machines in an IC fabrication facility. Personnel coverage of all the machines is prohibitively expensive and effective strategies need to be determined which deliver the best compromise of cost and operator availability that will satisfy

factory performance objectives. There are several options available, including: (1) multi-skilled people able to operate several types of machines, (2) effective scheduling of people around product requirements for movement and processing, and (3) the staffing level. Study of such strategies is recommended.

APPENDIX A - SIMULATION DATA MODELS

This appendix contains the data used in the flow shop models in this thesis. Each model except for the full process model is described with both a flow diagram and a data table. The full process model is described by a data table only.

A.1 - Isolated Hub Models

A.1.1 - Furnace A Workstation Model

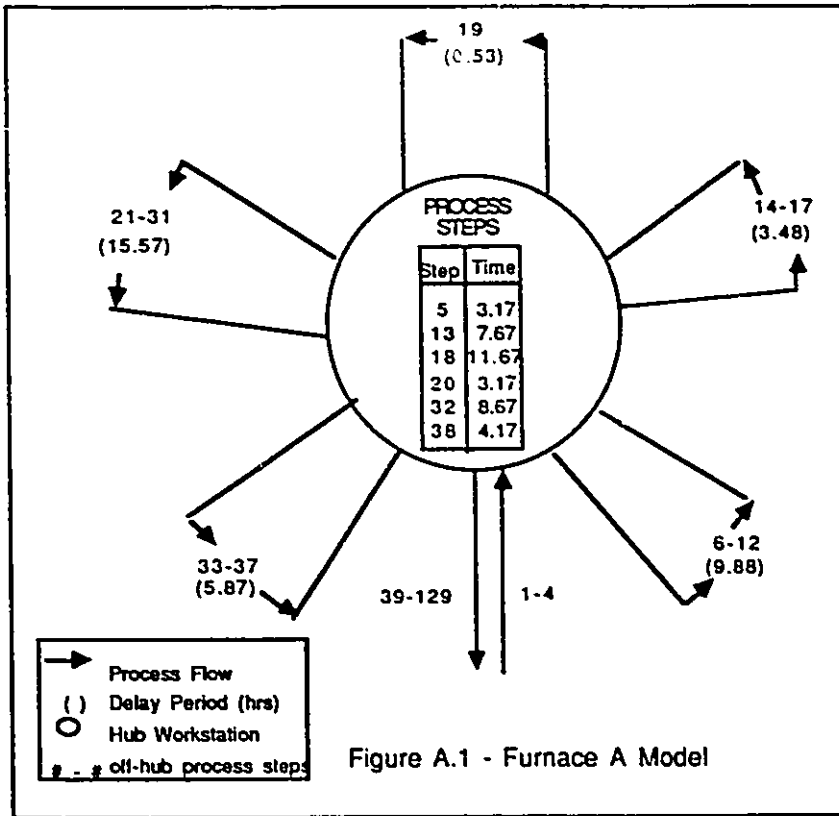


Table A.1 - Furnace A Data Model

Step	Node Type	Processing Time (hrs.)	Delay Period (hrs.)
1	furnace	3.17	
2	delay		9.88
3	furnace	7.67	
4	delay		3.48
5	furnace	11.67	
6	delay		0.53
7	furnace	3.17	
8	delay		15.57
9	furnace	8.67	
10	delay		5.87
11	furnace	4.17	

A.1.2 - Furnace B Model

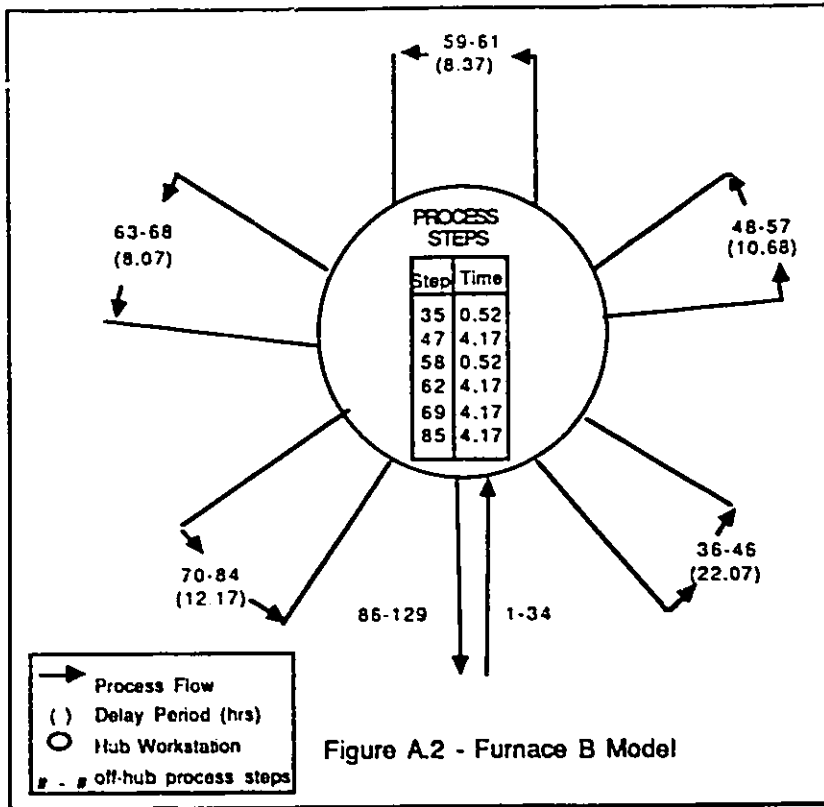


Figure A.2 - Furnace B Model

Table A.2 - Workstation B Data Model

Step	Node Type	Processing Time (hrs.)	Delay Period (hrs.)
1	furnace	0.52	
2	delay		22.07
3	furnace	4.17	
4	delay		10.68
5	furnace	0.52	
6	delay		8.37
7	furnace	4.17	
8	delay		8.07
9	furnace	4.17	
10	delay		12.17
11	furnace	4.17	

A.1.3 - Stepper Model

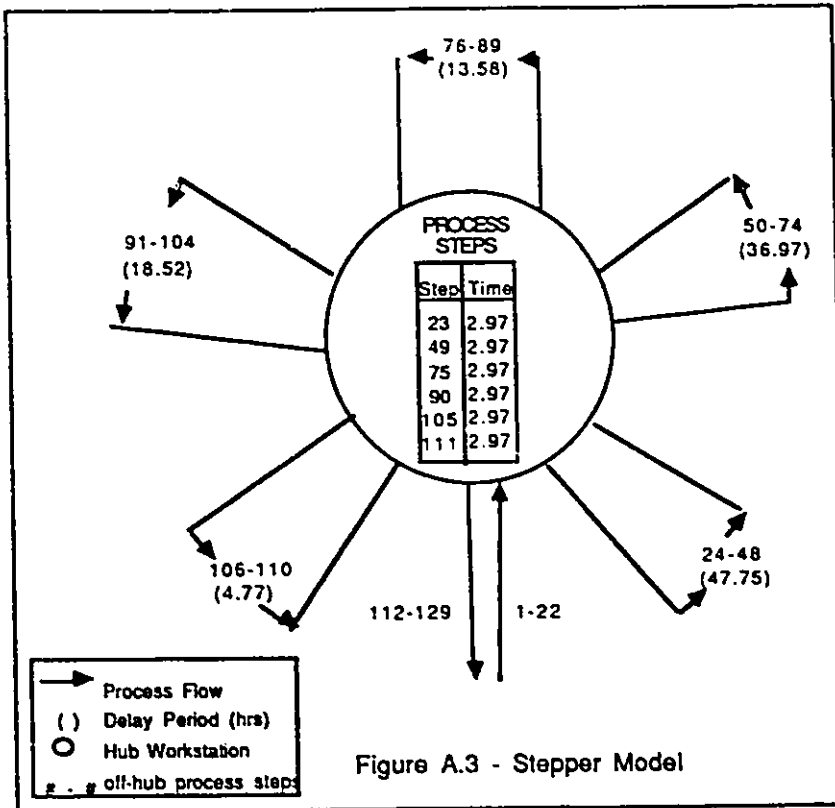


Table A.3 - Stepper Data Model

Step	Node Type	Processing Time (hrs.)	Delay Period (hrs.)
1	stepper	2.97	
2	delay		47.75
3	stepper	2.97	
4	delay		36.97
5	stepper	2.97	
6	delay		13.58
7	stepper	2.97	
8	delay		18.52
9	stepper	2.97	
10	delay		4.77
11	stepper	2.97	

A.2 - Hub Network Model

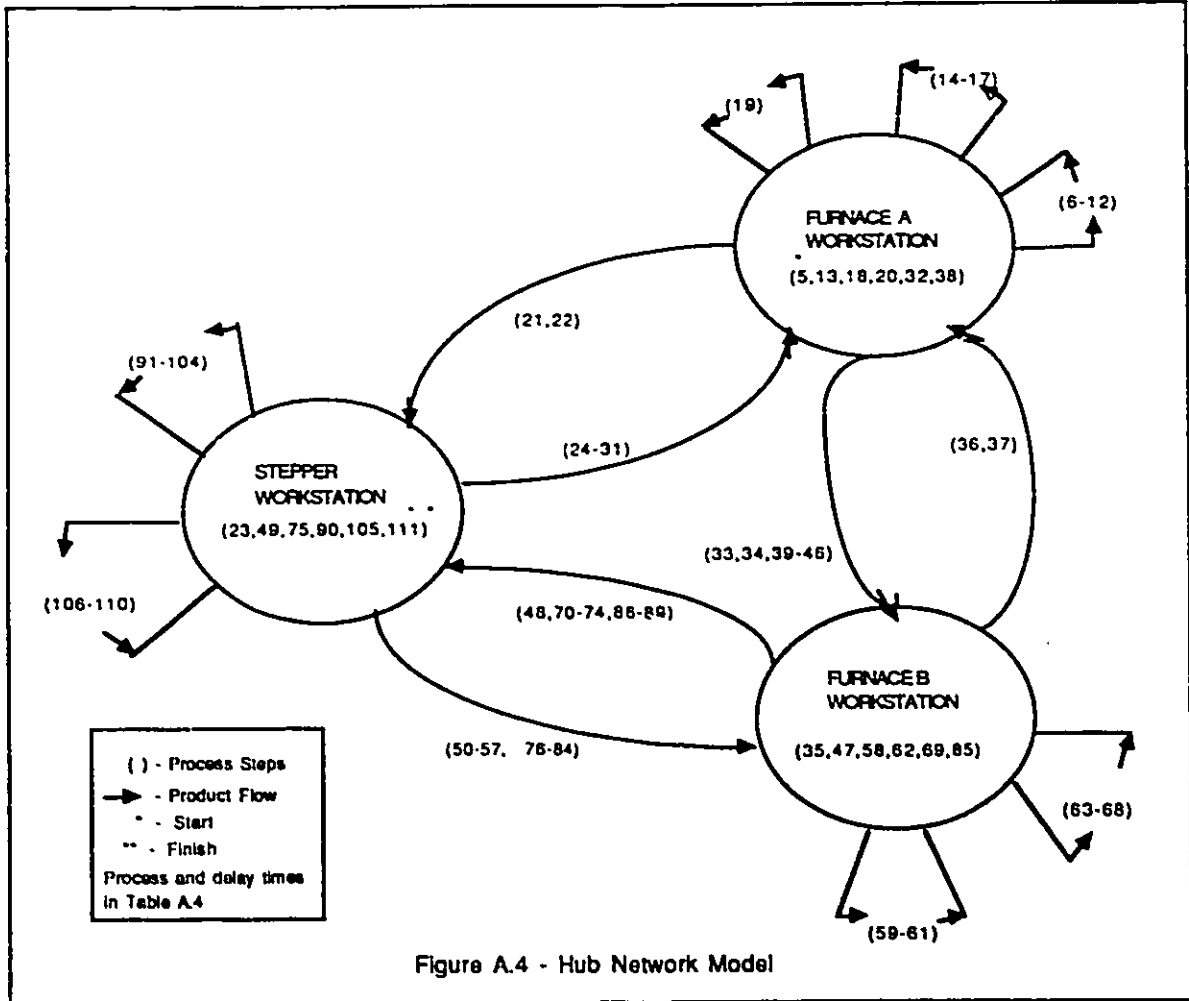


Figure A.4 - Hub Network Model

Table A.4 - Hub Network Data Model

Step	Node Type	Processing Time (hrs.)	Delay Period (hrs.)
1	furnace A	3.17	
2	delay		9.87
3	furnace A	7.67	
4	delay		3.47
5	furnace A	11.67	
6	delay		30.53
7	furnace A	3.17	
8	delay		6.25
9	stepper	2.97	
10	delay		5.23
11	furnace A	8.67	
12	delay		2.73
13	furnace B	0.52	
14	delay		0.95
15	furnace A	4.17	
16	delay		15.62
17	furnace B	4.17	
18	delay		0.27
19	stepper	2.97	
20	delay		6.17
21	furnace B	0.52	
22	delay		8.37
23	furnace B	4.17	
24	delay		8.05
25	furnace B	4.17	
26	delay		2.05
27	stepper	2.97	
28	delay		6.08
29	furnace B	4.17	
30	delay		2.32
31	stepper	2.97	
32	delay		18.5
33	stepper	2.97	
34	delay		4.75
35	stepper	2.97	

A.3 - Full process Model

Process Step #	Time (hrs.)	Machine Type	Process Step #	Time (hrs.)	Machine Type
1	0.18	1	38	4.17	furnace A
2	0.77	2	39	0.52	17
3	0.37	3	40	0.37	18
4	0.73	4	41	6.37	19
5	3.17	furnace A	42	0.33	9
6	4.37	5	43	0.37	11
7	0.25	6	44	4.47	20
8	0.53	7	45	1.62	21
9	1.08	8	46	0.37	11
10	0.33	9	47	4.17	furnace B
11	0.87	10	48	0.25	6
12	0.37	11	49	2.97	stepper
13	7.67	furnace A	50	0.55	22
14	2.07	12	51	1.58	23
15	0.37	13	52	0.87	16
16	0.33	9	53	0.55	24
17	0.37	11	54	0.25	15
18	11.67	furnace A	55	0.90	25
19	0.53	14	56	0.87	16
20	3.17	furnace A	57	0.37	13
21	4.37	5	58	0.52	furnace B
22	0.25	6	59	4.37	26
23	2.97	stepper	60	3.12	21
24	0.25	15	61	0.37	11
25	1.25	8	62	4.17	furnace B
26	0.87	16	63	0.25	6
27	0.25	15	64	0.37	7
28	0.53	7	65	1.12	21
29	0.33	9	66	0.87	10
30	0.87	10	67	0.37	11
31	0.37	11	68	4.17	27
32	8.67	furnace A	69	4.17	furnace B
33	0.52	17	70	0.25	6
34	2.07	12	71	0.57	28
35	0.52	furnace B	72	0.37	16
36	0.52	17	73	0.37	13
37	0.37	13	74	0.25	6

Process Step #	Time (hrs.)	Machine Type	Process Step #	Time (hrs.)	Machine Type
75	2.97	stepper	114	0.37	30
76	0.55	22	115	0.33	35
77	0.82	29	116	0.65	24
78	0.87	30	117	3.42	40
79	0.65	24	118	2.25	41
80	0.37	13	119	0.25	42
81	0.43	31	120	0.37	7
82	1.07	29	121	0.37	28
83	0.37	30	122	0.50	30
84	0.53	13	123	0.08	44
85	4.17	furnace B	124	0.37	30
86	0.37	28	125	0.25	45
87	0.92	32	126	0.83	46
88	0.38	33	127	0.33	47
89	0.53	11	128	0.50	30
90	2.97	stepper	129	0.08	49
91	0.55	22			
92	1.57	34			
93	0.87	30			
94	0.33	35			
95	0.65	24			
96	3.75	36			
97	0.25	15			
98	4.10	37			
99	0.87	30			
100	0.37	28			
101	0.52	38			
102	0.65	24			
103	2.25	36			
104	0.25	15			
105	2.97	stepper			
106	0.55	22			
107	1.57	39			
108	0.87	30			
109	0.65	24			
110	1.00	60			
111	2.97	stepper			
112	0.55	22			
113	1.57	34			

APPENDIX B - EXPERIMENTAL RESULTS

This appendix contains a complete listing of all the experimental results. Each set of results is preceded by an outline of the experimental design. The format of the results tables and the design outlines are the same as in the main body of the thesis. All table elements consist of the mean total queue time per lot and the associated standard deviation (in brackets).

B.1 - Isolated Hub Experiments

B.1.1- Isolated Furnace A Experiment - Constant Arrival Rates

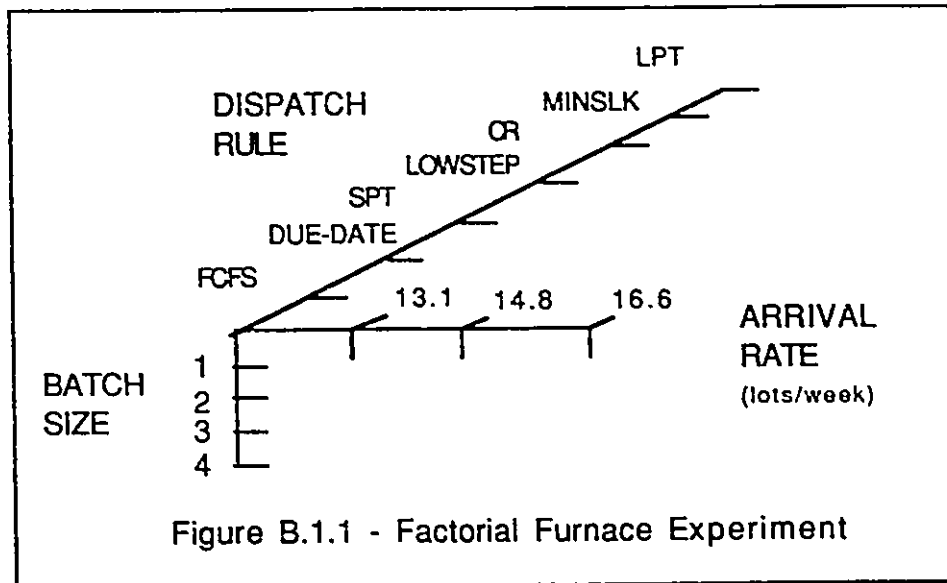


Table B.1.1 - Furnace A Model - 16.6 l.p.w.

	Furnace Batch Size (in lots)			
	1	2	3	4
FCFS	2.53 (0.7)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)
DUE-DATE	5.82 (2.3)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)
SPT	2.53 (0.7)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)
LOWSTEP	2.28 (0.8)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)
CR	2.28 (0.8)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)
MINSLK	5.47 (2.2)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)
LPT	51.17 (14.5)	5.05 (5.1)	10.13 (8.3)	15.17 (11.3)

* - queue times measured in hours
 - sampling procedure per Section 3.2

Table B.1.2 - Furnace A Model - 14.8 l.p.w.

	Furnace Batch Size (in lots)			
	1	2	3	4
FCFS	3.28 (2.5)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
DUE DATE	3.10 (2.6)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
SPT	3.20 (2.5)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
LOWSTEP	3.77 (1.8)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
CR	3.77 (1.8)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
MINSLK	3.13 (2.5)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
LPT	3.13 (2.1)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)

* - queue times measured in hours
 - sampling procedure per Section 3.2

Table B.1.3 - Furnace A Model - 13.1 l.p.w.

	Furnace Batch Size (in lots)			
	1	2	3	4
FCFS	0.00 (0.0)	6.42 (6.4)	12.80 (10.5)	19.25 (14.4)
DUE DATE	0.00 (0.0)	6.42 (6.4)	12.80 (10.5)	19.25 (14.4)
SPT	0.00 (0.0)	6.42 (6.4)	12.80 (10.5)	19.25 (14.4)
LOWSTEP	0.00 (0.0)	6.42 (6.4)	12.80 (10.5)	19.25 (14.4)
CR	0.00 (0.0)	6.42 (6.4)	12.80 (10.5)	19.25 (14.4)
MINSLK	0.00 (0.0)	6.42 (6.4)	12.80 (10.5)	19.25 (14.4)
LPT	0.00 (0.0)	6.42 (6.4)	12.80 (10.5)	19.25 (14.4)

* - queue times measured in hours
 - sampling procedure per Section 3.2

B.1.2- Isolated Furnace B Experiment - Constant Arrival Rates

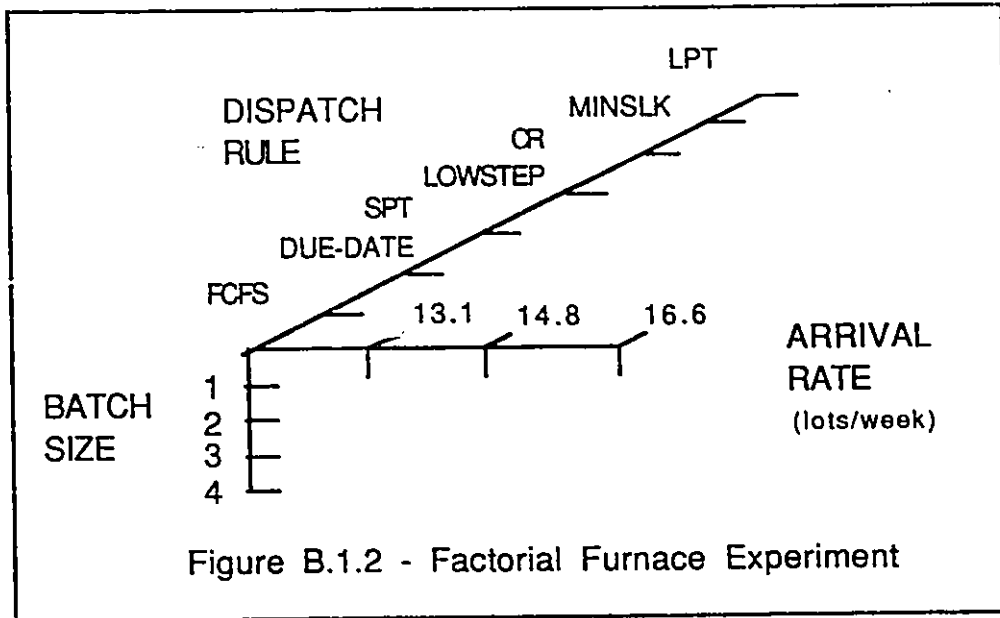


Figure B.1.2 - Factorial Furnace Experiment

Table B.1.4 - Furnace B Model - 16.6 l.p.w.

	Furnace Batch Size (in lots)			
	1	2	3	4
FCFS	8.98 (4.1)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
DUE DATE	7.52 (4.6)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
SPT	5.98 (2.1)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
LOWSTEP	5.93 (3.8)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
CR	3.57 (4.4)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
MINSLK	2.78 (3.1)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)
LPT	5.55 (4.8)	5.07 (5.1)	10.13 (8.3)	15.18 (11.3)

* - queue times measured in hours
 - sampling procedure per Section 3.2

Table B.1.5 - Furnace B Model - 14.8 l.p.w.

	Furnace Batch Size (in lots)			
	1	2	3	4
FCFS	3.37 (1.1)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
DUE DATE	9.47 (4.8)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
SPT	3.50 (1.2)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
LOWSTEP	3.25 (1.2)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
CR	3.73 (1.2)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
MINSLK	8.52 (3.6)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)
LPT	9.88 (4.0)	5.65 (5.7)	11.30 (9.2)	16.98 (12.7)

* - queue times measured in hours
 - sampling procedure per Section 3.2

Table B.1.6 - Furnace B Model - 13.1 l.p.w.

	Furnace Batch Size (in lots)			
	1	2	3	4
FCFS	7.28 (3.8)	9.22 (6.8)	12.80 (10.5)	19.25 (14.4)
DUE DATE	6.10 (3.4)	9.22 (6.8)	12.80 (10.5)	19.25 (14.4)
SPT	6.03 (2.7)	9.00 (6.8)	12.80 (10.5)	19.25 (14.4)
LOWSTEP	6.82 (1.9)	9.00 (6.8)	12.80 (10.5)	19.25 (14.4)
CR	6.10 (2.9)	9.22 (6.8)	12.80 (10.5)	19.25 (14.4)
MINSLK	7.03 (2.0)	9.22 (6.8)	12.80 (10.5)	19.25 (14.4)
LPT	6.18 (3.3)	9.22 (6.8)	12.80 (10.5)	19.25 (14.4)

* - queue times measured in hours
 - sampling procedure per Section 3.2

B.1.3- Isolated Furnace A Experiment - Constant Utilization Levels

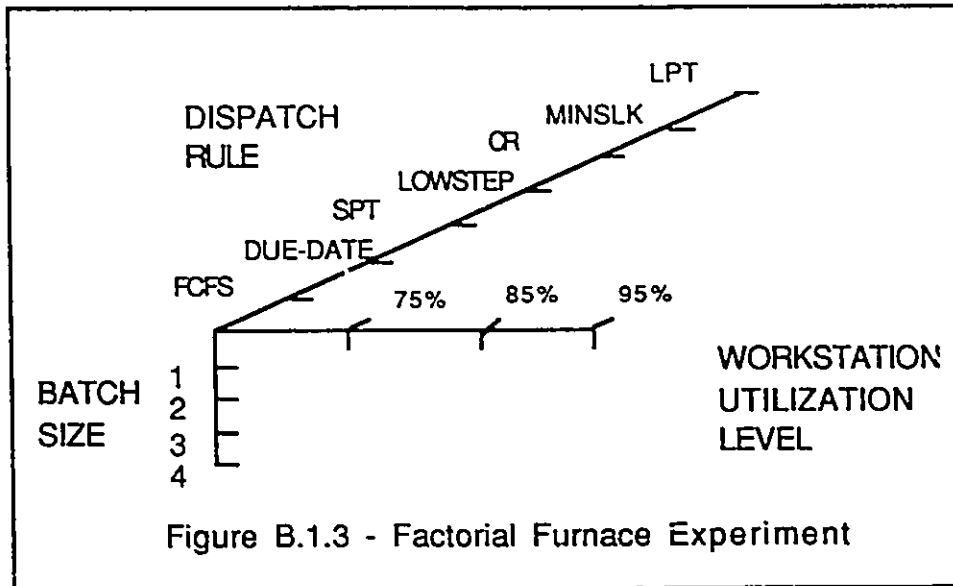


Table B.1.7 - Furnace A Model - 95%

	Furnace Batch Size (in lots)			
	1	2	3	4
Arrival Rate	16.6	33.2	49.8	66.4
FCFS	2.53 (0.7)	5.03 (2.6)	5.83 (2.8)	6.32 (2.9)
DUE-DATE	5.82 (2.3)	8.65 (3.2)	9.75 (3.9)	9.92 (3.4)
SPT	2.53 (0.7)	5.03 (2.6)	5.83 (2.8)	6.32 (2.9)
LOWSTEP	2.28 (0.8)	4.82 (2.7)	5.73 (2.8)	6.07 (3.0)
CR	2.28 (0.8)	5.20 (2.7)	5.90 (2.8)	7.20 (3.2)
MINSLK	5.47 (2.2)	5.03 (2.6)	5.88 (2.8)	6.28 (2.9)
LPT	51.17 (14.5)	42.87 (12.9)	40.43 (11.3)	37.05 (11.7)

- * - queue times measured in hours
- arrival rate = 95% utilization level of furnace A workstation.
- sampling procedure per Section 3.2

Table B.1.8 - Furnace A Model - 85%

	Furnace Batch Size (in lots)			
	1	2	3	4
Arrival Rate	14.8	29.6	44.4	59.2
FCFS	3.28 (2.5)	6.07 (3.8)	7.02 (4.0)	7.40 (4.0)
DUE DATE	3.10 (2.6)	5.97 (3.9)	6.92 (4.1)	7.27 (4.1)
SPT	3.20 (2.5)	5.98 (3.8)	6.92 (4.0)	7.25 (4.0)
LOWSTEP	3.77 (1.8)	6.42 (3.4)	7.37 (3.6)	7.60 (3.7)
CR	3.77 (1.8)	5.92 (3.7)	6.85 (4.0)	7.10 (3.9)
MINSLK	3.13 (2.5)	5.92 (3.8)	6.85 (4.0)	7.10 (3.9)
LPT	3.13 (2.1)	5.92 (3.5)	6.85 (3.8)	7.52 (3.9)

* - queue times measured in hours
 - arrival rate = 85% utilization level of furnace A workstation.

Table B.1.9 - Furnace A Model - 75%

	Furnace Batch Size (in lots)			
	1	2	3	4
Arrival Rate	13.1	26.2	39.3	52.4
FCFS	0.00 (0.0)	3.20 (3.2)	4.30 (3.5)	4.83 (3.6)
DUE DATE	0.00 (0.0)	3.20 (3.2)	4.30 (3.5)	4.83 (3.6)
SPT	0.00 (0.0)	3.20 (3.2)	4.30 (3.5)	4.83 (3.6)
LOWSTEP	0.00 (0.0)	3.20 (3.2)	4.30 (3.5)	4.83 (3.6)
CR	0.00 (0.0)	3.20 (3.2)	4.30 (3.5)	4.83 (3.6)
MINSLK	0.00 (0.0)	3.20 (3.2)	4.30 (3.5)	4.83 (3.6)
LPT	0.00 (0.0)	3.20 (3.2)	4.30 (3.5)	4.83 (3.6)

* - queue times measured in hours
 - arrival rate = 75% utilization level of furnace A workstation.

B.1.4- Isolated Furnace B Experiment - Constant Utilization Levels

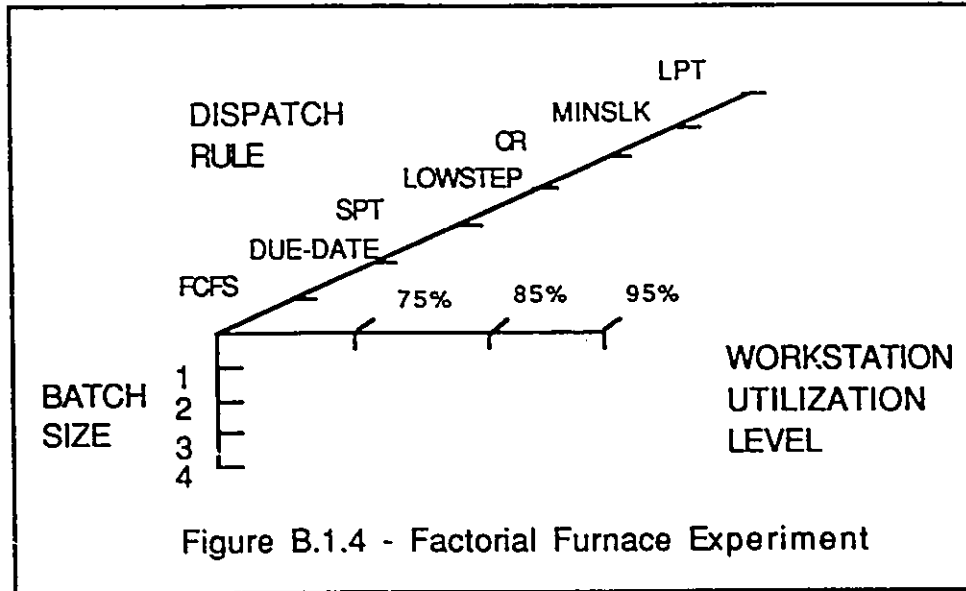


Table B.1.10 - Furnace B Model - 95%

	Furnace Batch Size (in lots)			
	1	2	3	4
Arrival Rate	16.6	33.2	49.8	66.4
FCFS	8.98 (4.1)	10.32 (4.2)	9.77 (3.7)	11.75 (4.5)
DUE-DATE	17.52 (4.6)	20.50 (6.2)	21.97 (6.5)	22.85 (6.4)
SPT	5.98 (2.1)	8.52 (3.3)	9.45 (3.6)	10.05 (3.4)
LOWSTEP	5.93 (3.8)	8.98 (3.8)	9.95 (4.5)	10.25 (4.0)
CR	13.57 (4.4)	15.98 (5.4)	18.25 (5.4)	18.42 (6.3)
MINSLK	12.78 (3.1)	15.02 (4.3)	16.27 (4.5)	16.33 (4.2)
LPT	15.55 (4.8)	17.03 (5.4)	19.45 (5.4)	18.35 (5.2)

- * - queue times measured in hours
- arrival rate = 95% utilization level of furnace A workstation.
- sampling procedure per Section 3.2

Table B.1.11 - Furnace B Model - 85%

	Furnace Batch Size (in lots)			
	1	2	3	4
Arrival Rate	14.8	29.6	44.4	59.2
FCFS	3.37 (1.1)	6.13 (3.0)	7.08 (3.3)	11.83 (4.4)
DUE DATE	9.47 (4.8)	10.75 (4.8)	11.70 (5.0)	7.62 (3.3)
SPT	3.50 (1.2)	6.13 (3.0)	7.08 (3.3)	7.22 (3.2)
LOWSTEP	3.25 (1.2)	6.02 (3.1)	6.97 (3.3)	7.22 (3.2)
CR	3.73 (1.2)	14.55 (5.3)	13.83 (5.8)	7.62 (3.3)
MINSLK	8.52 (3.6)	11.98 (4.6)	12.88 (4.8)	7.47 (3.3)
LPT	9.88 (4.0)	12.37 (5.1)	13.35 (5.3)	7.47 (3.3)

* - queue times measured in hours
 - arrival rate = 85% utilization level of furnace A workstation.

Table B.1.12 - Furnace B Model - 75%

	Furnace Batch Size (in lots)			
	1	2	3	4
Arrival Rate	13.1	26.2	39.3	52.4
FCFS	7.28 (3.8)	10.32 (4.3)	11.07 (4.5)	11.83 (4.4)
DUE DATE	6.10 (3.4)	9.33 (4.7)	10.32 (4.9)	11.02 (4.9)
SPT	6.03 (2.7)	9.82 (4.0)	11.12 (4.3)	11.93 (4.3)
LOWSTEP	6.82 (1.9)	10.03 (3.8)	11.00 (4.0)	10.95 (4.7)
CR	6.10 (2.9)	9.33 (4.7)	10.32 (4.9)	11.02 (4.9)
MINSLK	7.03 (2.0)	10.23 (3.8)	11.33 (4.0)	11.90 (4.1)
LPT	6.18 (3.3)	9.50 (4.6)	10.53 (4.9)	10.62 (4.8)

* - queue times measured in hours
 - arrival rate = 75% utilization level of furnace A workstation.

B.1.5- Stepper as a System Bottleneck

Two sets of simulations were performed at an arrival rate of 17.9 lots per week at batch size two on both isolated furnace workstation models. The experiment is described in Section 4.1.3.

Table B.1.13 - Stepper as the System Bottleneck

	Furnace A	Furnace B
FCFS	4.68 (4.7)	5.83 (4.9)
DUE DATE	4.68 (4.7)	5.83 (4.9)
SPT	4.68 (4.7)	5.83 (4.9)
LOWSTEP	4.68 (4.7)	5.83 (4.9)
CR	4.68 (4.7)	5.83 (4.9)
MINSLK	4.68 (4.7)	5.83 (4.9)
LPT	4.68 (4.7)	5.83 (4.9)

- * - queue times measured in hours
- furnace batch size is two
- arrival rate is 17.9 lots per week

B.1.6- Stepper Workstation Experiments

Twenty-eight simulations were performed on the isolated stepper workstation model. The simulations examined four job arrival rates and seven dispatch rules in a matrix design. The experiment is described in Section 4.2.

Table B.1.14 - Stepper Workstation

	Arrival Rate (l.p.w.)			
	17.9	16.6	14.8	13.1
FCFS	7.17 (2.2)	9.12 (2.1)	2.93 (1.3)	5.18 (2.0)
DUE DATE	113.10 (5.4)	6.87 (2.8)	2.53 (1.8)	4.03 (2.1)
SPT	6.98 (2.1)	9.12 (2.1)	2.93 (1.3)	5.17 (2.0)
LOWSTEP	6.55 (1.3)	10.35 (2.7)	3.05 (1.3)	4.60 (2.8)
CR	6.45 (1.3)	8.03 (2.7)	3.05 (1.3)	5.17 (2.0)
MINSLK	7.95 (1.9)	7.53 (2.6)	3.50 (2.0)	3.98 (2.2)
LPT	6.93 (2.1)	9.12 (2.1)	2.93 (1.3)	5.17 (2.0)

* - queuc times measured in hours

B.2 - Hub Network Model

B.2.1 - Factorial Network Experiment

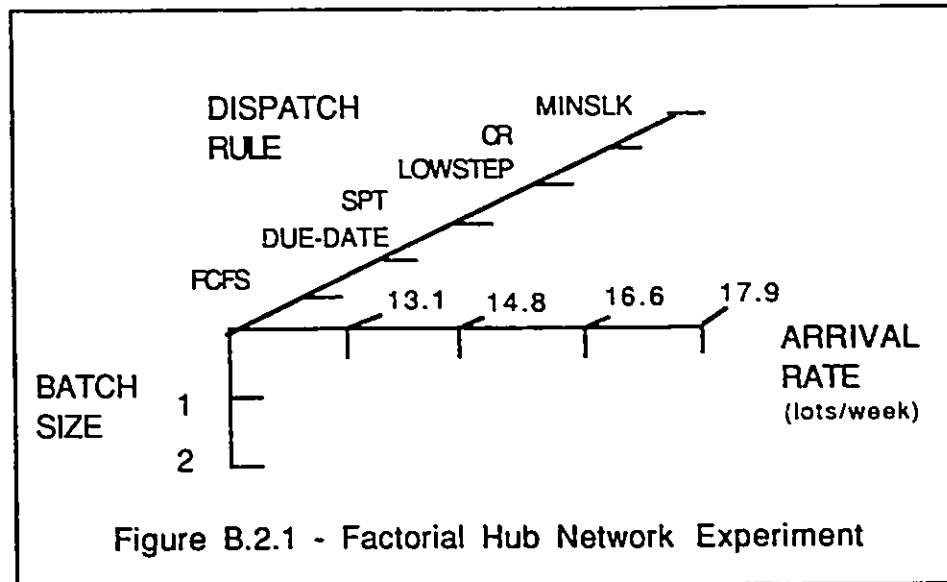
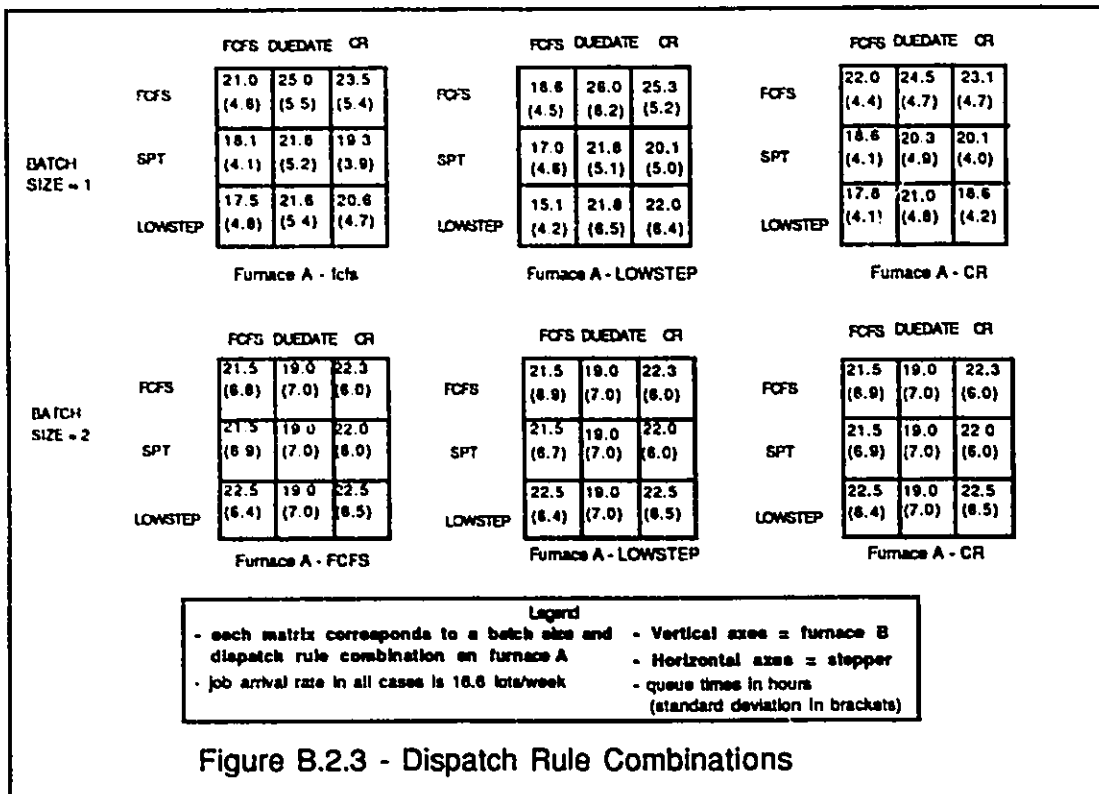
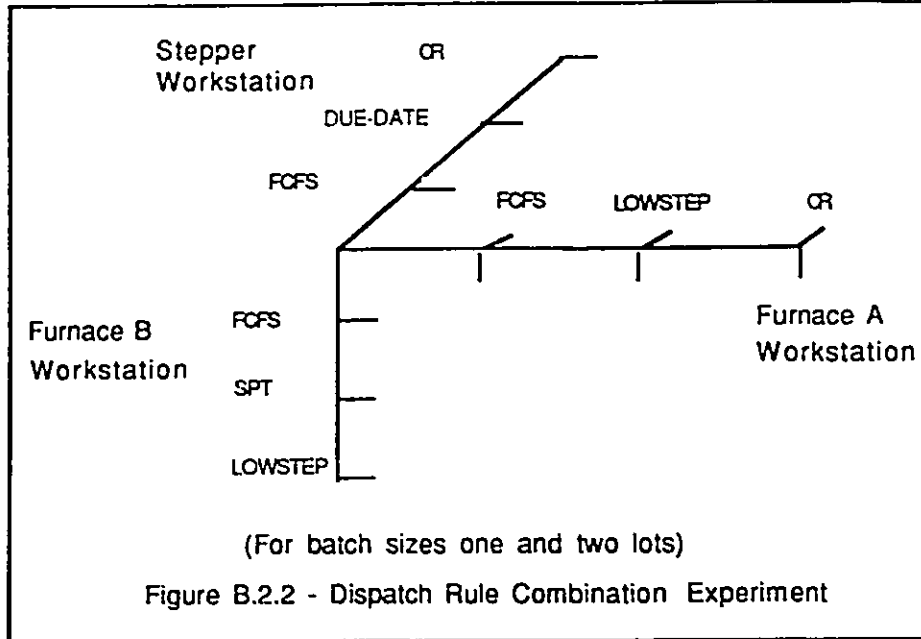


Table B.2.1 - Factorial Hub Network Results

Arrival Rate (lpw)	17.9	16.6		14.8		13.1	
Batch Size	2	1	2	1	2	1	2
FCFS	25.8 (7.6)	21.0 (4.6)	21.5 (6.9)	12.8 (4.0)	15.0 (6.2)	9.1 (3.7)	18.5 (7.7)
DUE-DATE	29.0 (8.8)	65.3 (9.4)	19.0 (7.0)	16.0 (4.7)	10.1 (6.0)	7.3 (3.7)	18.0 (8.8)
SPT	27.5 (7.6)	18.0 (4.4)	22.5 (6.4)	12.1 (3.6)	15.0 (6.2)	8.0 (3.4)	20.8 (7.1)
LOWSTEP	26.3 (7.9)	16.3 (5.1)	26.5 (8.5)	12.3 (4.1)	14.3 (6.7)	7.6 (3.4)	16.6 (7.4)
CR	30.1 (7.2)	36.5 (6.6)	22.5 (4.4)	13.5 (3.5)	28.3 (5.9)	9.6 (3.6)	20.1 (7.6)
MINSLK	30.1 (6.5)	25.6 (4.1)	21.3 (5.0)	13.8 (3.5)	21.8 (8.4)	8.0 (3.9)	20.1 (5.9)

* - all queue times measured in hours

B.2.2 - Dispatch Rule Combinations



B.3 - Full Process Model

B.3.1 - Standard Factorial Experiment

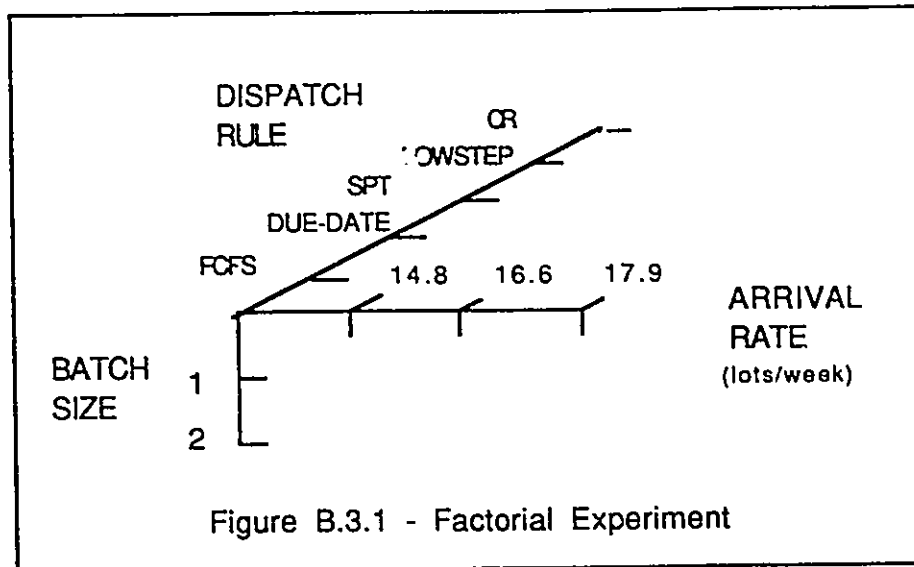


Table B.3.1 - Standard Factorial Experiment - Results

Arrival Rate (lpw)	17.9		16.6		14.8	
Batch Size	2	1	2	1	2	
FCFS	75.0 (8.0)	33.1 (4.9)	66.3 (8.1)	22.1 (4.5)	64.6 (9.1)	
DUE DATE	101.3 (13.4)	71.6 (9.4)	70.3 (7.2)	24.0 (5.0)	63.3 (6.9)	
SPT	73.1 (7.5)	32.8 (5.0)	66.3 (7.6)	20.8 (4.5)	64.1 (8.5)	
LOWSTEP	72.5 (10.2)	31.1 (6.8)	67.5 (9.9)	21.3 (4.9)	69.1 (8.5)	
CR	101.3 (13.5)	31.3 (3.1)	71.0 (7.3)	21.1 (4.2)	63.3 (6.3)	

* - queue times measured in hours

B.3.2 - Stochastic Job Inter-Arrival Experiment

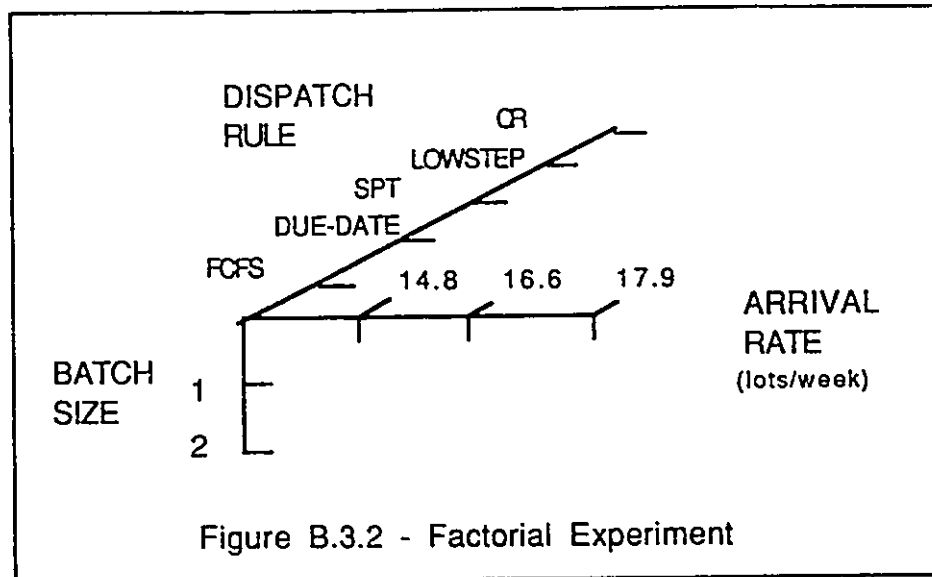


Table B.3.2 - Stochastic Job Inter-Arrival - Results

Arrival Rate (lpw)	17.9		16.6		14.8	
Batch Size	2	1	2	1	2	
FCFS	204.3 (110.2)	208.0 (146.7)	155.3 (84.3)	98.0 (82.4)	100.3 (55.4)	
DUE DATE	188.8 (97.9)	154.8 (97.5)	135.8 (59.3)	81.5 (58.6)	95.3 (39.8)	
SPT	203.6 (108.4)	171.5 (117.8)	152.3 (83.0)	84.3 (69.9)	98.5 (52.6)	
LOWSTEP	340.3 (181.5)	1108.5 (636.7)	437.8 (371.0)	399.8 (241.8)	198.6 (210.8)	
CR	177.3 (86.9)	155.5 (99.5)	140.8 (63.4)	87.3 (61.5)	95.6 (40.9)	

* - queue times measured in hours

B.3.3 - Gated Arrival Policy Experiment

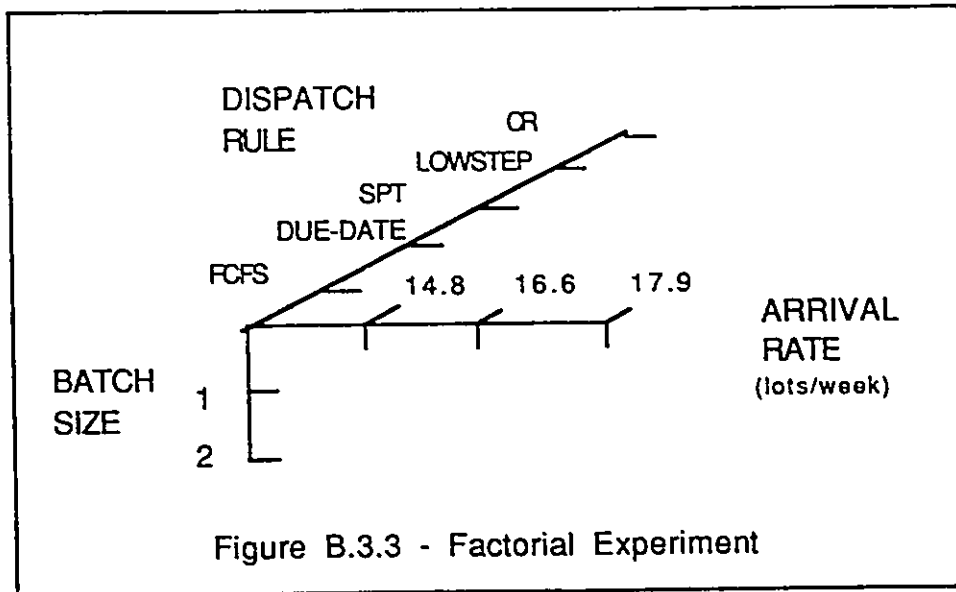


Table B.3.3 - Gated Arrival Policy - Results

Arrival Rate (lpw)	17.9		16.6		14.8	
	2	1	2	1	2	
FCFS	116.1 (7.9)	81.6 (5.8)	104.5 (10.4)	62.5 (6.6)	95.5 (10.9)	
DUE-DATE	134.6 (10.0)	95.6 (8.3)	110.0 (9.4)	61.6 (7.2)	95.8 (10.1)	
SPT	115.0 (18.0)	75.5 (16.7)	103.0 (18.1)	61.1 (16.1)	94.6 (17.9)	
LOWSTEP	126.0 (21.0)	85.1 (18.3)	111.0 (20.1)	65.8 (15.4)	102.1 (20.8)	
CR	139.3 (14.0)	78.5 (6.7)	104.3 (8.5)	59.5 (5.8)	94.8 (8.7)	

* - queue times measured in hours

B.3.4 - Machine Failure Experiment

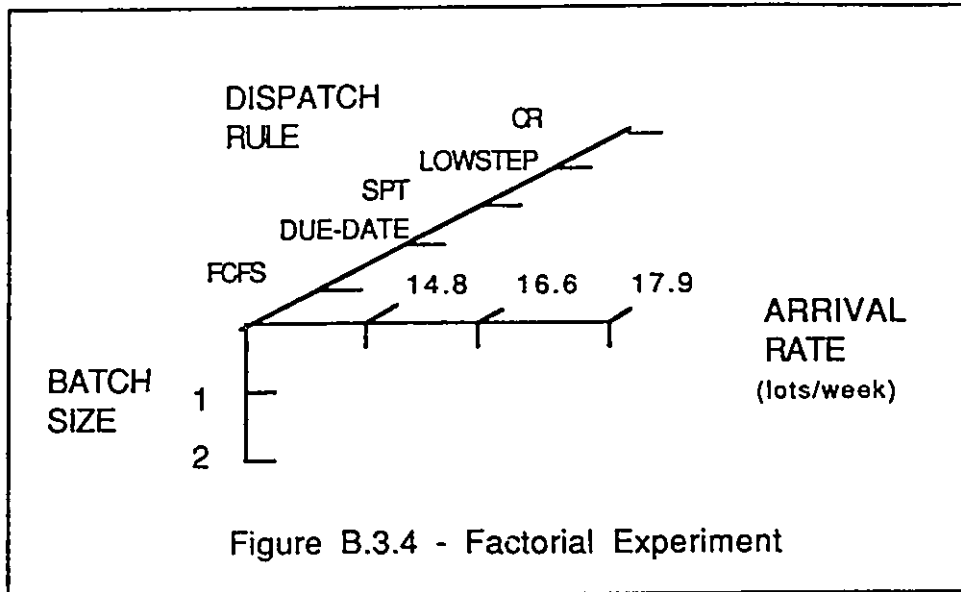


Figure B.3.4 - Factorial Experiment

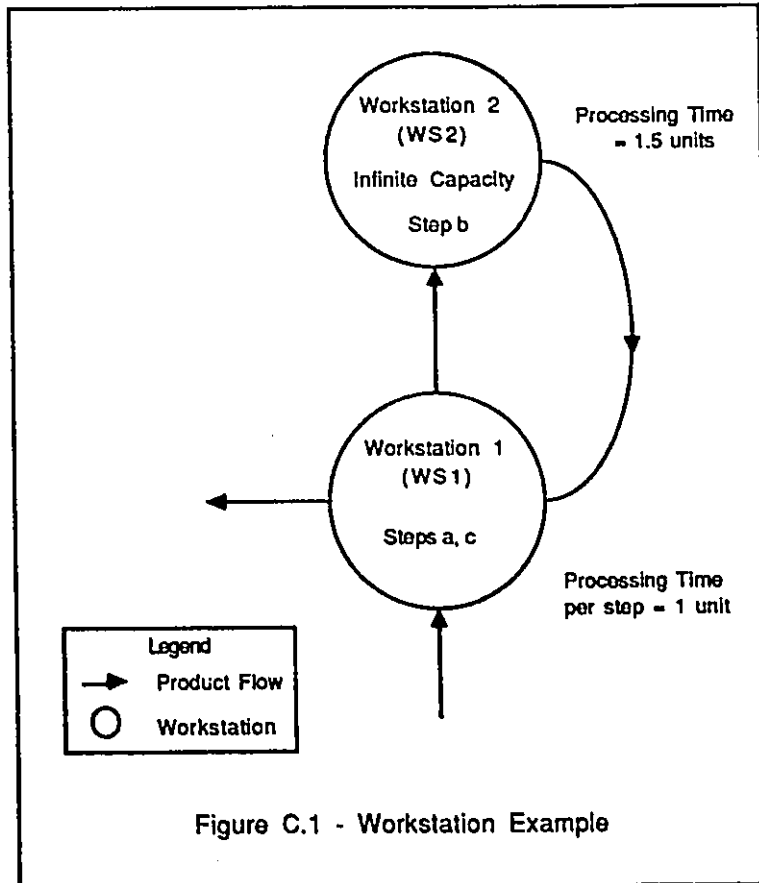
Table B.3.4 - Machine Failure - Results

Arrival Rate (lpw)	17.9		16.6		14.8	
Batch Size	2	1	2	1	2	
FCFS	80.1 (8.9)	220.0 (62.4)	67.6 (8.7)	30.3 (5.9)	67.6 (9.1)	
DUE-DATE	114.0 (14.2)	246.0 (35.0)	77.5 (9.5)	52.0 (8.6)	65.5 (7.2)	
SPT	77.1 (8.2)	195.0 (53.7)	67.1 (7.8)	28.5 (6.2)	65.5 (8.8)	
LOWSTEP	77.3 (10.3)	1141.6 (388.7)	67.0 (9.5)	32.3 (10.7)	69.3 (9.2)	
CR	114.0 (10.9)	203.6 (36.6)	72.8 (7.3)	30.6 (4.5)	66.5 (7.0)	

* - queue times measured in hours

APPENDIX C - QUEUEING EXAMPLE

This appendix contains an example of the 'zig-zag' queue time effect referred to in Chapter 4. The effect is illustrated on a two workstation model outlined in Figure C.1.

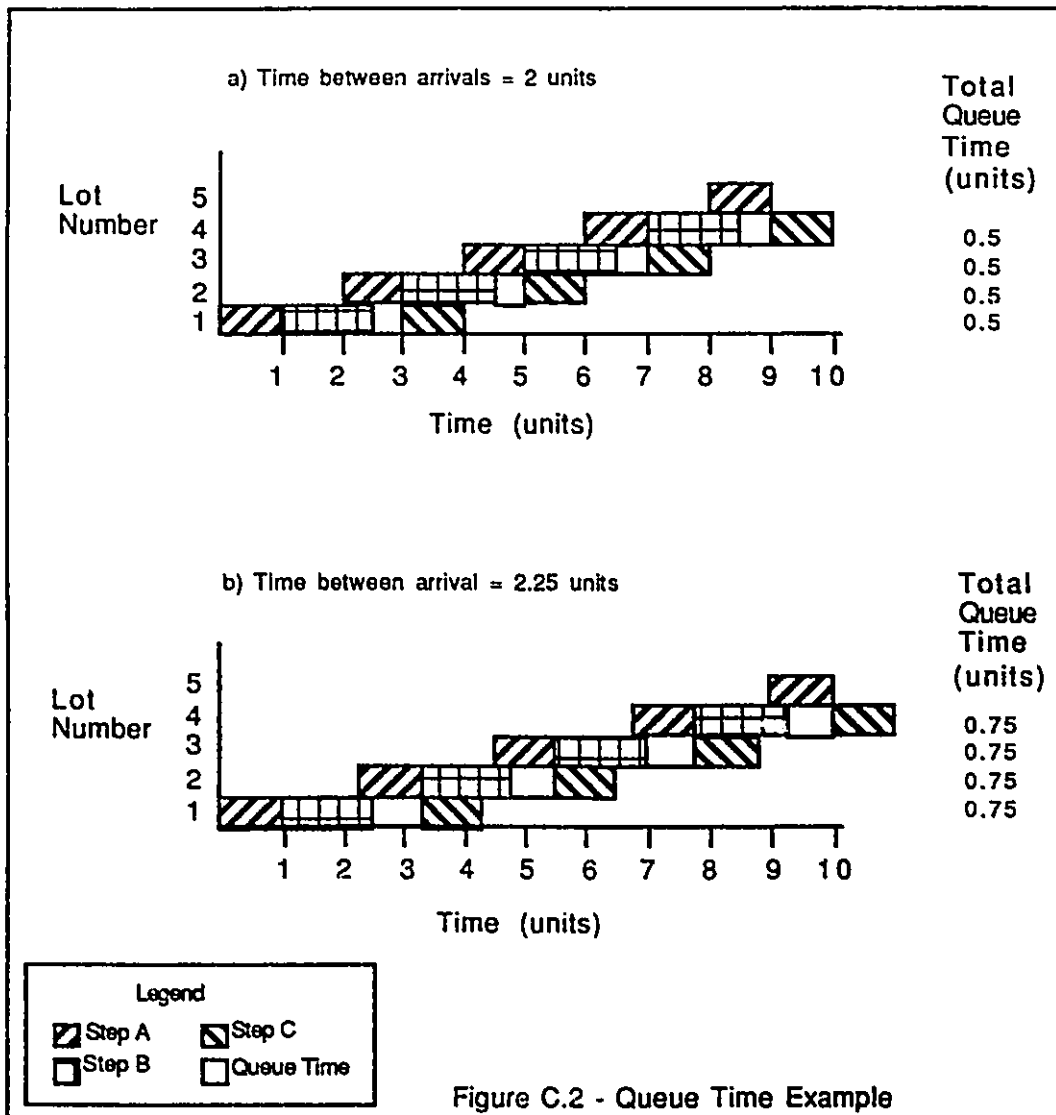


Workstation 1 (WS1) is visited on the first and last process steps in the model. WS1 can process only a single lot at a time with a machine process time of one unit for both of the steps. The dispatch rule used at WS1 is FCFS with a tie breaking rule of 'highest step first'. i.e. in the event of equal arrival times at WS1 the lot at step c is processed first.

Workstation 2 (WS2) is visited on the second step and has an infinite processing capacity. The result is that the workstation acts as a 1.5 time unit delay for lots in between visits to WS1.

The system has a capacity of 1 lot per two units of time dictated by the processing rate of WS1.

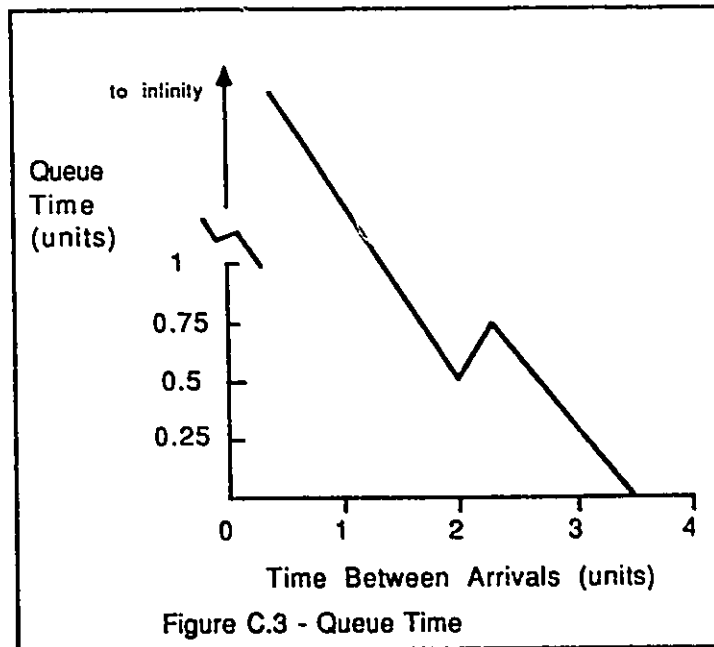
Two scenarios are presented, each with a different job arrival rate. The first scenario has a job arrival rate of one lot every two time units (the system capacity). The second scenario has a lot arrival rate of one lot every 2.25 time units. Two important points should be noted: (1) any scenario with a job arrival rate less than one lot every 3.5 time units (the total service time) has no queuing whatsoever, (2) as the time between job arrivals approaches zero the queue time per lot approaches infinity. In light of these points the scenarios for time between arrivals of 2 and 2.25 time units are illustrated in Figure C.2.



The results are summarized in Table C.1. and graphed in Figure C.3 to illustrate the counter intuitive 'zig-zag' effect. i.e. Intuitively one expects lower arrival rates to result in lower queue times.

Table C.1 - Queue Time Results

Time Between Arrivals(units)	0	2	2.25	> 3.5
Queue Time (units)	infinity	0.5	0.75	0



BIBLIOGRAPHY

- Adlakha, V.G., and Fishman, G.S., 1982, 'Starting and Stopping Rules for Simulations Using a priori Information', *E.J.O.R.*, 10, 379-394.
- Atherton, R.W. and Dayhoff, J.E., 1986, 'Signature Analysis: Simulation of Inventory, Cycle Time, and Throughput Trade-Offs in Wafer Fabrication, *IEEE Trans. on Components, Hybrids, and Manufacturing Technology*, Vol. CHMT-9, 4, December, 498-507.
- Atherton, R.W., 1987, 'Factory Scheduling Using Simulation Models', *ECS Extended Abstracts*, Abstract No. 658.
- Baker, R.B., 1984, 'Sequencing Rules and Due-Date Assignments in a Job Shop', *Management Science*, 30, 9, September, 1093-1103.
- Bitran, G.R. and Tirupati, D., 1988, 'Planning and Scheduling for Epitaxial Wafer Production Facilities', *Operations Research*, 36, 1, Jan/Feb., 34-49.
- Blackstone, J.H., Phillips, D.T., and Hogg, G.L., 1982, 'A State of the Art Survey of Dispatching Rules for Manufacturing Job Shop Operations', *Int. J. Prod. Res.*, 20, 1, 27-45.
- Bogk binder, J.H. and Noor, A.I., 1985, 'Setting Job-Shop Due-Dates with Service-Level Constraints', *Journal of the Operational Research Society*, 36, 11, 1017-1026.
- Bulkin, M.H., Colley, J.L., and Steinhoff, H.W., 1966, 'Load Forecasting, Priority Sequencing, and Simulation in a Job Shop Control System', *Management Science*, 13, 2, (B-29)-(B-50).
- Bunnag, P. and Smith, S.B., 1985, 'A Multifactor Priority Rule for Job Shop Scheduling Using Computer Search', *IIE Transactions*, 17, 2, June, 141-146.
- Burman, D.Y., Gurrola-Gal, F.J., Nozari, A., Sathaye, S., and Sitarik, J.P., 1986, 'Performance Analysis Techniques For IC Manufacturing Lines', *AT&T Bell Labs Technical Journal.*, 46-56.

- Buzacott, J.A., 1985, 'Modelling Manufacturing Systems', *Robotics and Computer-Integrated Manufacturing*, 2, 1, 25-32.
- Buzacott, J.A. and Shanthikumar, J.G., 1985, 'On Approximate Queueing Models of Dynamic Job Shops', *Management Science*, 31, 7, July, 870-887.
- Chen, H., Harrison, J.M., Mandelbaum, A., Ackere, A.V., and Wein, L.M., 1988, 'Empirical Evaluation of a Queueing Network Model for Semiconductor Wafer Fabrication', *Operations Research*, 36, 2, Mar/Apr., 202-215.
- Chang, Y.L., Sullivan, R.S., and Bagchi, U., 1985, 'Experimental Investigation of Real-Time Scheduling in Flexible Manufacturing Systems', *Annals of Operations Research*, Switzerland, 3, 355-377.
- Davies, B., 1987, 'The Virtual Fab Modelling System', *ECS Extended Abstracts*, Abstract No. 657.
- Dannenbring, D.G., 1977, 'An Evaluation of Flow Shop Sequencing Heuristics', *Management Science*, 23, 11, July, 1174-1182.
- Dar-El, E.M. and Wysk, R.A., 1982, 'Job Shop Scheduling - A Systematic Approach', *Journal of Manufacturing Systems*, 1, 1, 77-88.
- Dayhoff, J.E., and Atherton, R.W., 1985, 'Simulation and Optimization of Wafer Fab Operations', *ECS Extended Abstracts*, Abstract No. 97.
- Dayhoff, J.E., and Atherton, R.W., 1984, 'Simulation of VLSI Manufacturing Areas', *VLSI Design*, December, 84-92.
- Dayhoff, J.E., and Atherton, 1986, 'Signature Analysis of Dispatch Schemes in Wafer Fabrication', *IEEE Trans. on Components, Hybrids, and Manufacturing Technology*, Vol. CHMT-9, 4, December.
- Dayhoff, J.E. and Atherton, R.W., 1987, 'A Model for Wafer Fabrication Dynamics in Integrated Circuit Manufacturing', *IEEE Trans. on Systems, Man and Cybernetics*, 26, 2, Jan/Feb, 91-100.

- Dobson, G., Karmarkar, U.S., and Rummel, J.L., 1987, 'Batching to Minimize Flow Times on One Machine', *Management Science*, 33, 6, June, 784-799.
- Driscolls, J., and Abdel-Shafi, A.A.A., 1985, 'A Simulation Approach to Evaluating Assembly Line Balancing Solutions', *Int. J. Prod. Res.*, 23, 5, 975-985.
- Economides, S., and Cunningham, F., 1987, 'Microcomputer Simulation Improves Production Scheduling for Microelectronics Manufacturers', *Journal of Manufacturing Systems*, 6, 4, 267-275.
- Eden, C.L., 1975, 'Rules for Scheduling Semi-automatic Machines with Deterministic Cycle Times', *Int. J. Prod. Res.*, 13, 1, 41-55.
- Edwards, B.A.T., Felix, D., Holvey, R., Poulsen, R., and Sutcliffe, V., 1987, 'A Versatile Simulator/Scheduler for Silicon Integrated Circuit Manufacture', *Proceedings ECS Symposium on Automation in IC Manufacturing*, Oct.
- Eilon, S., and Chowdhury, I.G., 1976, 'Due Dates in Job Shop Scheduling', *Int. J. Prod. Res.*, 14, 2, 223-237.
- Elvers, D.A., 1973, 'Job Shop Dispatching Rules Using Various Delivery Date Setting Criteria', *Production and Inventory Management*, 4th Quarter, 62-70.
- Elvers, D.A., 1974, 'The Sensitivity of the Relative Effectiveness of Job Shop Dispatching Rules with Respect to Various Arrival Distributions', *AIIE Transactions*, 6, 4, March.
- Fishman, G.S., 1971, 'Estimating Sample Size in Computing Simulation Experiments', *Management Science*, 18, 1, 21-37.
- Fishman, G.S., 1978, 'Grouping Observations in Digital Simulations', *Management Science*, 24, 5, 510-521.
- Fredericks, A.A., 1986, 'Performance Analysis Modeling for Manufacturing Systems', *AT&T Technical Journal*, 65, 4, July/Aug., 25-34.

- Gere, W.S., 1966, 'Heuristics in Job Shop Scheduling', *Management Science*, 13, 3.
- Glasse, C.R. and Resende, M.G.C., 1988, 'Closed-Loop Job Release Control for VLSI Circuit Manufacturing', *IEEE Transactions on Semiconductor Manufacturing*, 1, 1, Feb., 36-46.
- Gupta, Y.P., Bector, C.R., and Gupta, M.C., 1988, 'A Review of Scheduling Rules in Flexible Manufacturing Systems', Dept. of Actuarial and Management Sciences, Faculty of Management, University of Manitoba.
- Huang, P.Y., Moore, L.J., and Russell, R.S., 1984, 'Workload vs Scheduling Policies in a Dual-Resource Constrained Job Shop', *Computer & Operations Research*, 11, 1, 37-47.
- Jackson, J.R., 1957, 'Simulation Research on Job Shop Production', *Naval Research Logistics Quarterly*, 4th quarter, 287-295.
- Jones, C.H., 1973, 'An Economic Evaluation of Job Shop Dispatching Rules', *Management Science*, 20, 3, November, 293-307.
- Karmarkar, U.S., 1987, 'Lot-Sizing and Sequencing Delays', *Management Science*, 33, 3, March, 419-423.
- Kiran, A.S. and Smith, M.L., 1984a, 'Simulation Studies in Job Shop Scheduling-I, A Survey', *Computer & Industrial Engineering*, 8, 2, 87-93.
- Kiran, A.S. and Smith, M.L., 1984b, 'Simulation Studies in Job Shop Scheduling-II, Performance of Priority Rules', *Computer & Industrial Engineering*, 8, 2, 95-105.
- Kleijnen, J.P.C., vander Ven, R., and Sanders, B., 1982, 'Testing Independence of Simulation Subruns: a Note on the Power of the Von Neumann Test', *E.J.O.R.*, 9, 92-93.
- Kleijnen, J.P.C., 1984, 'Statistical Analysis of Steady-state Simulations: Survey of Recent Progress', *E.J.O.R.*, 17, 150-162.
- Kleinrock, L., 1976, 'Queuing Systems', Volumes I,II, John Wiley & Sons, New York, N.Y.

- Kochhar, S. and Morris, R.J.T., 1987, 'Heuristic Methods for Flexible Flow Line Scheduling', *Journal of Manufacturing Systems*, 6, 4, 299-314.
- Krigline, A.G., Kuzdrall, P.J., and Villeda, R.R., 1988, 'Simulating Assembly Line Operations', Presented to the CORS National Meeting, Montreal, May.
- Lageweg, B.J., Lenstra, J.K., and Rinnooy Kan, A.H.G., 1978, 'A General Bounding Scheme for the Permutation Flow Shop Problem', *Operations Research*, 26, 1, Jan/Feb., 53-57.
- Lane, D.E., 1988, Private communication.
- Lane, D.E. and Sidney, J.B., 1988, 'Batching and Scheduling in FMS Hubs: Flow Time Considerations', Presented to TIMS/ORSA Annual Meeting, Washington, D.C., April 25-27.
- Lau, H.S. and Zaki, A., 1984, 'The Effect of Using Discrete Time Units in Queueing Simulation', *E.J.O.R.*, 17, 163-168.
- Leeke, S.D., Davies, B., and Saraswat, K., 1987, 'The Virtual Fab Modelling System', *ECS Extended Abstracts*, Abstract No. 657.
- Machida, A., 1987, 'Using Simulation to Plan and Examine a Semiconductor Fab Operation', *ECS Extended Abstracts*, Abstract No. 660.
- Maimon, O.Z. and Gershwin, S.B., 1988, 'Dynamic Scheduling and Routing for Flexible Manufacturing Systems That Have Unreliable Machines', *Operations Research*, 36, 2, Mar/Apr., 279-292.
- Mamiliias, A.G., Bilalis, N.G., and Konstantinidis, 1987, 'On Simulation Modelling for FMS', *Simulation*, 48, 1, January, 19-23.
- Neelamkavil, J. and Thompson, V., 1986, 'Use of Dispatching Rules in Manufacturing Applications', 5th Canadian CAD/CAM and Robotics Conf., Toronto, CANADA, June 17-19, MS86-720.
- Nelson, E.A. 1983, 'The Use of a Generic Simulation Model in Evaluating Manufacturing Processes', *Computers & Industrial Engineering*, 7, 3, 217-224.

- Nelson, R.T., 1967, 'Labour and Machine Limited Production Systems', *Management Science*, 13, 9, May, 648-671.
- Nelson, R.T., Sarin, R.K., and Daniels, R.L., 1986, 'Scheduling with Multiple Performance Measures: The One Machine Case', *Management Science*, 32, 4, April, 464-479.
- O'Grady, P.J., and Harrison, C., 1985, 'A General Search Sequencing Rule for Job Shop Sequencing', *Int. J. Prod. Res.*, 23, 5, 961-973.
- Park, Y.B., Peyden, C.D., and Ensore, E.E., 1984, 'A Survey and Evaluation of Static Flowshop Scheduling Heuristics', *Int. J. Prod. Res.*, 22, 1, 127-141.
- Panwalker, S.S. and Iskander, W., 1977, 'A Survey of Scheduling Rules', *Operations Research*, 25, 1, Jan/Feb., 45-61.
- Poulsen, R. and Sutcliffe, V., 1987, 'Sensitivity Analysis and Best Compromise Optimization of Furnace Batch Sizes in IC Factory Simulation', *Proceedings ECS Symposium on Automation in IC Manufacturing*, October.
- Pritsker, A.A.B., 1986, 'Introduction to Simulation and SLAMII', 3rd Edition, Halsted Press, John Wiley & Sons, New York.
- Rochette, R. and Sadowski, R.P., 1976, 'A Statistical Comparison of the Performance of Simple Dispatching Rules for a Particular Set of Job Shops', *Int. J. Prod Res.*, 14, 1, 63-75.
- Rowe, A.J., 1958, 'Sequential Decision Rules in Production Scheduling', unpublished doctoral dissertation, University of California, Los Angeles, August.
- Russell, R.S., Dar-El, E.M., and Taylor, B.W., 1987, 'A Comparative Analysis of the COVERT Job Sequencing Rule Using Various Shop Performance Measures', *Int. J. Prod. Res.*, 25,10, 1523-1540.
- Russell, R.S. and Taylor, B.W., 1985, 'An Evaluation of Scheduling Policies in a Dual Constrained Assembly Shop', *IIE Transactions*, 17, 3, September, 219-231.

- Sen, T. and Gupta, S.K., 1984, 'A State-of-Art Survey of Static Scheduling Research Involving Due Dates', *Omega-The International Journal of Management Science*, 12, 1, 63-76.
- Sidney, J.B., 1977, 'Optimal Single Machine Scheduling with Earliness and Tardiness Penalties', *Operations Research*, 25, 1, Jan/Feb., 62-69.
- Smith, M.L. and Seidmann, A., 1983, 'Due Date Selection Procedures For Job-Shop Simulation', *Computer & Industrial Engineering*, 7, 3, 199-207.
- Sze, S.M., 1983, 'VLSI Technology', McGraw-Hill, New York.
- Till, W.C. and Luxon, J.T., 1982, *Integrated Circuits: Materials, Devices, and Fabrication*, Prentice-Hall, Inc., Englewood Cliffs, N.J.
- Wein, L.M., 1988, 'Scheduling Semiconductor Wafer Fabrication', *IEEE Transactions on Semiconductor Manufacturing*, 1, 3, August, 115-130.
- Worrall, B.M., and Mert, B., 1980, 'Application of Dynamic Scheduling Rules in Maintenance Planning and Scheduling', *Int. J. Prod. Res.*, 18, 1, 57-71.
- Yao, D.D., and Buzacott, J.A., 1985, 'Modelling the Performance of Flexible Manufacturing Systems', *Int. J. Prod. Res.*, 23, 5, 945-959.
- Zeigler, B.P., 1984, 'System-Theoretic Representation of Simulation Models', *IIE Transactions*, 16, 1, March.

END

05.12.90

FIN