Proceedings of the 1994 Winter Simulation Conference
ed. J. D. Tew, S. Manivannan, D. A. Sadowski, and A. F. Seila

INTRODUCTION TO MANUFACTURING APPLICATIONS

Gordon M. Clark

Department of Industrial, Welding and Systems Engineering
The Ohio State University
Columbus, Ohio 43210, U.S.A.

ABSTRACT

This tutorial introduces manufacturing applications of simulation through three illustrative example applications. These examples illustrate the additional understanding of system behavior gained by the use of simulation models. Individuals using simulation should use a structured process in applying simulation. The second example illustrates this structured process. The examples also illustrate the use of both stochastic and deterministic variables in modeling manufacturing systems.

1 INTRODUCTION

Manufacturing is one of the earliest simulation application areas (Naylor et al 1966), and the attendance at the manufacturing application track of the Winter Simulation conferences indicates that manufacturing remains as one of the most popular application areas. We use simulation to improve the performance of manufacturing systems because:

- Many manufacturing systems are too complex to be analyzed and improved by simply thinking and talking about possible approaches, and
- Simulation can predict system performance resulting from interactions among system components.

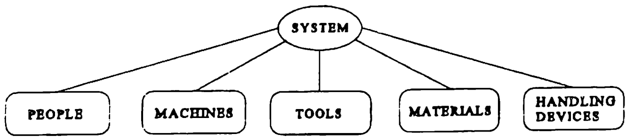

Figure 1: System Components

System components can be people, machines, tools, material handling devices, and materials as depicted in Figure 1. The result of interactions among these components may be very difficult to predict without the use of a model, and a simulation model is frequently the easiest model to use.

This tutorial introduces simulation applications to manufacturing systems by illustrating:

- Diverse uses of simulation,
- Use of random and deterministic variables in simulation models
- A structured process for applying simulation

This tutorial uses three example applications to illustrate the above points. The first example is a single-machine-failure model. It illustrates the value of using a model even when studying the performance of a single machine. The second example is a study to determine effective operating policies for a cell. This example illustrates the steps and process that one ought to follow in a simulation study. The third example is the use of simulation in an on-line mode to schedule a manufacturing system. The simulation model is a completely deterministic model in this application.

2 SINGLE-MACHINE-FAILURE MODEL

Results from the single-machine-failure model illustrate the need for a model in analyzing a machine purchasing decision, and the model shows requirements for using a simulation. Law and Kelton (1991) present similar results in their example 13.4. A company is considering the purchase of a machine from a vendor that advertises that the machine has the following performance characteristics:

- Constant time to process a job, i.e., 1 minute
- A mean time to failure of 540 minutes, i.e., 9 hours, of actual operating time
- A mean time to repair a failure of 60 minutes, i.e., 1 hour

Operating time does not include idle time. The company wants to predict the effect of the machine characteristics on throughput rate, work-in-process (WIP), and throughput time. Throughput time is the total time at the machine including both queueing time, processing time, and waiting on a repair to be completed.

The above characteristics permit one to calculate an upper limit on the capacity or the throughput rate of the machine. The first step is to obtain a mean service time for a part on the machine. Service time will include the processing time and any time a job has to wait for the machine to be repaired. Define a cycle starting when the machine is repaired, operates until failure, undergoes repair, and then starts another cycle. The cycle length is 10 hours, and the machine is being repaired for 1 hour or 10% of the cycle length. Thus, the effective mean time to service a job is 1/.9 minutes. An equivalent viewpoint is that a job will require 1 minute of processing plus any time to remedy a failure. On the average, one out of every 540 jobs will experience a failure, and have to wait for repair. Since the mean repair time is 60 minutes, the mean service time for a part is 1 + 60/540 = 1/.9 minutes. The mean service time of 1/.9 minutes implies a mean service rate of 1/(1/.9) = .9 jobs per minute. Define the following quantities, having units of jobs per minute, that apply over an extended time period.

λ = the average arrival rate of jobs to the machine
θ = the average throughput rate

In this case, θ = minimum(λ, .9). The above example illustrates that we can calculate the mean throughput rate without a simulation model.

To predict the average WIP and throughput time, we will use a simulation. The variability in job interarrival times and job service times is important in predicting queueing and queueing times. However, the distribution of times to failure and of machine repair times are unknown. Assume that the probability distribution of job inter-arrival times is negative exponential with mean 1.25 minutes. Then, λ = 1/1.25 = .8 jobs per minute. Thus, θ = .8 jobs per minute. For the simulation of job service times, consider the three cases shown in the following table.

Case	Description
STOFAIL	Operating times to failure have a negative exponential distribution with mean 540 minutes Repair times have a lognormal distribution with mean 60 minutes and standard deviation 20 minutes
DETFAIL	Operating times to failure are a constant 540 minutes Repair times are a constant 60 minutes
NOFAIL	Service times are a constant 1/.9 = 1.111 minutes

Results for the above three cases will show the effect of variability in service times. We simulated 100 independent shifts of 8 hours production for each case. A machine under repair at the end of a shift will be operable at the start of the following shift. Also, each shift starts with the machine and its associate queue being empty and idle. The following table gives 90% confidence intervals for mean throughput time, WIP, and number in the queue.

Case	Throughpu t Time	WIP	Number in Queue
STOFAIL	17.6 ± 4.0	14.4 ± 3.0	13.5 ± 3.0
DETFAIL	12.3 ± 1.6	10.2 ± 1.3	9.3 ± 1.3
NOFAIL	5.02 ± .32	4.06 ± .29	3.18 ± .28

Observations based on the above table are:

- Service time variability has an important effect on WIP and throughput time.
- The timing of machine failures and/or stoppages must be explicitly represented as opposed to allocating an average down time to the processing time of each job. Law and Kelton (1991) make this observation.
- A deterministic representation of time to failure and time to repair gives larger throughput time and WIP estimates than allocating average down time to each job.
- The STOFAIL and DETFAIL confidence intervals overlap; thus, one should run more than 100 replications for these two cases.

This single-machine-failure model illustrates an important property of simulation models. A simulation is an atomistic model in that the simulation describes the performance of individual system components and their interactions. Given this description, the simulation model can trace the changes in system state over time. This means that the simulation user must specify a description of the system components and their performance characteristics to implement a simulation. In the above example, the user specified the distribution of job inter-arrival and service times for each case. To represent the variation in machine failure times, the user must specify or select a particular distribution.

Why is variability so important in estimating WIP and throughput time? Any utilization close to one will result in excessive work-in-process (WIP) if there is any variation in service times or times between arrival of lots to the respective machines Figure 2 illustrates the effect of fluctuations in job inter-arrival times and job operation times on WIP for a single machine. The top sequence, called case 1, of job arrivals and service times for each arrival follows a perfect uniformly spaced pattern that has no variation. That is, the times between each arrival are all equal and the service times are also all equal. The proportion of time the machine is busy represents the machine utilization which is close to one. Because of this lack of variation, case 1 gives no queueing and no instances of WIP greater than 1. The lower sequence, called case 2, of service and inter-arrival times have precisely the same mean and gives the same overall machine utilization which is the proportion of time the machine is busy. However, this statistical fluctuation increases the WIP which becomes as large as three. The shaded area in plot at the bottom of the figure shows the jobs waiting in the machine queue.

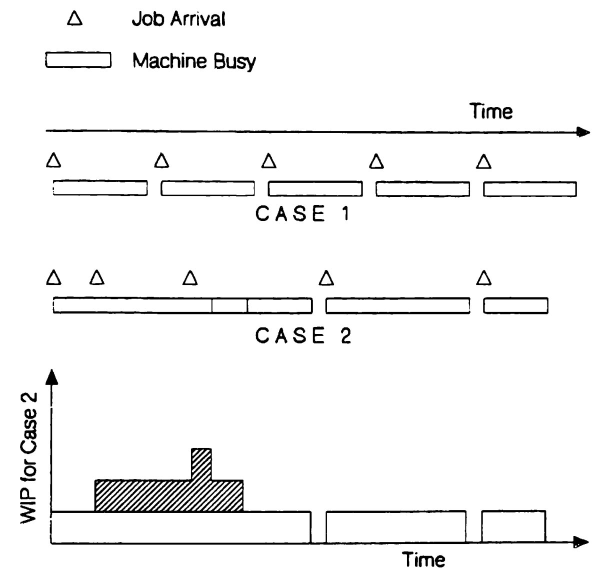

Figure 2: Effect of Statistical Fluctuations on WIP

Potential sources of variations in job service times are:

- Tooling failures
- Machine cycle length changes due to different types of jobs, i.e., a machine performs operations on non-identical parts. For example, a machine processes an XYZ123 job and then an ABC123.
- Machine failures and adjustments
- Variations in human paced task times

Variations in inter-arrival times could result from:

- Any variation in the times between release to production due to the company planning system or customer order times, e.g., job release times that vary with the hour of the day or the day of the week.
- Variations in the times materials arrive from vendors
- Variations in initiation of production caused by tooling not being available
- Variations in the times jobs depart from upstream work stations in the job's route

See Law and Kelton (1991) for another list of potential sources of statistical variations in manufacturing simulations.

3 GEAR MANUFACTURING THROUGHPUT TIME

Figure 3: Gear Production Flow

Clark and Cash (1993) used simulation in a study to identify preferred operating policies for a rough steel cell used in the manufacture of precision gears. Figure 3 depicts the three cells used to produce gears. The manufacturer produces gears to order, rather than making gears for stock. Customer orders may specify a gear that the manufacturer has produced in the past, but the elapsed time between repeat orders is so long that making gears for stock is not economical. Each customer order specifies a quantity of gears that varies over a wide range. The flow allowance is the lead time quoted to the customer specifying the promised delivery date. Currently, the type of gear determines the flow allowance, but no allowance is made for the number of gears in the customer order. The manufacturer currently releases work to production as soon as raw materials are available to produce the customer order; thus, the number of gears in a job has a considerable range of variation. The manufacturer uses manual procedures for tracking and scheduling work in the plant.

This study illustrates the process of using simulation to generate recommendations for management action. This process includes the following steps.

- Specify study objectives
- Specify performance measures
- Determine alternatives to investigate
- Describe systems to be simulated
- Specify system experimental conditions
- Create simulation model
- Prepare input data
- Formulate experimental design
- Conduct simulation experiments
- Analyze results
- Make recommendations

3.1 Study Objectives

The manager of manufacturing engineering and the director of engineering requested a study to determine policies for scheduling work in the rough steel cell. These scheduling policies consist of policies for controlling the release of work to the cell and sequencing work in the cell. The objectives for these policies are to:

- Reduce throughput time through the cell
- Reduce WIP
- Reduce quoted lead time
- Reduce tardiness

- Reduce cost

The tardiness objective requires establishing flow allowances and due dates specifically for the rough steel cell. This study emphasized simplified procedures for scheduling because of:

- The objective of reducing cost and
- The lack of a computerized procedure for tracking work in the plant

3.2 Performance Measures

The primary performance measures are:

- Average WIP
- Average system time
- Average number of tardy jobs per year
- Average time a tardy job is late
- Quoted lead times for each type of job

3.3 Alternatives Investigated

The alternatives investigated included fixed capacity buffers, a modified due-date procedure, an upper limit on job size, and a sequencing rule. The following paragraphs describe the alternatives.

Fixed Capacity Buffers: The use of fixed capacity buffers at each work station is a simplified means for reducing WIP. A station is blocked, becoming inactive, when it completes work on a job and the next station in a job's route has a full buffer. Reducing WIP also simplifies the scheduling problem for work in the cell. If the buffers do not significantly reduce capacity, by the occurrence of blocking, the reduced WIP will reduce throughput time. The use of buffers forces incoming orders to wait in a backlog when the first work station in the processing plan has a full buffer. Thus, the use of buffers introduces a control on the timing of production release. A similar alternative is to define an upper limit on the number of jobs in the entire cell. This alternative is known as the CONWIP alternative (Spearman et al 1990).

Modified Due-Date Procedure: We defined a modified due-date procedure that incorporated the number of gears in an order to determine the flow allowance. The modified flow allowance has two components. One for the aggregate setup time, and one for the aggregate run time per gear. The run time component is proportional to the order quantity. For most customer orders, the modified procedure has a shorter flow allowance than the current flow allowances.

Job Size: An upper limit on job size or the number of gears in a job will reduce the large variation in the number of gears in a job. Large jobs tend to create floating bottlenecks. A customer order for more gears than the job size limit will result in multiple jobs to fill an order.

Sequencing Rules: We investigated two sequencing rules for work at a work station. They were first-in-first-out (FIFO) and earliest due date (EDD).

3.4 System Description

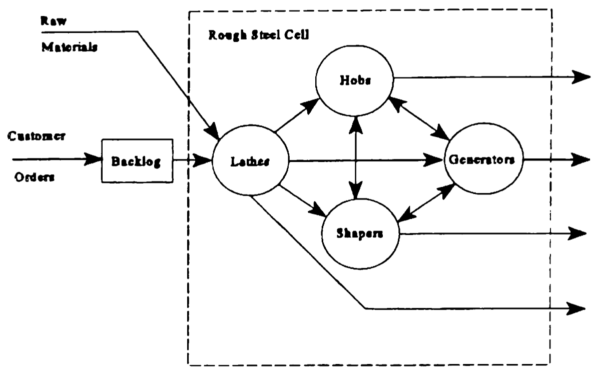

Figure 4: System Studied

Figure 4 depicts the system studied. The rough steel cell has the following work stations: lathes, hobs, shapers, and generators. Each work station may have a buffer and multiple machines. The buffer sizes and number of machines in each work station are inputs. The service time for a job at a machine has a setup time and a run time component. The setup times and single part run times are lognormal random variables. The total service time for a job is the sum of the setup time and lotSize independent run times, where lotSize is the number of gears in the job. The system represents numGears different gear types, where each gear type has its own processing plan. A processing plan gives the route for a gear type through the cell and the standard setup and run times. The arrival times of customer orders are exogenous, deterministic inputs.

3.5 Experimental Conditions

The director of engineering and the manager of manufacturing engineering selected 50 gear types for analysis. That is, numGears = 50. The gears selected are representative of future business. They supplied the process plans for each gear type.

The manager of information systems supplied historical job release times over the previous four years. These data became the basis for the exogenous customer order times. The study used three different customer order patterns, known as release schedules, i.e., RS1, RS2, and RS3. Each release schedule gives specifies the time materials are available for production for each customer order over a one-year period. The intent is to represent more than a single scenario to increase the robustness of study conclusions. These release schedules present the lathe work station with average utilization levels of 65%, 85%, and 95% for RS1, RS2, and RS3, respectively. These averages apply over a one year period.

3.6 Study Requirements

The five previous steps; i.e., specify study objectives, performance measures, alternatives to investigate, system to simulate, and experimental conditions; place requirements on the study. They dictate the detail in the simulation model and the data to be collected. All concerned parties should review the results of these steps prior to making simulation runs and recommendations.

3.7 Simulation Model

The simulation model was programmed in WITNESS which permitted animation of the simulations. The animated display was effective in showing company management the nature of the simulation. Two additional programs, written in C++, simplified the use of WITNESS considerably. These programs prepared inputs for WITNESS and analyzed the WITNESS output data. The extensive inputs required to represent the large number of different gear processing plans, i.e., 50, and their flow allowances motivated the input program.

3.8 Prepare Input Data

An analysis of shop labor records, supplied by the manager of information systems, provided historical data on actual times to implement the process plans for the fifty gears. The study assumed that the coefficient of variation for setup and run times at a machine group is the same for all 50 gears. That is, the ratio between the standard deviation and mean of a setup (run) time is a constant for a machine group. The estimation of these coefficients of variation used historical data.

3.9 Experimental Design

The primary objective of the first set of simulation experiments was to determine the effectiveness of buffers in limiting WIP without significantly reducing capacity. This set of experiments imposed no limit on job size. These experiments had four factors, i.e., buffer configuration, flow allowance procedure, release schedule, and sequencing rule. The following table shows the levels of each factor. Each possible combination of the levels for each factor was simulated in the first set of experiments for a total of 72 simulation runs.

Factor	Levels
Buffer Configuration	Buffer size of 1 at each station Buffer size of 2 at each station Buffer size of 3 at each station CONWIP with WIP limited to x CONWIP with WIP limited to y No WIP Control (unlimited buffer size)
Flow Allowance	Current procedure Modified
Release Schedule	RS1 RS2 RS3
Sequencing Rule	FIFO EDD

Stochastic simulations of this type present two experimental problems (Law and Kelton 1991). That is, the initial condition effect and run length so that confidence intervals are sufficiently narrow. The release schedules apply over a one-year period. Each simulation run consisted of eleven consecutive years by repeating the appropriate release schedule eleven times. Thus, the final simulation state at the end of December became the initial condition for the next January. The C++ post-processor program deleted the first year to reduce the initial condition effect. The analysis assumed that statistics for each subsequent year are independent and identically distributed, which is the batch means procedure. The post-processor program employed these assumptions in calculating 90% confidence intervals which were sufficiently narrow.

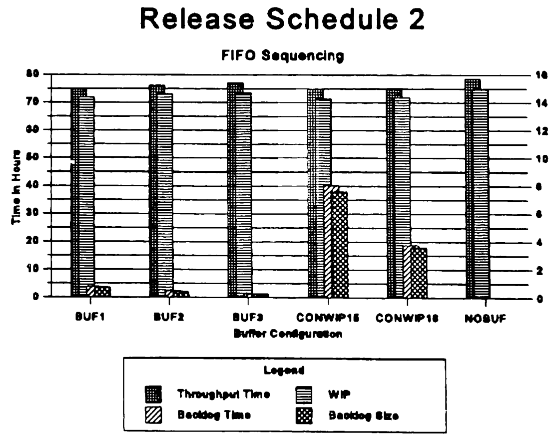

Figure 5: Throughput Time and Wip

Based on results from the first set of experiments, the analysis identified a preferred buffer configuration, sequencing rule, and flow allowance procedure. Further simulation experiments investigated the upper limit on job size.

3.10 Simulation Results

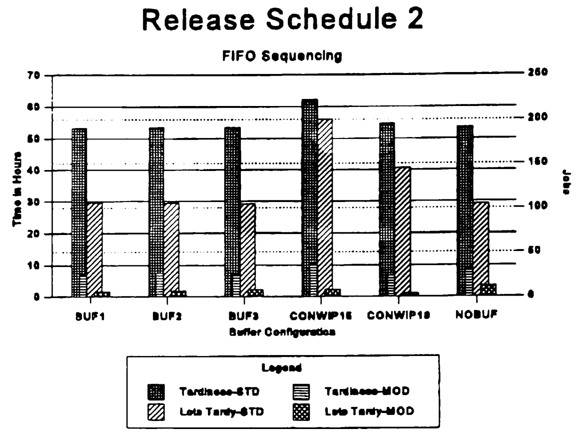

Figure 6: Tardiness

Figures 5 and 6 illustrate the results from the simulation experiments. In both figures, the use of buffers at each station dominates the CONWIP results. The system and WIP performance measures apply to the shop after leaving the backlog. Throughput time and WIP are less with a buffer size of 1 at each station. However, the total of backlog time and throughput time are slightly larger than the results with no buffers. Figure 6 clearly show the superiority of the modified due-date procedure. Also, for the modified procedure, the tardiness results for a buffer size of 1 are slightly less than tardiness with no buffers. The manufacturer prefers a buffer size of 1 since:

- WIP is less reducing costs, improving quality, and simplifying scheduling
- Tardiness is lower

3.12 Major Points Illustrated

The gear manufacturing throughput time example illustrates the overall steps required to apply simulation and influence management decisions. An important milestone is to review the first five steps with all concerned parties before collecting data and programming the model. Then, affected individuals will feel they are a part of the study. Also, the simulation experimental results can address the study objectives and provide the proper outputs. The effort in programming the simulation is usually a small part of the overall study effort. Data collection can require a major part of the study effort.

4 SIMULATION-BASED SCHEDULER

FACTOR (Pritsker Corporation 1989) is an example of a simulation-based scheduling system. A scheduler will use FACTOR in an on-line mode. That is, FACTOR will take inputs from an existing data base and then generate schedules after a short time delay such as a half hour. The data base will specify the status of all jobs in the system, process plans for these jobs, standard setup and run times, and the status of resources such as machines. For many applications, the principal output for the simulation is a schedule giving the times that jobs are processed by resources. Shop personnel can use this schedule to insure that other resources such as tools are available when the schedule requires them. The schedule also identifies which jobs will be probably be late. The simulation can do "what if" comparisons. As an example, the simulation may compare sequencing rules such as earliest due date and shortest processing time. The motivation for using simulation is that it automates the scheduling task and simulation is capable of generating realistic schedules. The schedules are realistic in the sense that the simulation represents the finite capacity of resources in a detailed manner.

FACTOR has a completely deterministic simulation. That is, FACTOR does not sample from probability distributions in generating a schedule. Since the scheduler must generate a single schedule, a deterministic representation simplifies this task. Also, by accessing a data base specifying the process plans and standard times for all jobs, the nature of each simulated task is known in more detail than simulating in a planning mode. For example, when simulating to identify preferred designs for a production line, the precise sequence of each job type may not be known.

Cheselka (1992) describes the use of FACTOR to schedule Timken's Gambrinus Thermal Treatment Facility. Scheduling that facility is challenging because the scheduler must balance two conflicting objectives.

- Complete orders by their due date
- Maximize furnace utilization
- Minimize energy costs

These objectives can conflict because maximizing utilization and minimizing energy costs would sequence jobs to avoid changes in furnace temperature and speed of material handling devices transporting jobs through the furnace. FACTOR uses a scheduling logic that first identifies the highest priority jobs using critical slack.

Critical_Slack = Firm_Plan_Date - Current_Time - Estimated_Processing_Time

If the critical slack for an job is less than 30 hours then the job is considered critical. The system assigns a higher priority to critical jobs, and they are scheduled first.

Within the same priority level, FACTOR will maximize furnace utilization by searching for a job that matches the current furnace setup after completing a job. The setup includes furnace temperature and speed of the material handling device.

The timeliness of schedules depends on the ability to quickly obtain inputs from an existing data base. At Timken, the data inputs to FACTOR include data from the following data bases.

- The VAPP data base supplies job due dates and current job work center locations.
- The RODS data base supplies detailed order information such as product size and special processing data
- The Heat Chemistry data base supplies a heat chemistry analysis for each job

5 CONCLUSIONS

The three examples summarized in this paper illustrate important applications of simulation in manufacturing. Simulation is a powerful approach to modeling manufacturing systems in that many complex and diverse systems can be represented. Simulation can predict system performance measures that are difficult to assess without a model. However, simulation requires data that characterizes the behavior of system components. Also, individuals contemplating the use of simulation should use a structured process such as the one described in Section 3.

REFERENCES

Cheselka, Michael C. 1992. The Timken Company's Gambrinus Thermal Treatment Facility Scheduling System. In *Proceedings of the 1992 Winter Simulation Conference*, ed. J.J. Swain, D. Goldsman, R.C. Crain, and J.R. Wilson, 833-841, Institute of Electrical and Electronics Engineers, San Francisco, California.

Clark, Gordon M. and Charles R. Cash. 1993. *Reduction of Gear Manufacturing Throughput Time: Rough Steel Cell Operating Policies*. Department of Industrial and Systems Engineering, The Ohio State University, Columbus, Ohio.

Law, Averill M. and W. David Kelton. 1991. *Simulation Modeling & Analysis*. McGraw-Hill, Inc., New York.

Naylor, T.H., J.L. Balintfy, D.S. Burdick, and K. Chu. 1966. *Computer Simulation Techniques*. John Wiley & Sons, Inc., New York.

Pritsker Corporation. 1989. FACTOR Version 4.0 Implementation Guide. Pritsker Corporation, Indianapolis, Indiana.

Spearman, M. L., D. L. Woodruff, and W. J. Hopp. 1990. CONWIP: A Pull Alternative to Kanbans. *International Journal of Production Research* 22:879-894.

AUTHOR'S BIOGRAPHY

GORDON M. CLARK is a professor in the department of Industrial, Welding and Systems Engineering at The Ohio State University. He received the B.I.E. degree from The Ohio State University in 1957, the M.Sc. in Industrial Engineering from The University of Southern California in 1965, and the Ph.D. degree from The Ohio State University in 1969. His current research interests include integrated manufacturing decision support systems in a CIM environment and the design and analysis of efficient simulation experiments. He is an associate editor for the *ORSA Journal on Computing*. In 1991, he served as program chair for the WSC.