Proceedings of the 1988 Winter Simulation Conference
M. Abrams, P. Haigh, and J. Comfort (eds.)

A statistical method for generating order information for simulation modeling and analysis

David L. Martin
Pritsker & Associates, Inc.
1305 Cumberland Avenue
P.O. Box 2413
West Lafayette, IN 47906

and

F. Bradley Armstrong, IV
Hughes Aircraft Co.
P.O. Box 9399
Bldg. C-5, MS: 2016
Long Beach, CA 90810

ABSTRACT

The flexibility and sensitivity of manufacturing systems to changes in customer orders, both quantity and composition, is of vital importance. Evaluation of systems using simulation models is sometimes restricted because of a lack of actual customer order data. This paper describes a method for generating orders for use as input to manufacturing simulation models. These orders have characteristics based upon the statistical properties of a representative sample of actual orders processed by the facility. Two examples of the application of the method to actual simulation studies are presented.

1. INTRODUCTION

The fundamental purpose of performing a simulation analysis is the accurate modeling of a physical system so that valid experiments can be performed and useful solutions can be obtained (Law and Kelton 1982; Pritsker 1986). Underlying this analytical process, as shown in Figure 1, is the assumed availability of actual data obtained from the facility being studied. As simulation is used to model larger portions of manufacturing facilities, the associated data requirements are becoming enormous. In non-computerized environments, even when the data is available, it must be input manually which is very time consuming and generally unacceptable. The importance of actual data (especially order data) for model validation and subsequent experimentation cannot be overemphasized. It is clear that if the straightforward approach to data collection is unable to satisfy requirements, alternative methods of obtaining data must be explored. The purpose of this paper is to describe one such alternative.

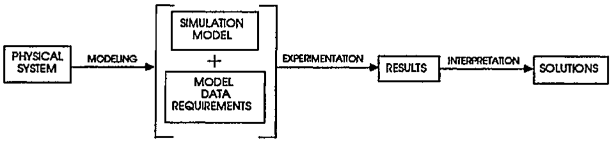


Figure 1. Simulation Analysis Process

This paper first presents a method of generating orders based upon the statistical properties of a representative sample of actual orders (Martin and Armstrong 1984). It then describes two simulation studies in which this method was applied. The last section of this paper discusses the benefits of the method and potential uses and applications that lend themselves to this tool.

2. ORDER GENERATION PROCESS

The objective of this effort is to use a representative sample of orders currently or recently involved in operations at a manufacturing facility to generate a larger set or additional sets of orders to drive a simulation model of the facility. This allows the analyst to employ more than one set of orders or a set of orders large enough to drive the model for a significant period of time, as input to the model, with which to effectively evaluate system performance. A requirement is that the generated orders should demonstrate characteristics similar to the actual orders. In this way the method will enable evaluation of the system's performance when limited data is available.

The steps, as shown in Figure 2, involved in this procedure are:

- Determine the pattern of order arrivals to the system.
- Determine the characteristics of each order.
- Validate the method.

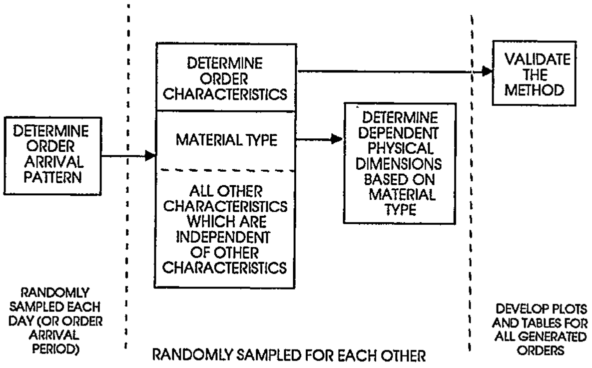


Figure 2. Steps in the Order Generation Process

The method described in the succeeding sections requires data concerning both order arrival patterns and order characteristics. The arrival pattern pertains to the quantity of orders arriving to the system or due to be completed during a given time period, while the order characteristics include scheduling data, such as due date, and physical data, such as material type, gage, width, weight, length, and number of pieces in the order.

2.1 Order Arrival Pattern

The method of determining the pattern of order arrivals is usually a heuristic method based on historical data from the facility and a forecast of market potential. Due to constantly changing economic conditions an exact order arrival pattern is difficult, if not impossible, to forecast. For our purpose either an *expert's* best estimate or a distribution based upon historical data of order arrival quantities is used.

2.2 Order Characteristics

Once an order has been generated it requires certain characteristics to be attached (i.e., due date, gage, etc.). The method of deriving these characteristics ranges from a heuristic method employing a well known probability distribution to sampling from an empirical probability distribution to generating samples from a multivariate probability distribution.

Two categories of order characteristics are generated by the method: 1) independent, such as due date, product type (i.e., sheet vs. strip), and material type, and 2) dependent, representing physical dimensions based on material type, such as gage, width, and length. Independent characteristics are those which can be determined without information concerning other characteristics, while dependent characteristics require knowledge of the values of other characteristics.

The first category, independent, may be determined by sampling from an empirical or a well known probability distribution. For example, in one application of this method the due date was determined to be between 25 and 50 days after order arrival. A Poisson random variable was used to create variability in the time lag between order arrival date and order due date. Specifically, the formula used to calculate the due date was:

$$\text{Due Date} = \text{Arrival Date} + 25 + P$$

where P is a sample from a Poisson distribution in the range [0,25].

The second category, dependent, determines the physical dimensions of the order: gage, width, weight, and length. All the physical properties of the order are determined via probability distributions based on the material type. A material type can be differentiated by size, weight, or type of material, such as alloy. These characteristics are extremely dependent upon each other, therefore a different method is employed to generate these characteristics. That method is described below.

For our example let us asume the weight of the order can be calculated using the gage, width, and length, along with the density of the material. Therefore, we need only generate three characteristics: gage, width, and length. In order to insure proper dependence (in this case dependence is measured by correlation), we generate samples from a trivariate probability distribution. The uniform transformation method was chosen as the method with which to generate these samples (Hull 1975; Kimeldorf and Sampson 1975; Kimeldorf and Sampson 1975). This method guarantees exact marginal distributions and correlations which may be close. The correlations are exact if the marginal distributions are normally distributed. As the marginal distributions move away from the normal there is a greater chance that the correlations will not be close. The marginal distributions used to represent the characteristics are empirical distributions derived from the actual orders. The correlations are also calculated from the actual order characteristics. Details of the uniform transformation method are discussed in Appendix A.

3. PROCEDURE VALIDATION

Once orders have been generated, the method must be validated to insure that the statistical properties of the orders have been retained. A portion of this validation process involves the examination of scatterplots, histograms, and tables. In addition, one must scan the individual orders to determine whether combinations of characteristics are logical.

Histograms of order characteristics (gage, width, length, etc.) by material type can be generated. These histograms can be used to compare the actual orders to the generated orders visually to insure that they have similar shapes and proportions of observations in each histogram cell. Figure 3 shows histograms of gage for actual and generated orders. Additionally, the average, standard deviation, and range statistics are computed for all orders within a given material type.

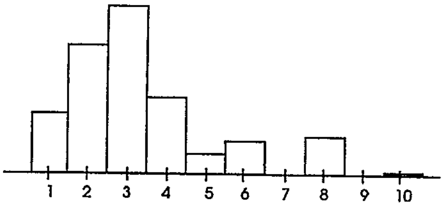
Scatterplots can be employed to check the validity of the uniform transformation method by plotting characteristics for both actual and generated orders of a given material type on the same grid. Shown in Figure 4 is a plot of gage versus width. The coordinates [(gage, width)=(X,Y)] represent order characteristics and can be compared on the scatterplot. If the order points are concentrated in the same areas on the plots for actual and generated orders, then we can conclude that the procedure generates realistic orders.

The method's accuracy is also checked via tables of statistical values derived form the actual and generated orders. Shown in Figure 5 is an example of a table containing order statistics. One can then compare the characteristics of all orders within a material type to test the method's accuracy.

Based on examination of the plots and statistics for sets of orders which have been generated using this method, it has performed quite well. The order characteristics, statistics, and correlations have been similar for actual and generated orders, which implies that the method is capable of generating representative orders.

4. SIMULATION STUDY OF A STEEL PRODUCTION FACILITY

The installation of a new 4-HI rolling mill for rolling raw stock into coils at a special alloy steel facility initiated a simulation study of the effects of the new equipment on production. Raw stock is fed into the 4-HI and rolled into coils (5,000-10,000 pounds). The coils, which have orders attached, are processed through rolling, annealing (furnace), and cutting operations. The result of those processes is completed customer orders, which are placed with stipulation on alloy, gage, weight, width, length, and other chemical and metallurgical properties. The flow of material through the facility is shown in Figure 6.



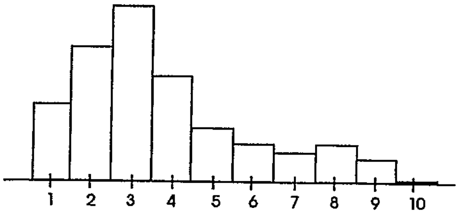
CELL STATISTICS

CELL NO.	OBS. FREQ	REL. FREQ	CUM. FREQ	UPPER BOUND
1	34	.1193	.1193	.0292
2	76	.2667	.386	.0434
3	95	.3333	.7193	.0576
4	41	.1439	.8632	.0718
5	10	.035	.8982	.086
6	13	.0456	.9439	.1002
7	0	0.	.9439	.1144
8	15	.0526	.9965	.1286
9	0	0.	.9965	.1428
10	1	.0035	1.	.157
	285			

CELL WIDTH = .0142

SAMPLE STATISTICS:

MEAN	.0516
STANDARD DEVIATION	.0247
MINIMUM VALUE	.015
MAXIMUM VALUE	.157



CELL STATISTICS

CELL NO.	OBS. FREQ	REL. FREQ	CUM. FREQ	UPPER BOUND
1	54	.1219	.1219	.028
2	99	.2235	.3454	.041
3	123	.2777	.623	.054
4	83	.1874	.8104	.067
5	26	.0586	.8691	.08
6	17	.0383	.9074	.093
7	13	.0293	.9368	.106
8	16	.0361	.9729	.119
9	11	.0248	.9977	.132
10	1	.0022	1.	.145
	443			

CELL WIDTH = .013

SAMPLE STATISTICS:

MEAN	.0531
STANDARD DEVIATION	.0252
MINIMUM VALUE	.015
MAXIMUM VALUE	.145

Figure 3. Comparison for Actual (top) and Generated (bottom) Orders

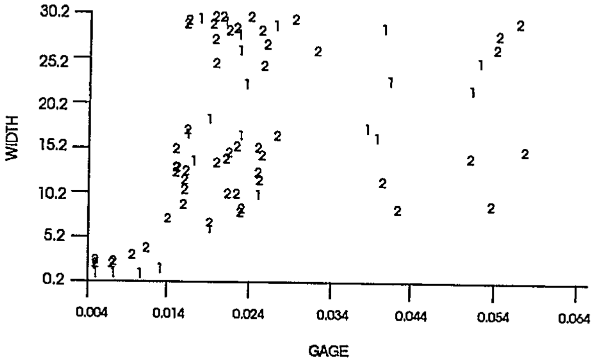


Figure 4. Scatterplot of Gage vs. Width (1=Actual Orders, 2=Generated Orders)

•••MATERIAL CODE 125•••

	ACTUAL ORDERS	GENERATED ORDERS
Number Of Orders	15	119
Fraction Of All Orders	0.020	0.020
Average Number Of Pieces	29.200	30.655
S.D. Of Number Of Pieces	119.930	285.070
Average Gage	0.070	0.069
S.D. Of Gage	0.021	0.053
Average Width	2.667	2.622
S.D. Of Width	4.309	9.526
Average Length	8.467	7.891
S.D. Of Length	32.957	85.775
Gage-width Correlation	-0.083	-0.069
Gage-length Correlation	0.145	0.051
Width-length Correlation	0.644	0.398

Figure 5. Table of Statistical Properties for Actual and Generated Order for a Material Code

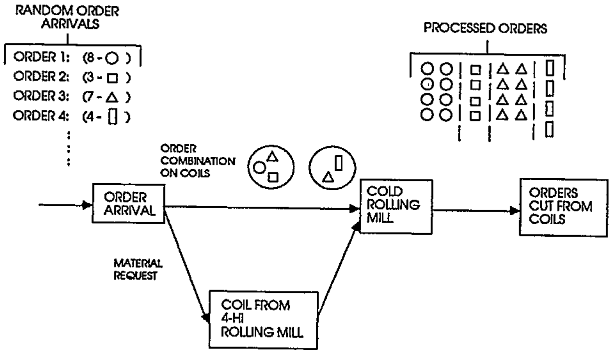


Figure 6. Material Flow Through Steel Facility

The purpose of the model was to assist in developing operating policies at the facility. The model was used to:

- Evaluate system capacity since the 4-HI increased raw stock input capacity.
- Identify bottlenecks due to increased work loads.
- Evaluate standard width policy (coil raw stock is cut into certain standard widths for use at the mill).
- Evaluate order provisioning rules (how orders are distributed onto coils).
- Evaluate order processing rules (the sequence in which coils are processed through various stages of the facility).
- Evaluate raw material requirements (the number of coils of each standard width required form the 4-HI to maintain smooth operation of the facility).

The model is driven by an order book which contains a list of orders to be processed by the facility. Once the model had been constructed the above evaluations were performed using a three month order book consisting of actual orders. The problem confronting the project team was that the order book at any point in time was not necessarily identical to that three month order book. In order to add credibility to the results, different order books should have been run through the simulation model, but no additional data was available. It was determined that the order generator could be applied in this case since the mix of orders in the actual three month order book was deemed representative of a *typical* order book. A schematic of the order generation system used is shown in Figure 7.

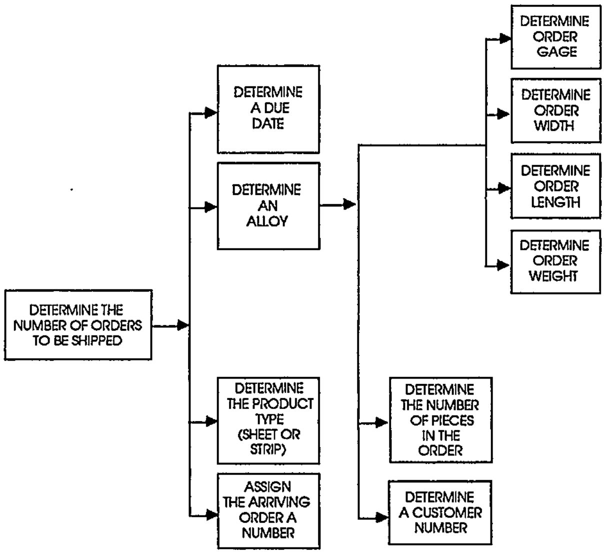


Figure 7. Steel Facility Order Generation Process

Several additional order books were generated and used by the model. The study results in a reduced number of standard widths, a detailed list of 4-HI production requirements, and investigation into the furnace since it was shown to be the bottleneck. Further evaluations of the facility are currently ongoing. The use of the order generator greatly expanded the usefulness and validity of the results from the model.

5. SIMULATION STUDY OF A SHEET METAL FABRICATION FACILITY

A critical function within aerospace production facilities is the fabrication of sheet metal parts. One of the more recent technological advancements has been the development of numerically controlled (NC) routers. These new routers are capable of producing more parts in less time and with less resultant scrap than was possible with manual operations. The primary advantage of these routers is their ability to process several orders simultaneously. The router function consists of routing and drilling in a numerically controlled path on a stack of metal sheets. Before processing can begin, however, the type and amount of material as well as the NC instructions must be specified.

The key activity associated with using the router is the process that selects which orders will be processed together. This process is referred to as nesting and the set of orders to be processed together is call a nest. The efficiency of the router hinges on the effectiveness of this nesting function which depends on two sets of information. The first set includes gage, material type, and number of parts required by the group of available orders. The second set includes the stack height and sheet size (length and width) that best utilize the material. Since the two sets of information are directly related to the characteristics of the orders, the nests built (and hence the router performance) are very sensitive to the number and types of orders in the system. A schematic of the router operations is shown in Figure 8.

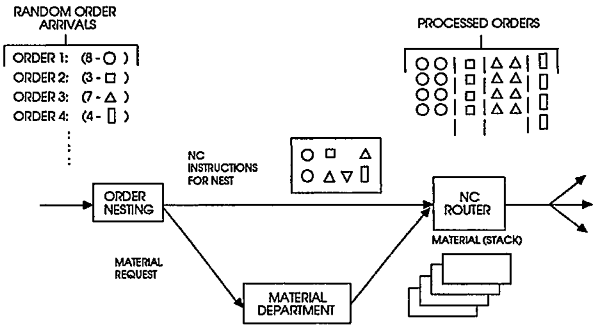


Figure 8. NC Router Operations Sequence

Although the introduction of these NC routers lowered production costs, the additional production complexities hampered the development of control strategies that best utilized the router capabilities. In an attempt to further reduce costs, a combined simulation-optimization model was developed to address the following (Pritsker & Associates 1983; Reilly *et al.*.):

- Manpower requirements for the NC router operations.

- NC router requirements for different levels of production activity.

- The effect on scrap production if the controls were modified so that the router could process more than one stack height.

Because of the nature of the nesting function and the long lead times for material, a model run length of six months was used. As a result, thousands of orders were needed for each execution of the simulation model. Unfortunately, only a fraction of the required number was available, and those had to be input manually. Although the number of orders was insufficient, the orders were representative. Therefore, to perform the evaluations, the previously described order generator was used to obtain the required number of orders. A schematic of the generation system used is shown in Figure 9.

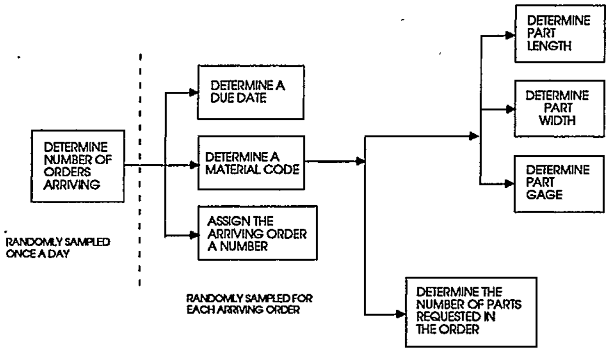


Figure 9. NC Router Order Generation Process

The results of the study indicated significant material scrap reductions could be obtained if variable stack height capabilities were added to the router. However, the variable stacks also resulted in nests having fewer orders per nest (on average) which caused router requirements to rise dramatically. In addition, the study also showed that the router itself was the bottleneck and that a second router would be necessary if the order volume increased significantly. Lastly, the study indicated that two of the manual operations (loading the stack onto the router and monitoring the router process) could be combined so that a single operator could perform both and not adversely affect system performance.

The ability to use the model as an analysis tool was due solely to the availability of the order generator. As is the case in many non-computerized production environments, the data required for simulation models is virtually impossible to obtain.

6. CONCLUSION

The order generator described in this paper is a technique that has been successfully utilized to generate orders having interdependent characteristics. The ability to generate orders for validation and experimentation, thereby freeing the analyst from actual data constraints, represents a major increase in the power of simulation as an analvsis tool.

As demonstrated by the two studies discussed in this paper, an order generator can be an invaluable aid when using simulation models to perform system analyses. For the steel facility study, the order generator allowed the modeler to evaluate and draw conclusions from several different model executions instead of only one. In effect, it freed the modeler from having to assume that the results from the single execution were representative. In the aerospace study the benefits were even more dramatic: *without the order generator no experimentation would have been possible.*

Although developed as an order generator for manufacturing applications, the method is not limited to generating orders for use as input to industrial simulation models. Any data that has interdependencies can be generated using this technique. For example, customer arrivals to a fast food restaurant, with orders associated with each customer, could be generated using this technique for purposes of evaluating staffing policies via a simulation model. Another example would be the arrival of patients to a hospital emergency room. In this case, ailments can be correlated with the age and sex of the patient for modeling purposes.

APPENDIX A: UNIFORM TRANSFORMATION METHOD

We now illustrate the Uniform Transformation Method. We want to generate a sample (with n characteristics) from a multivariate probability distribution $F(X_1,...,X_n)$ which has marginal cumulative distribution functions (CDFs) $F_1(X_1),...,F_n(X_n)$ and correlations $CORR(X_i,X_j)$; $i=1,...n$; $j=1,...n$.

Step 1

Generate a sample from a multivariate normal distribution (V_i; $i=1,...,n$) with correlations $CORR(V_i,V_j)$. (Note: $CORR(V_i,V_j)$ is estimated by $CORR(X_i,X_j)$.)

Step 2

Calculate (U_i; $i=1,...,n$) = ($F_i(V_i)$; $i=1,...,n$) which yields a multivariate uniform distribution. (Note: $F_i(V_i)$ are independent normal distributions.)

Step 3

Calculate (X_i; $i-1,...,n$) = ($F_i(-1)(U_i)$; $i-1,...,n$) which is the desired sample from $F(X_1,...,X_n)$.

REFERENCES

Hull, J.C., "Dealing with Dependence in Risk Simulations," *Operational Research Quarterly*, Vol. 28, No. 1, ii, 201-213 (1975).

Kimeldorf, George and Allan Sampson, "One Parameter Families of Bivariate Distributions with Fixed Marginals," *Communications in Statistics*, 4(3), 293-301 (1975).

Kimeldorf, George and Allan Sampson, "Uniform Representations of Bivariate Distributions," *Communications in Statistics*, 4(7), 617-627 (1975).

Law, Averill M. and W. David Kelton, *Simulation Modeling and Analysis,* New York: McGraw-Hill, Inc. (1982).

Martin, David L. and F. Bradley Armstrong, "Generation of Orders for Use as Input to Simulation Models," Presented at the Joint National Meeting, TIMS/ORSA, San Francisco, CA (May 15, 1984).

Pritsker, A. Alan B., *Introduction to Simulation and SLAM II,* Third Edition, Systems Publishing Corporation, West Lafayette, IN (1986).

Pritsker & Associates, Inc., "IDSS Build 1: Case Studies in Applications of Decision Support in Aerospace Industries, Vol. 4," ICAM Project Priority 8205 (1983).

Reilly, Charles H., F. Bradley Armstrong, and A. Alan B. Pritsker, "Solving a Trim Loss Problem in a Simulation Model: Exact and Heuristic Methods," Unpublished.

AUTHOR'S BIOGRAPHY

DAVID L. MARTIN is a Systems Consultant at Pritsker & Associates, Inc. (Pritsker). He holds a Bachelor of Arts in mathematics and economics from Hiram College and a Master of Science from the School of Industrial Engineering at Purdue University. Since joining Pritsker, Mr. Martin has been involved in a variety of projects including development of simulation models of a steel production facility and an offshore oil shuttle tanker system, development of simulation models to design, analyze, and schedule a variety of manufacturing systems, including FMSs, transfer lines, job shops, and AGVSs, and integration of simulation and order scheduling systems to analyze the effects of schedule modifications on manufacturing systems. Mr. Martin has written and presented several technical papers concerning simulation of manufacturing and mining systems. He is a member of Institute of Management Sciences (TIMS), Institute of Industrial Engineers (IIE), Society for Computer Simulation (SCS), and Society of Manufacturing Engineers (SME).

F. BRADLEY ARMSTRONG is a Staff Engineer at Hughes Aircraft Company, Long Beach, California. He received a B.S. in Mechanical Engineering from the University of Texas at Austin in 1981 and an M.S. in Industrial Engineering from Purdue University in 1986. He is currently an internal consultant providing manufacturing simulation consulting, training, and software support at Hughes. Prior to joining Hughes in 1986, he was a Senior Systems Analyst for Pritsker and Associates, Inc. (1982-1986), and an Operations Analyst for General Dynamics, Fort Worth, Texas (1981-1982). He is a professionally registered engineer in Indiana, a senior member of both SME and IIE, and a member of SCS.

David L. Martin
Pritsker & Associates, Inc.
1305 Cumberland Avenue
West Lafayette, IN 47906
(317)463-5557

F. Bradley Armstrong, IV
Hughes Aircraft Co.
P.O. Box 9399
Bldg. C-5, MS: 2016
Long Beach, CA 90810