

USING A SIMULATION TO GENERATE THE DATA TO BALANCE AN ASSEMBLY LINE

Mark R. Grabau

United States Air Force
14169 Compton Valley Way
Centreville, Virginia 20121-5003, U.S.A.

Ruth A. Maurer

Walden University
155 Fifth Avenue South
Minneapolis, Minnesota 55401, U.S.A.

Dennis P. Ott

Coors Brewing Company
9084 Holland Street
Westminster, Colorado 80021, U.S.A.

ABSTRACT

Coors Brewing Company's 16 ounce can production line was generating over \$2 million annually in scrap production. Flow scrap is generated when a machine stops because it is working faster or slower than its adjacent machines. The production line flow problem was that the machines were not working together in such a way to keep the flow of cans constant from the start to the finish of the line. This work demonstrates an analysis process that minimizes the generation of flow scrap in the context of the assembly line balancing problem. The process is 1) develop a verified and validated simulation model; 2) design an experiment to run using the simulation and to collect output data from the simulation; 3) using the experimental design to determine values of the independent variables and the output data as the dependent variable, develop a metamodel; and 4) optimize the metamodel using response surface methodology. Applying this process to the Coors problem generated an annual savings of \$1.87 million in 1996 dollars.

1 GENERAL IMPORTANCE

Many manufacturing systems today are automatic, and products are created at an amazing rate. The assembly line allows these products to be produced by a small number of people. In many instances the human factor is nothing more than a maintenance function, fixing problems when they arise. Otherwise, the machines run by themselves. These machines do not have the physical limitations that humans do and, therefore, can operate at significantly higher speeds than a human.

The goal of any assembly line is to continuously

generate throughput by keeping the flow of materials constant from start to finish. Generally, the back of the assembly line pulls processes in the front. More specifically, each machine pulls production from the machine immediately preceding itself. In some manufacturing situations, if a machine runs too fast or too slow, it will cause other machines to stop. When these machines stop, objects currently being manufactured may be scrapped because they fall off the assembly line. The items may be damaged when they fall; budgets do not permit additional personnel to be hired solely to place these objects back on the assembly line. The more objects that finish the production process, the greater the throughput. So a restatement of the goal becomes, "maximize throughput and meet production objectives by minimizing scrap."

The rapid rate at which the whole process is occurring, the interaction between machines, and different transition times between machines make it increasingly more difficult for a human being to make the correct decisions regarding how fast each machine should be working to continue the pulling process, while at the same time keeping scrap low and throughput at an acceptable level. Likewise, a casual observer cannot just observe the system and make these decisions. Many times the system must be intensely studied.

It is possible to design experiments to test a system by setting the machines at certain speeds and observing what happens. However, even this scientific approach to studying a manufacturing system may not be economically feasible. It is difficult to comprehend and anticipate the reaction of the system to certain experimental conditions on the spot. Requiring a human to make these decisions has the potential of

creating catastrophic problems just to observe one level of the experiment. In addition, this process can take an inordinate amount of time. These economic and time constraints preclude actual experiments being run on the system. A more economic and timely approach such as simulation is needed.

2 GENERAL SOLUTION

Once a simulation model is verified and validated, experiments can be run that do not have any direct impact on the system. The only costs are the time to develop the simulation and the computer resources to run the experiments.

Running the experiments correctly using the simulation should be cost-efficient. Management may need a timely response to their questions about the system under study which may preclude testing every possible design point. If the simulation is large and complex, running it at different design points, let alone replications at each design point, may be neither economically feasible nor possible in the time allotted. Experimental design is a field in and of itself, and much effort has been put into developing designs that are the best for certain situations. Even though an efficient design may reduce the number of runs, the amount of output to analyze may still be large.

Simulations have the potential of generating copious amounts of output. Once the experiments are run, a methodology is needed to analyze the results and synthesize them into a form that is presentable to the decision maker. As Arthur Geoffrion (1976) best put it, "The purpose of analysis is insight, not numbers." Metamodeling and response surface methodology are the analysis techniques that lead to the efficient use of simulation output for studying the assembly line optimization problem.

Metamodeling approximates a relationship between a dependent variable and one or more independent variables by using a mathematical function (Banks, Carson, and Nelson 1996, p. 514). Once a metamodel is developed, it can replace another model, e.g., a simulation or linear program. A metamodel allows the data to be combined in a manner that is manageable and provides the necessary insight to answer very specific questions about the system under study. Subsequently, once a system is metamodeled, there is no longer a need to run additional simulations to answer questions regarding the dependent variable, given different levels of the independent variables. These types of "what if" questions can be answered with a calculator. Response surface methodology can be used to optimize the metamodel and to obtain the values of the system input parameters, machine speeds

in this example.

A cost-effective method for studying rapid manufacturing systems was needed in order to make improvements to current operations. It was decided to use simulation modeling as a means to shed insights into the current problem, incorporating efficient experimental designs, metamodeling, and response surface methodology, to answer questions about the and determine the ideal operating conditions.

The general plan for the system analysis is a series of steps: 1) develop a verified and validated simulation model using the method outlined by Banks, Carson, and Nelson (1996); 2) design an experiment to run using the simulation and to collect output data from the simulation; 3) using the experimental design to determine values of the independent variables and the output data as the dependent variable, develop a metamodel; and 4) optimize the metamodel using response surface methodology.

3 GENERAL PROBLEM DESCRIPTION

Coors Brewing Company is in a cooperative partnership with Valley Metal Container that produces 16 ounce and 12 ounce aluminum cans. The 16 ounce can production line produces 200 million cans for Coors products annually. In addition, they produce cans having over 50 other additional labels for other beverage companies. The Coors Light ® label accounts for 140 million cans of the 200 million of Coors products. Because of the simple design of its label, this is the easiest and cheapest can to produce, and it accounts for 56% of the annual production. For this reason, this label is of specific importance to Coors. Since it is the easiest and most economical can to produce, it should represent the lower bound on scrap generation and the upper bound on production throughput.

In 1995, the 16 ounce can production line lost over \$2 million gross in production due to scrap. If a can falls down while on the production line, it becomes scrap since manpower constraints do not allow the positioning of personnel throughout the line to set the fallen cans up. The majority of the cans fall over when the can-pack throughout the line is not tight. While in tightly packed groups, the cans lend each other support to keep them from falling over. The authors' goal was to demonstrate how to make the stages of the production line work in concert in order to maintain a tight pack of cans throughout the line and thereby reduce the amount of annual scrap to at least an acceptable level.

4 STEPS OF THE SIMULATION STUDY

At the beginning of the study, the perception was that a flow scrap problem existed on the 16 ounce production line. The machines were not working together in a way that minimized flow scrap and kept a constant flow of cans from start to finish. The line was too complex to use heuristics or common sense, and actual experiments on the line were not economically feasible. Once built, the model was used to run experiments to determine optimal settings of the machines to minimize flow scrap.

4.1 Setting of Objectives and Overall Project Plan

The primary goal of the simulation project was to develop a realistic simulation of the 16 ounce production line. The model needed to be general enough to be easily extended to other labels, yet specific enough to answer the following questions: 1) what are the optimal settings of the machines to reduce flow scrap, and 2) where is the scrap being generated?

4.2 Model Conceptualization

Figure 1 contains a schematic of the 16 ounce production line. The Coors can manufacturing process uses the following stages:

- 1) A cup is formed from a sheet of aluminum at one of four cuppers.
- 2) The cup is stretched to the correct length and the rough top edge is trimmed at one of fifteen body makers. The cup is now a can.
- 3) The can is washed and dried in the washer to remove the grease from the first two stages.
- 4) A label is applied at the printer.
- 5) The paint is dried and the bottom is coated.
- 6) The inside of the can is coated at one of nine coaters.
- 7) The can is cured in an oven.
- 8) The top of the can is prepared for a cover at the necker/flanger.
- 9) The can is checked for holes and flaws at the tester.
- 10) The can is placed on one layer of a fourteen layer pallet at the palletizer.

Each of these steps occurs at a rate of 1400 to 1600 cans per minute. Scrap may be generated at any one of these ten steps. In order to focus on the controllable flow scrap generation, the authors divided scrap into 3 distinct categories: flow, break, and random.

If the printer or the coaters stop because there are not enough cans behind them or because there are too many cans in front of them, *flow scrap* is generated. The printer clears its wheels. Thermal energy rising in the oven knocks over the unsupported cans that are standing on the outer edges of the group.

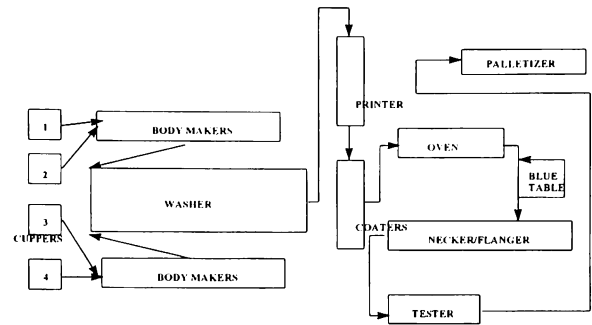


Figure 1: Coors 16 Ounce Production Line

Break scrap is caused by a printer problem or the necker/flanger jamming. *Random scrap* is generated in the washer, in the oven, and at the printer. Cans fall over in the washer when an edge of a group of cans is exposed to the sprayers, and cans fall over in the oven when an edge of a group of cans is exposed to thermal energy. Even if the washer and the oven are full, the outside edges of the flow of cans are still exposed. Random scrap at the printer is generated when a can is not properly seated on a wheel.

The reduction of flow scrap was the goal of this work. If the machines work together in such a way that there are no gaps in the line from start to finish, then the absolute amount of flow scrap will be minimized. The speed at which the line operates and the inability of any human being to comprehend all of the possible variables at once preclude a common sense approach based on observation and a subjective decision. This problem needed extensive study, yet economic constraints did not permit actual experiments to be performed on the line. A technique was needed that allowed the system to be studied in a non-invasive environment at no physical cost to the actual line.

4.3 Data Collection

A simulation model is only an empty shell without data to drive it. Based on the conceptualization, the following data were needed: travel times on the air tables and mechanized conveyors, scrap generation rates, time between machine breaks, and machine down times. BestFit (Palisade 1995) software package was used to determine which probability density function statistically best fits the collected data. SLAM II simulation language was used for this work. A second requirement imposed on the choice of density function chosen was that it must be supported by SLAM II (Pritsker 1996). It is assumed here that the reader has a basic familiarity with SLAM II.

4.4 Model Translation

An abstract representation of each stage of the production line was needed in order to make this complex problem manageable. Each stage of the can line is a separate area where distinct operations and queuing principles are applied. The conceptual model of Figure 1 was translated into the SLAM II simulation language.

In the model, cups are represented by entities. These same entities represent the cans when the cups are processed into cans. In order to keep the simulation time at a reasonable length, 1 entity represents 16 cups or cans. This simplification does not cause a loss of fidelity since several cups or cans normally travel together. All times are in seconds.

Each machine on the line is modeled as a resource. The cuppers, body makers, and coaters are aggregated into one resource. The speed of the resource is the speed of one machine times the number of machines running. The machine's air table or chutes act as queues. When an entity arrives at an air table, it waits in a first-in-first-out queue until the resource is available. Resources can also be preempted, simulating a broken machine.

A machine break-down is modeled as a separate entity arriving at a resource which moves to the front of the queue. The break entity then uses the resource until it is fixed. Resources can also balk and block.

If the body maker resource queue is full, then the entity balks to the track above the body makers where it cycles until the body maker queue has room. If the necker/flanger queue is full, then the entity balks to the blue table where it cycles until there is room in the necker/flanger queue.

If the printer queue, coater queue, tester queue, or palletizer queue gets too full, then it will block the preceding activity. This causes entities to stop moving through the system and to wait until the blocking queue has enough room to proceed.

Disjointed networks represent the control logic on the line. A diagnostic for each machine checks the queue in front of it to be sure it is not too full. The diagnostic also checks to make sure its own queue is not too low. If either of these conditions exist, an entity is sent to the resource and placed at the front of the queue. This entity then holds the resource until the queue length is at an acceptable level.

Several statistics are calculated throughout the simulation. The amount of scrap at each location is accumulated. At the palletizer, the time for an entity to move through the system is calculated. The interarrival time of pallets and the number of pallets are also accumulated. Since a label change is not an important

part of the study, the simulation is allowed to warm up for an hour of simulation time, and then these statistics are cleared. Statistics are then kept for another eight hours of simulation time.

4.5 Verification

SLAM II makes verification easy. The TRACE option prints the entire event list allowing the entity to be tracked through the system. Because of the size of the simulation and the number of entities being processed each second, the networks were verified one at a time. Each condition that causes a state change was tested. The queue levels that the diagnostics check were also reduced to smaller levels so the TRACE report could be kept to a reasonable size.

4.6 Validation

Data were collected based on the number of pallets over a four hour period. Validation was based on the number of pallets and uses the methodology in Banks, Carson, and Nelson (1996).

4.7 Experimental Design

The first part of the design of the experiment is to determine how many replications are necessary for statistical significance using the method outlined in Banks, Carson, and Nelson (1996). It was determined that five replications were needed for each run of this simulation.

At the beginning of each week, the production line begins empty. The first shift must fill the line with cans, and then the remainder of the week there are cans throughout the production line. This start-up condition must also be simulated. Statistics during this part of the simulation may negatively bias the final results since the line takes time to "warm up" and begin operating consistently in steady state.

Once the simulation reaches steady state, the statistics need to be reset to zero so unbiased statistics may be calculated. A machine break-down is a significant random shock to the system which may cause the simulation to never approach steady state; therefore, the simulation was run for a nine hour period where the machines were not allowed to break. Filling the line during the first hour is similar to a start-up after a weekend or after a label change. This system levels off at fifteen pallets per hour after the first hour. For this reason, the simulation was then designed to clear the statistics after a one hour warm-up period. Then statistics are calculated for one eight hour shift.

Theoretically, the machine speed should be able to

increase to a certain point where the system is overloaded and flow scrap gets worse instead of better. This effect can be modeled as a quadratic effect; therefore, the experiment was run at three levels for each machine in order to estimate this quadratic effect. Since adjacent machines interact with each other by checking the level of the queue in front of them, the design also accounts for interaction effects. For reasons external to this study, it was necessary to design the experiment in such a way that less than 48 hours of time on a Pentium 100 MHz processor was required.

The ultimate goal of the simulation was to estimate a metamodel and optimize the response surface. A design for this type of analysis is a Box-Behnken design (Schmidt and Launsby 1989, pp. 3-23). This type of design for six factors with three levels and five center point runs requires 54 runs. This is equivalent to 40 hours and 30 minutes, or less than 2 days on the same computer. This is definitely feasible and also models the needed linear, quadratic, and interaction effects. The Box-Behnken design also allows valid statistical analysis with as few runs as possible. Drawbacks of the Box-Behnken design are 1) corner points are not tested, and 2) there are enough runs to estimate all interactions and quadratic effects whether the analyst wants to or not (Schmidt and Launsby 1989, pp. 3-18). Since the goal of the study is to minimize flow scrap, the response for the experiment is the average number of cans per hour of flow scrap

4.8 Analysis

In order to determine if the metamodel and response surface optimization are beneficial, a baseline must be established. Table 1 shows the simulation results for observed settings of the machines during the data collection phase.

Table 1: Baseline Simulation Output

flow scrap/hour	219 cans
pallets/hour	13.1

The next step was use of the experimental data from the simulation to determine the metamodel and calculate the response surface. The first metamodel for the response surface includes linear, quadratic, and interaction terms for all six machines. Statistical Analysis Software (SAS) was used to calculate the parameters in the model, specifically, the procedure to estimate a response surface using ordinary least squares, SAS PROC RSREG. In summary, this model did a reasonable job of predicting flow scrap and could

be used as a proxy for the simulation to answer the "what if" questions about the impact of different machine speeds on flow scrap.

Using all of the coefficients in the first metamodel, the next step was to find the optimal combination of machine speeds that minimized this metamodel by using response surface methodology. This was done by taking partial derivatives of the first metamodel with respect to every variable and setting the result equal to zero. This system of equations was then solved to determine the optimal point of the metamodel.

Eigenvalues and eigenvectors are used in a multidimensional model to determine if the stationary point is a maximum, minimum, or a saddle point. If the eigenvalues are all positive, the stationary point is a minimum. If the eigenvalues are all negative, the stationary point is a maximum. If there are positive and negative eigenvalues, the stationary point is a saddle point. (Schmidt and Launsby 1989, pp. 5-15).

Since there were both positive and negative eigenvalues in the first metamodel, the stationary point was a saddle point. Additional experimentation was needed in order to locate the local optimum. Additionally, the printer and the body maker were very significant in the first metamodel. Therefore, additional experimentation using these two inputs while keeping the others constant was done in the hope of finding the minimum. Using SAS PROC RSREG, a gradient search along the path of steepest descent was performed to determine at which values to fix the other variables and where to center the two variables of interest. It was decided that the copper, coater, and necker should be set at a middle value, while the tester should be set at a high value. Finally, additional testing was done with the body maker and printer around their middle values. Since the problem was reduced to only two factor with three levels, a full factorial model was feasible. The next step was to use the experimental data to determine the final metamodel and calculate the response.

This model included all linear effects, quadratic effects, and interaction effects for these two factors. Using SAS PROC RSREG, the parameters were calculated. The second metamodel had much better results and may be used as a proxy for the simulation to answer the "what if" questions about different machine speeds, provided the copper, coater, and necker are at a middle value, and the tester is set at a high value. The second metamodel was also significantly less complicated than the first in regards to the number of terms. The next step was to find the optimal combination of machine speeds that minimizes this metamodel by using response surface methodology. Since both eigenvalues were positive, this would be a

minimum. Using the optimal settings, the first metamodel predicted 82 cans of flow scrap, the second metamodel predicted 128 cans of flow scrap, and the simulation estimated an average of 140 cans of flow scrap per hour while increasing to 13.4 pallets per hour. This is a 36% reduction in scrap while increasing the line's throughput and shows the improvement in predictability from the first to the second metamodel.

4.9 Documentation and Reporting

A formal report and briefing were given to the plant manager at Coors, who accepted the results enthusiastically. The credibility gained by the authors actually working the line and taking data, along with their intuitive understanding of the line, were extremely beneficial. The results were understandable and believable. The plant manager took the results for action and is ready for another study to be done on another label. He is interested in extending the study to the 12 ounce production lines as well.

4.10 Implementation

This study resulted in the reassignment of several personnel and machine speeds were changed to reflect the analytical results, e.g. 8 coaters are now run instead of 7. Several months after completing the study, Coors verified an annual savings of \$1.87 million in 1996 dollars.

5 SYNOPSIS

The most difficult portions of this simulation study were the data collection and the model translation steps. They were also the most time consuming. Over 200 hours were spent on data collection, and over 100 hours were spent on model translation, accounting for over 80% of the total project hours. The efficient experimental design was crucial to keeping the computer time to a minimum while producing the desired result--a balanced line.

Experimental design, metamodeling, and response surface methodology make an extremely powerful analysis toolbox. Solving a problem from this point of view significantly reduces the complexity found in this manufacturing system.

6 CONCLUSIONS

If the resources are available, metamodeling of a production line through simulation is extremely insightful. Not only does it point out the key inputs of the line, but it also provides a simple algebraic equation

for the response under study. The benefit of the metamodel is that anyone with a calculator can determine the level of the response given certain inputs. No knowledge of simulation, simulation languages, or computers is required. The metamodel is also significantly faster than running the simulation, a few seconds versus forty five minutes on a Pentium 100 MHz processor.

REFERENCES

- Banks, Jerry, John S. Carson III, and Barry L. Nelson. 1996. *Discrete-event system simulation*. Upper Saddle River: Prentice Hall.
- Geoffrion, Arthur M. 1976. The purpose of mathematical programming is insight, not numbers. *Interfaces* 7 (1): 81-92.
- Schmidt, Stephen R. and Robert G. Launsby. 1989. *Understanding industrial designed experiments*. Colorado Springs: Academy.
- Palisade. 1995. *BestFit user's guide*. Newfield: Palisade.
- Pritsker, Alan B. 1995. *Introduction to simulation and SLAM II*. New York: Wiley.

AUTHOR BIOGRAPHIES

MARK R. GRABAU earned MS degrees in Mineral Economics (Operations Research) and in Mathematical and Computer Sciences (Statistics) at the Colorado School of Mines (CSM) in 1997. His BS is in Operations Research from the United States Air Force Academy. He earned the CSM Operations Research Guild Diamond Stickpin award while a student at CSM. Currently, he is a mobility analyst at the Pentagon in Washington, D. C.

RUTH A. MAURER earned her doctorate in Mineral Economics, (Operations Research & Applied Statistics) at Colorado School of Mines in 1978. She holds a CSM Operations Research Guild Diamond Stickpin award, is Professor Emerita of Mathematical and Computer Sciences at CSM, has taught at United States Military Academy (West Point) and the University of Lulea in Sweden. Currently, Dr. Maurer is Chair, Applied Management and Decision Sciences Division, Walden University, Minneapolis, Minnesota.

DENNIS P. OTT earned his MS in Computer Science in 1988 from the University of Denver. Currently, he is working on his MS in Mathematical and Computer Sciences at the Colorado School of Mines and is a Quality Engineer at the Adolph Coors Company in Golden, Colorado.