Leveraging Reliability Data to Reduce Capital Expenditures

Carol Vesier
President
RonaMax, LLC
46 Decidedly Lane
Bear, DE 19701
Email: cvesier@ronamax.com
Phone: (302) 832 0666

Richard Pettigrew
Director of Corporate Asset Management
Rohm and Haas Co.
P.O. Box 547
Bristol, PA 19067
Email: richard-pettigrew@rohmhaas.com
Phone: (215) 785 7529

Abstract

Reliability data combined with Discrete Event Simulation offers a gold mine of opportunity to improve financial performance through lower capital expenditures, lower costs, and higher throughput. Many organizations are failing to take advantage of this opportunity because they are unfamiliar with the modeling tool and its data requirements. Many are afraid that developing a model fueled with sufficient data would be too costly or time consuming. In our experience, this is not true. To illustrate why this is a fallacy; this paper will review the minimum data requirements, data sources, and model function for reducing capital expenditures. Finally, this paper will review a case study that illustrates how a model developed with minimal data was leveraged into the knowledge needed to reduce a $14 Million project to $12 Million.

Introduction

Frequently, capital budget projects resemble the tale of Goldilocks and the Three Bears, some are too small, some are too large, and some are just right. Consistent design of the "just right" facility requires understanding the impact of unpredictability on facility performance as a function of design parameters and market conditions. There are two common reasons for failing to consistently design the "just right" facility. First is an unawareness or unfamiliarity with the tool that can enable a "just right" design, Discrete Event Simulation. Discrete Event Simulation predicts how failures[1] and other unpredictable events[2] impact system function[3] and behavior[4]. The second reason is the perception that developing a Discrete Event Simulation model and populating it with sufficient data are insurmountable obstacles. In our experience, this is not true. Developing a model capable of receiving input data is typically as quick as creating a process flow diagram with a CAD program. Input data is readily obtained from operator logs, production logs, inventory history and operator interview. The resulting models have had an accuracy of +/-5%. The first step to understanding how this is possible is familiarity with the technology, data requirements, and possible data sources.

What is Discrete Event Simulation and how does it work?

Discrete Event Simulation predicts the impact of events and operational strategy on throughput, inventory, and system response time. An event represents a state change. Examples of events include failures, changes in process yield, changes in product quality, returning equipment to service, filling a tank, emptying a tank, arrival of a customer order, starting a batch, or planned maintenance. A calendar

---

[1] Examples of failures include fouling of a heat exchanger, leaks, seal failure, and off-spec production.
[2] Examples of unpredictable events include bad weather, random customer orders, and product turnover.
[3] Examples of system function would be productive capacity or ability to make on spec product
[4] System behavior are the characteristics of the system as it performed its function such as inventory levels and response time

keeps track of all of the events. At the start of a simulation, the model defines the next occurrence for every possible event and posts the events to a calendar. Once all of the events have been posted to the calendar, the model selects the time step size that will advance the model's clock to the event that is closest on the time horizon.

The model creates events according to one of two types of user input: a schedule (e.g. shutdown schedule) or a probability distribution (e.g. Weibull distribution). To simulate events based on a probability distribution, Discrete Event Simulation uses a random number generator. In this section, the interaction between the random number generator and probability distributions will be explained. Although the explanation may seem complex, please remember that this complexity is hidden from the user inside the model.

The model uses the random number generator in calculating a variety of probabilistic variables. A probabilistic variable is a variable that is defined by a probability function. Examples of probabilistic variables include times between failures, times to return to service, times between orders, order sizes, process yields, batch sizes, and quality losses.

The model uses the random number generator the same way, regardless of the variable it is calculating. First, an event triggers the model to request a random number. The random number generator produces a number between 0 and 1. This number represents the cumulative probability for a specific value of that variable. The model then calculates this specific value using the user-defined probability distribution or condition. A probability distribution can be a Weibull function or a probability array. For example, let's say that the model contains a bearing that fails randomly with a mean time to failure of 3 years. When the bearing is returned to service, the model asks the random number generator to provide a number that represents the cumulative probability for the next time to failure. If the random number generator were to output 0.63, the operating time to failure would be set to 3 years.[5]

Other examples of how the model uses a random number generator are shown in Figures 1 and 2. Figure 1 explains how the random number generator is used to simulate infant mortality while Figure 2 shows how the generator simulates a randomly varying return to service period.

---

[5] The cumulative probability of failure for a randomly failing bearing with a mean time to failure of 3 years is

$$F(t) = 1 - \exp(-t/3 \text{ yrs})$$

Since $F(t=3 \text{ yrs}) = 1 - \exp(-3/3) = 0.63$, a cumulative probability of 0.63 corresponds to a time to failure of 3 operating years. Another way of looking at this is given a probability of 0.63, the time between failures corresponding to this probability is

$$\text{TIME\_BETWEEN\_FAILURES} = -\text{MTBF} * \text{Ln}(1-0.63)$$

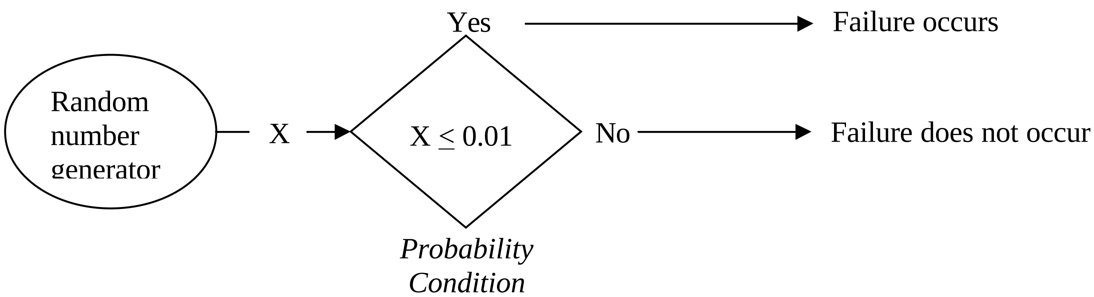

Figure 1. Defining if a specific printed circuit board will fail given an infant mortality rate of 0.01. In this case the model would ask the random number generator to produce a number (X) every time a board was initially returned to service. If X is less than or equal to 0.01 failure occurs; otherwise it does not. Boards not failing initially will result in the model recalculating time to next failure using another failure mode specified by the user.

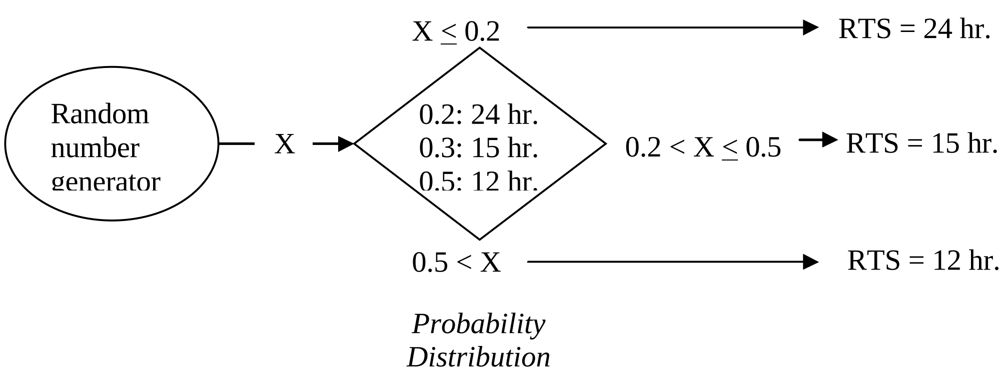

Figure 2. Defining return to service time (RTS) given a histogram of times. Return to service is the time lapse between loss and restoration of function. This time includes time spent bringing down the unit, time waiting to start repair, repair time, and start up time. In this case the model would ask the random number generator to produce a number (X) when a failure occurred. Each failure mode can have its own unique return to service times (probability distribution).

The amount of time before the occurrence of an event can also be a function of the other events. An example of this is shown in Figure 3.

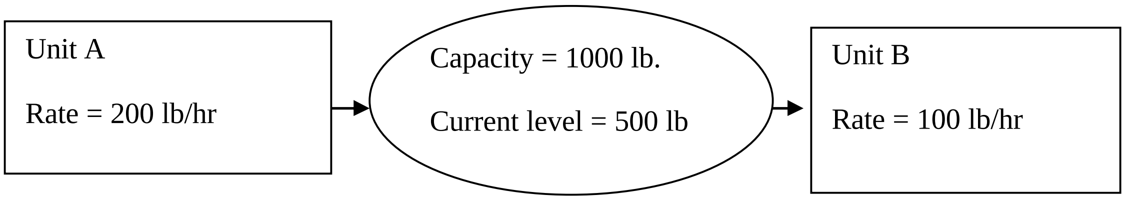

Figure 3. Failure in Unit B has just forced Unit B to drop rates to from 200 lb./hr. to 100 lb./hr. When the rate cut occurred, the tank between Unit A and B was at 500 lb. and Unit A rate was 200 lb./hr. The next event for the tank will be in 5 hours when the tank is full. If either Unit A or Unit B rate changes prior to filling the tank, the next event time for the tank will be recalculated based on a mass balance around the tank.

In summary, a Discrete Event Model marches forward in time, simulating the occurrence of failures and other events, such as filling a tank[6]. These events define the instantaneous system capability. As the model marches forward in time, material flows through the system at a rate based on the instantaneous system capability. The material may be a continuous stream or discrete batches. The material may also have attributes such as product code, yield, or age. As the material flows through the model, the model records data about the system such as the reasons for reduced rates, inventory levels, flow rates, and time for the material to pass through the system. This recorded data is used to generate model output.

What is the minimum data needed to develop a model?

There are three categories of data needed to develop a model.

1. Baseline operation: This includes a product-specific definition of process flows and unit capacities during optimal operation. For a continuous unit, the capacity is its maximum rate. For a batch unit, maximum batch size and shortest cycle time define capacity. For a buffer, capacity is its working or usable capacity.

2. Upset operation: This is a description of operational logic during upset conditions and includes

   - How feed will be shared during scarcity;
   - If and how rates will be cut based on tank level; and
   - Unit turndown

3. Upsets: This data defines the probability that all or part of a system will not be able to perform its function at highest efficiency and includes

   - Planned shutdown schedule;

---

[6] If realistic representation of your system requires modeling continuous flow through a tank, you should verify that the tool set allows modeling continuous flow. Many of the commercially available tool sets are unable to accurately handle continuous flow processes.

- Probability of a failure (failure rate, mean time between failures, etc) for a failure category;
- Consequences of a failure (shutdown, rate reduction, lost batch, off-spec material production, yield losses etc); and
- Duration of a failure (how long it takes to return to service). Failure duration may be explicitly defined with a single time (6 hours) or a probability distribution (50% of failures require 6 hours, 50% of failures require 12 hours). It may also be implicitly defined by describing what must occur for the unit to come up (repairs will require 6 hours once a mechanic is available)

For some industries, facility sizing is a function of order size, frequency, and lead time requirements. For these industries a fourth category of data is required that defines the customer-manufacturing interface. This category includes definition of order frequency, order size, order lead-time policies, and inventories management logic.

How is upset data obtained?

Typically, the most daunting task to developing a model is obtaining the data needed to describe operational upsets. We have found that we can develop very accurate models by using single parameter failure equations. This simplification greatly reduces the burden of gathering data.

Another key to obtaining data that will represent system performance is recognition that the data can not include any interaction effects. Interaction effects are starved and blocked losses resulting from operations represented in the model. These interaction losses are model output not model input.

An adequate picture of the frequency and severity of the upsets can be developed from many sources including

1. <u>Internal capacity loss database.</u> This is typically the best source of data for a model because it represents data typical for facility. It is also the rarest source of data. To be a standalone source of data, this database must account for all capacity losses such that a time and mass balance can be closed. In other words, a database representing a year of production must account for every instant and all lost production such that the sum of the actual production plus all lost production equals the time period multiplied by the maximum rate. If production losses occur because of upsets upstream or downstream of a unit, those losses must be denoted as a starved (upstream) loss or blocked (downstream) loss.

2. <u>External equipment reliability database</u>. Several databases are available that provide a range of failure rates associated with different types of equipment. Possible sources for this data include the vendor, Nuclear Plant Reliability Data System (NPRDS), Offshore Reliability Database (OREDA), Reliability Analysis Center, Government Industry Data Exchange Program (GIDAP), Center for Chemical Process Safety (CCPS), and RAMShipNet.

3. <u>Operations or production logs combined with inventory or demand history</u>. Most facilities record daily, hourly, or shift production as well as inventory levels. Production history will indicate if a capacity loss occurred. Comparison of production history with upstream and downstream inventory levels will indicate if the loss was the result of a unit upset or an interaction loss. The resulting data can be used to create model input.

4. <u>Site personnel interview</u>. Most sites have one or two people expert in the daily operation of a unit or facility. Through a guided interview, these experts can bracket the frequency of outages or slowdowns of varying severity. Their “gut-feels” can be fed to the model and the model will predict how the system as a whole behaves based on their “gut”. The experts can then adjust their estimates until the pieces of input create an accurate output.

The combination of site personnel interview with one or more of the other methods can produce a surprising accurate predication of facility performance. Combining these methods is analogous to building an air tight legal case of circumstantial evidence. The case is air tight because only one explanation is consistent with all of the evidence.

Case Study: Change Process Chemistry of Single Product, Semi-Continuous Facility

Market changes had created a need for a higher quality product in order to maintain market share. Meeting this need necessitated changing the process chemistry of an existing facility. Changing the chemistry required the installation of additional processing equipment. The initial project scope and equipment sizes were based on Theory of Constraints analysis of existing facility. This analysis resulted in a preliminary cost estimate of $14 Million. $2 Million of the investment was for expansion of the existing solvent recovery and batch reaction systems. The $2 Million expansion was included in the project scope because the business could not risk reducing production capacity as a result of changing the process chemistry.

Prior to proceeding with the $14 Million investment, the business decided to use a Discrete Event Simulation model to determine if they could reduce project cost by eliminating the $2 Million expansion. A model was developed to define the minimum unit sizes needed to guarantee no loss of production capacity. To develop the model a custom library of “blocks” were created using EXTEND, a commercially available simulation package. Blocks represent equipment that must operate together to perform a function, such as solvent extraction. These blocks are used to represent the manufacturing process by creating process flow diagram and model input. A process flow diagram is created in a manner similar to using a CAD program; the user copies blocks from a library and pastes them on a worksheet. The resulting model is shown in Figure 4.

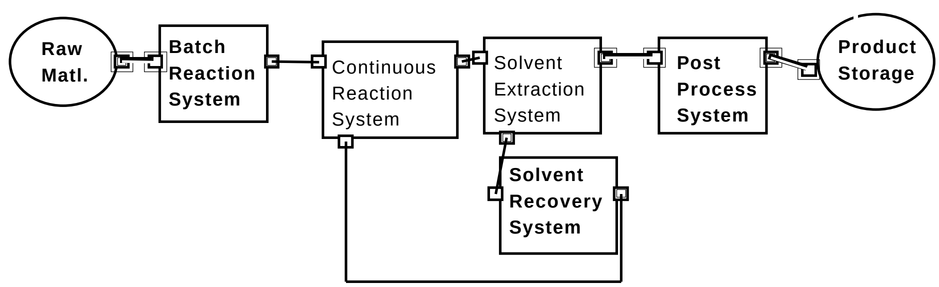

Figure 4. In this model of a semi-continuous process, a Batch Reaction System is fed raw materials. From the Batch Reaction System, the batch is transferred to a Continuous Reaction System along with solvent from the Solvent Recovery System. After flowing through the Continuous Reaction System, the Solvent Extraction System separates the solvent from the active ingredient. The solvent is sent to the Solvent Recovery System while the active ingredient goes to the Post Process System before Product Storage.

The next step was to gather and input the model data. This model only required two of the three categories of data: baseline operating data (operation in the absence of failures) and upset data (probability that all or part of a system will not be able to perform its function). Upset operating data was not required to simulate this process. The baseline operating data was obtained from process documentation. The upset data was obtained from multiple sources because formal data for the existing facility was limited. One of the unit operations (the original bottleneck) had a capacity loss database; however, the rest did not. This capacity loss data was supplemented primarily by interviewing operators, reviewing operator logs, and examining strip charts. Approximately one week on-site was required to assemble sufficient data from these three sources. The resulting reliability data for each block was entered into the model through a dialog box interface. This interface with example input and output data is shown in Figures 5a to 5e.

Number of Plan Shutdowns = 1

Rate During StartUp and ShutDown 24

Duration of Shutdown period (days) 1

Duration Of StartUp Period (days) 1

	Starting Day	Duration (day	Interval (Day
0	180	16	1460
1			

Figure 5a. Planned shutdown input defines the planned shutdown schedule that will be used in the model and can be defined for every block in the model. In this case, this unit's planned shutdown starts half-way into the first simulated year and includes one day to bring the unit down, 16 days of unit downtime, and one day to bring the unit up. The shutdown is repeated every 1460 days or 4 years.

All rates are in M Lbs/day

Maximum Rate = 40.00

Capacity Loss Data

Number of rows for unplanned downtime = 49

	Cause	Prob (#/hr)	Duration	Rate
14	Unknown	0.002212145	3	0
15	Unknown	0.002212145	3.25	30
16	Unknown	0.001171135	3.5	0

Figure 5b. Continuous unit input always includes the maximum rate and capacity loss data. The capacity loss data is input as a probability distribution describing the failures. Each event has its own independent failure probability and severity (Duration and rate). Assuming probability of failure does not change with time (uniform probability), the failure probability can be estimated by dividing the number of failures by the operating hours. Since these failures are all independent, it is possible for multiple failures to occur simultaneously.

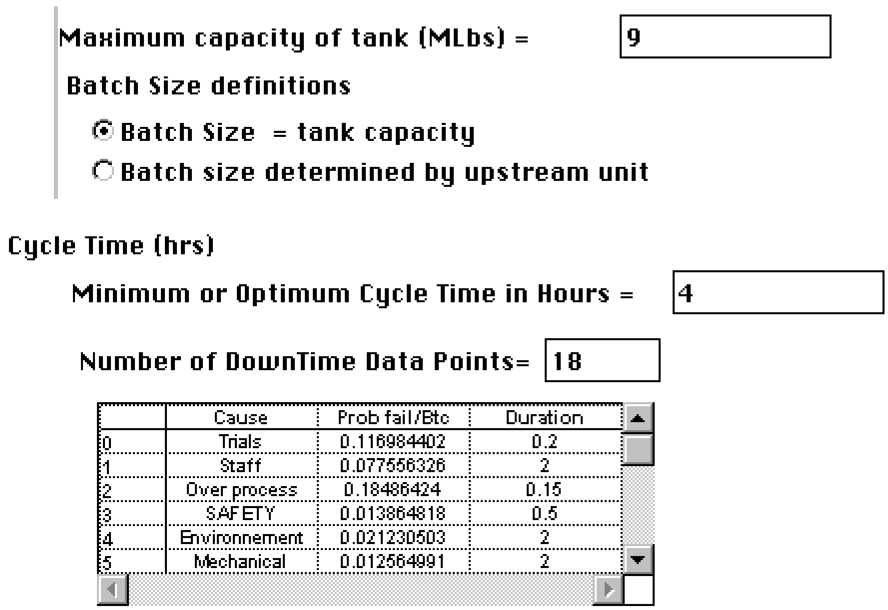

Figure 5c. Batch operation input includes a definition of the batch size, minimum cycle time and any cycle time delays. The cycle time delays are all independent failures. This means that it is possible to have a very long cycle time that is the sum of the minimum plus multiple failures. You will note that the failure probability is given as the probability per batch not per operating hour. Typically, a batch-based probability is easier to compute than a time-based probability because the number of batches is readily available while the number of operating hours is not.

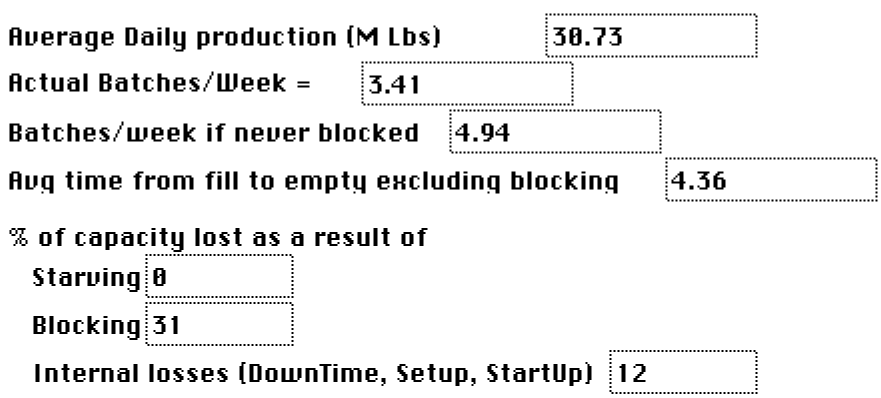

Figure 5d. Batch unit output also includes the average production rate and how capacity was lost. In this case additional output was added to facilitate the area operator is assessing the accuracy of their downtime estimates.

Daily Production (Mlbs)= 29.76

Percent of capacity lost due to

	Internal Losse	Starving	Blocking
0	10	15	0

Figure 5e. Continuous unit output always includes the average production rate and how capacity was lost. Additional output may also be added to indicate variability in production rates and capacity losses between simulated periods.

Once the model was complete with data, the next step was to use the model to determine whether or not the $2 Million capital expansion could be safely dropped from the project scope. To do this, two basic scenarios were simulated: a before scenario (before the process chemistry change) and an after scenario (after the process chemistry change). The before scenario was needed to validate the model by confirming that the model with its data would accurately represent historic performance. Once the model was validated, the next step was to determine system capacity if only the changes needed to implement the process chemistry change were implemented; i.e. without the $2 Million expansion. This second scenario established that the $2 Million expansion could be eliminated. If the model had predicted a drop in production capacity as a result of the process chemistry changes, the model would have been used to size the additional equipment.

Based on the insight gained from the model, the project scope was reduced from $14 Million to $12 Million by eliminating all costs associated with the expansion of the existing batch reaction and solvent recovery systems. The facility with its new process chemistry has been operational for about a year. The process chemistry changeover was accomplished without a loss of production capacity. ***In fact the model's predictions of production capacity after the changeover were within 5% of actual output.*** Today, the modeling tool developed for this project is still used by the business to develop cost-effective strategies for increasing production.

Conclusion

Reliability Data combined with Discrete Event Simulation can be leveraged into substantial capital cost savings. The task of assembling the data can be initially daunting, particularly if a capacity loss database is not available. Even in the absence of a capacity loss database, typically enough data exists amongst all sources to support design of the minimal cost facility. The payoff for projects that include reliability considerations in facility design is substantial.