Proceedings of the 1992 Winter Simulation Conference
ed. J. J. Swain, D. Goldsman, R. C. Crain, and J. R. Wilson

A COMBINED SIMULATION/OPTIMIZATION APPROACH TO PROCESS PLANT DESIGN

J. F. Faccenda
R. F. Tenga

E. I. Du Pont de Nemours & Company
Louviers Building
655 Paper Mill Road
Newark, DE 19714–7515

ABSTRACT

It is now critically important to understand all aspects of process plant design, because of market pressures to deliver high-quality, specialized products at lower cost, and with faster response to customer requests. Key design elements which are often overlooked include the operations and scheduling consequence of the design. In this paper, we present an approach to incorporating the key operational elements into the design of a process plant. The elements of the approach are described, along with an example application of the techniques.

1 INTRODUCTION

Process plants are generally very capital-intensive, and have traditionally been designed for high-volume production of a relatively few number of products. Thus, the primary design objective has been to minimize the investment required to produce a specified quantity of material. A number of things have occurred which substantially changed this picture. Customers are demanding a wider variety of more specialized products, with tighter quality specifications. The "JIT" (Just-In-Time) delivery philosophy which is rooted in discrete manufacturing has now taken hold in process manufacturing. Competitive pressures are driving the need to reduce production costs. And, it is being recognized that building large inventories is no longer a workable strategy to manage the business in light of these changes.

We describe a combined simulation/optimization approach which incorporates the process plant production operations into the design of a facility. This approach attempts to combine the best elements of a dynamic simulation with the benefits of optimization. Dynamic simulation is an excellent tool for predicting how a particular plant design will perform as different products are scheduled over the plant's resources to meet final product demands. The simulation model also allows incorporation of random events that occur in such operations. Analysis of such a system is normally accomplished through a series of simulation case runs, based off of an initial base design which was arrived at by static calculations. The problem with this approach is that it has the potential of overlooking some important design alternatives that were not considered, because of the inability to consider dynamic operations and scheduling in the static design calculations. Coupling the simulation model with an optimizer allows key design parameters to vary as independent variables in the optimization. The optimizer can then drive these variables to values which satisfy a specified objective.

Our approach consists of three key elements, each of which will be described in more detail. The first is the underlying simulation model. The simulation must be capable of representing the key process plant operations, interacting with a variety of resources over a range of product types. The simulation model must also be capable of reporting on the performance of a plant design with respect to the important business metrics used to evaluate a design, such as overall net present value, customer response time, and flexibility. It is also necessary that the simulation model be able to easily interface with the optimization algorithms that will be applied.

The second key element is the optimizer. We are assuming that the optimization objective function can only be observed by running a dynamic simulation model of the process. The optimization problem cannot, therefore, be handled by traditional analytical methods. The "function" generated by a simulation will contain a mixture of discrete and continuous parameters, with highly non-linear responses. An optimization algorithm which has proven to be capable of dealing with this type of problem is an adaptive search procedure known as the "genetic algorithm". The genetic algorithm is linked to the simulation to provide the capability of optimizing the simulated system.

The third element is the user interface, which is important for several reasons. First, it is the vehicle through which product, process, and facility designers are able to manipulate the numerous variables of interest. It is also key to aiding the interpretation of the large amounts of data which are generated by the analysis.

2 SIMULATION MODULE

The design objective for the simulation module was to establish a standard generic interface for modeling and subsequent optimization of process plant designs. Another objective was that the simulator run on a wide variety of computer platforms. In order to reach the widest possible user-base, the simulator was designed for ease of use, including automatic generation of statistics and reports.

The simulation design objectives were met by developing a set of "simulator" blocks within the SIMSCRIPT II.5® general purpose simulation language. This approach has several advantages. Coding requirements are limited, since the basic model construction consists of data-driven descriptions of equipment, combined with simulator block logic. However, for special modeling situations, a simulator block is included which allows interfacing to custom-coded SIMSCRIPT routines.

The simulator is based around a process-oriented world view. The primary modeling construct is a path which describes how a product utilizes resources, as well as how it interacts with other products. A product can be characterized with attribute values, and each individual batch may also carry it's own unique set of attributes. Named global variables are included for use in coordinating product flows, representing semi-finished inventories, miscellaneous calculations, as well as other similar functions. A unique feature which was incorporated to facilitate modeling process operations is the ability to pass a resource to a cloned batch in a disassembly operation. This often proves to be useful because of the large number of operations which combine with or spawn off other operations.

The standard simulator interface also allows pre-defined connection to optimization objective function calculations. Process metrics that are typically used in process plant design are automatically computed as part of the simulation run. These metrics include financial metrics such as net present value, as well as non-financial metrics such as product flow cycle times.

3 OPTIMIZATION MODULE

3.1 Characteristics of Optimization Problem

Generally, the economic optimization of plant design and operation parameters is a discrete, non-linear problem which may also be constrained in many different ways. Examples of the parameters that one might wish to optimize against include: number of machines, machine sizes, machine rates, machine and/or storage capacity, amount of labor, quality, reliability, and process control limits. Some of these parameters, such as the number of machines, have a fairly straightforward relationship to cost and production. Others such as machine rate and quality require a functional definition of the relationship to cost. For example, in some cases a machine which runs twice as fast as another may cost more than twice as much, while in another case there may be some other functional relationship between cost and speed. These relationships needs to be precisely defined for the optimization.

In addition, one might wish to constrain variables such as inventory levels, response time, or investment. It may be desirable to build constraints on variables such as inventory levels into the simulation itself, to explicitly account for the operational consequences of this type of policy decision. However, many other such constraints are only reasonably handled using penalty functions.

Other important parameters may represent operating rules, scheduling logic, or process control limits, and may not correspond to physical objects. These parameters cannot be ignored because they may have a large impact on capacity, cycle time, or other metrics which are vital to plant operations. They often have no cost associated with them, but they may need to be included in optimization to gain the best global understanding of combined equipment and operational consequences.

Because of the mix of discrete, continuous and possibly even qualitative variables, there is no rigorous algorithmic approach to solving these problems which is based on mathematical programming techniques. Most of the problems will be highly non-linear and contain many local optima. The best alternatives available are experimental design, simulated annealing, genetic algorithms, or other adaptive search techniques. All of these methods are based on sequential sampling of the objective function to search for an optimum solution. The method selected for this implementation is a genetic algorithm, chosen because of the flexibility to apply the algorithm to virtually any type of formulation without modification of the algorithm itself.

3.2 Genetic Algorithm Description

The genetic algorithm is an adaptive search procedure based on a model of population genetics, described by Goldberg and Samtani (1986), and also by Grefenstette and Fitzpatrick (1985). The basic idea is as follows. Candidate solutions (individuals) are represented by their parameter values (genetic code). The fittest individuals are those with the better objective function values. A population of individuals is maintained from iteration to iteration. Natural selection is used to choose which individuals survive and breed in future generations. Individuals pass along some of their genetic code to others by means of crossing over, i.e. the genetic codes of two individuals are broken at a random point and paired with the complementary half of the other individual. New individuals are also created via mutation, where bits of their genetic code are changed from 0 to 1 or vice versa.

The implementation of the genetic algorithm requires the specification of several control parameters. The population size is the number of individuals present in each generation. The percent of the population that is crossed over and mutation probability are examples of two others. The performance of the algorithm can be significantly different for different values of these parameters. For example, picking a small population size might not include enough diversity in individuals and the population might stagnate prematurely. Conversely, picking a very large population size can consume a lot of time in evaluating more individuals than are needed.

Another important parameter is the number of generations. This is the only stopping criterion besides manual intervention, so it must be chosen large enough to allow time for the algorithm to find a good solution. The population is saved at the end of every generation so it is possible to resume the algorithm from any stopping point if further optimization is needed. Additional implementation details are described by Rhoads (1986).

3.3 Interface to Simulator

In order to use the genetic algorithm (GA) in our environment, the simulator must be invoked to evaluate the objective function value of each individual. A degree of independence is maintained between the simulator and the optimizer by restricting communication between the GA and the simulator to external file data exchange.

The GA execution is keyed from a file which defines the parameters for the algorithm. It also contains parameter names, ranges and increments for the simulation parameters. Each individual generated by the GA has parameter values within the specified ranges.

When the GA generates a new individual it writes the parameter values of the individual to a communication file, and then spawns a new process which invokes the simulator to evaluate the individual. The simulator works from this file in combination with auxiliary calculations of dependent variables.

At the completion of the simulation, the user-specified objective function is evaluated. The objective function value is written back to the communication file, where it is read by the GA. The GA could potentially regenerate an individual which it has generated before (i.e. the same genetic code). If this occurs the old objective function value is looked up and the individual is not reevaluated.

4 USER INTERFACE

The user interface is designed to allow easy interaction with all of the data elements in the analysis, on both the input and the output sides. The interface is based on a spreadsheet model, with GUI (Graphical User Interface) input screens created using a hyper-scripting language. This allows a set of standard input screens to be easily combined with application-specific screens that can be created relatively quickly. The simulator/optimizer output can be analyzed via time plots, histograms, and a several types of business graphics. Animated output is created via communication files output by the simulation.

5 EXAMPLE 1: NEW BATCH CHEMICAL PROCESSING FACILITY

5.1 Problem Description

In this example, we review the design of a new batch chemical processing facility. The primary areas of interest include two types of finished product production, and two types of intermediate production. A third type of intermediate production was fixed in this case, and is not considered in the design problem.

The basic process flows are shown in Figure 1. Each of the major processing areas consist of some number of processing units which are capable of producing one class of product. Each product class will include a variety of product types. A given product type may or may not be able to be produced on a specific processing unit, depending on the design decision for that unit. In addition to the primary processing facilities shown, additional resources will be required at each production stage, depending on the product type.

The materials that are involved in the production of any particular intermediate or final product batch are dependent on the product type. An intermediate will typically be used in several different types of other intermediate and final products. Raw materials may be introduced at any production stage, depending on the product being produced. Intermediate storage is permitted between each production stage and after final product production, but most products have a limited shelf-life.

The processing of a batch of material is made up of several individual processing operations. The time to process a batch of a specific product type depends on the basic design of processing unit, the size of the unit, and the product itself. The processing time also contains a random component. At different points in the production process, the batch is tested for several key properties. If these properties are not within specified limits, the batch must be adjusted by adding materials and then repeating one or more processing operations. These factors also effect the use of additional supporting resources, such as the testing equipment and the labor force.

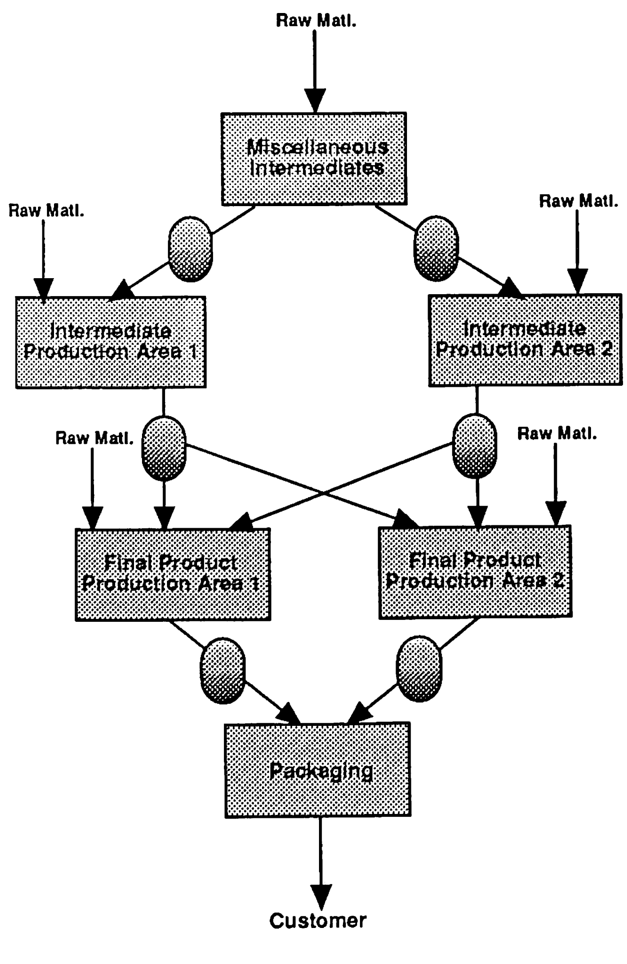

Figure 1: Process Flow

Changeover times are sequence-dependent, i.e., they depend on both the product that was previously produced as well as the new product that is about to be produced. Finished product is packaged into a variety of different-sized containers, specified by the customer. The combination of the finished product batch size, and the customer container size will affect the time to package the material as well as the material losses incurred in the packaging operation. The material losses are of particular importance here, as well as at other production stages, since the materials are generally very expensive.

The primary design question is: "How many processing units of a given type and size should be installed in each processing area?" The overall objective of the business is to maximize the net present value of the facility over a specified time horizon. There are also a number of other explicit or implied objectives and constraints, including:

- The facility must be able to produce the forecasted demand, including volume and product mix. It should also be flexible enough to produce a reasonable range of different volumes and product mixes, because of the inability to accurately forecast future sales.
- The facility must be responsive to customer demands. This can be interpreted to mean that given an overall demand scenario with a random component, a specified percentage of customer orders should be produced within a given time. The maximum order production time may also be specified.
- Although the net present value objective captures the important financial aspects of the decision, there may also be other constraints imposed on the total investment capital, as well as the working capital tied up in inventory.

The facility design objectives and constraints are clearly tied to how the facility is scheduled and operated. For example, given a finished product demand at some point in time, there is a question of what final product batch size to produce. This will depend in part on the design decisions made around equipment size, and will affect the changeover times, the processing times, the inventory levels, the material losses, and the customer delivery times. This decision also cascades back to the intermediate production batch size decision, and is complicated by the fact that the intermediates are generally also used in other products.

5.2 Solution Methodology

The simulation model constructed for this facility contains definitions of each type of processing unit being considered, at a range of sizes. The user interface allows all of the information that is specific to a processing unit type to be accessible through one input screen. The products at all levels are defined by the processing operations that they require. The processing times are generally nonlinear functions of equipment size, and may also contain logical references to the equipment type. Furthermore, these processing time functions may contain embedded references to random step functions. The definition of a product type is also accessible through one input screen.

The final product demand is input as a volume percentage for each product and container type combination. The overall volume can be controlled independently, so that the product mix and total volume sensitivities can be varied separately or together.

The production scheduling logic is rule-based. It considers product attributes such as monthly demand volume, equipment attributes such as size, product to equipment relationships such as manufacturing capability, and product to product relationships such as changeover time.

The primary optimization variables in this analysis are the numbers of each size and type of equipment to put into the facility. The structure also allows some switching of control logic by the optimizer. An example of this is a logical variable specifying whether or not to dedicate a given piece of equipment to a specific product type. There are options for the choice of objective function. Constraints are expressed as penalty functions.

5.3 Analysis Process

The summary output of an optimization analysis run is a facility design specified in terms of the numbers of each size and type of equipment. The robustness of this solution depends on how the objective function and constraints were expressed. For example, some of the goal constrains such as customer response time may be specified tighter than required, to allow for flexibility across potential product mix variations. In any case, the optimum solution is only valid for the combination of assumptions and constraints specified.

The nature of the genetic algorithm optimization process produces additional information that we can take advantage of. We save key information from each function evaluation, and then transform this information for use in sensitivity analysis. Figure 2 is an example of how average production cycle time varies with different total numbers of processing units, across a range of feasible solutions. This plot can be interpreted in several ways. One interpretation is an illustration of the design vulnerability to product-mix changes, as the design moves towards a fewer number of more specialized equipment.

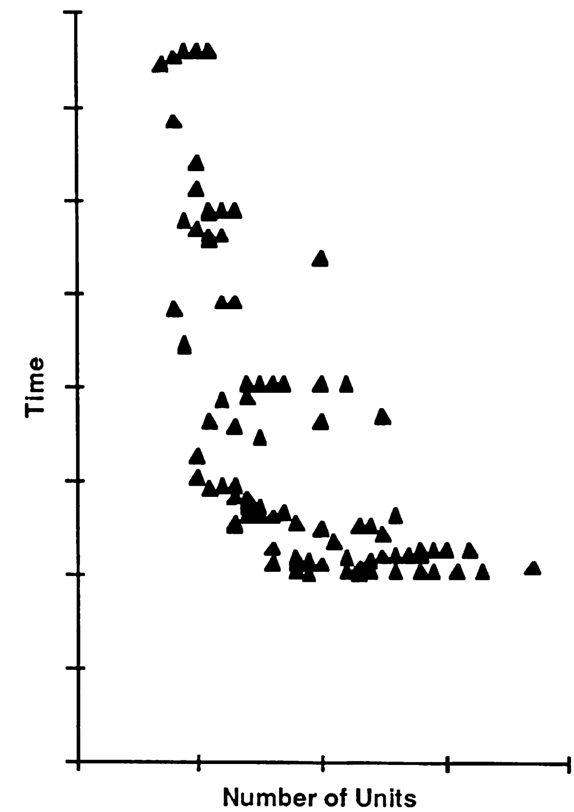

Figure 2: Production Cycle Time vs. Number of Units

Several types of post-optimization analyses are used to verify and fine-tune the design specifications. The robustness of the solution is studied by re-running the simulation with different random number seeds, and different product demand assumptions. Another area of analysis in this example is around the operation time estimates, since several of the operations rely on new processing equipment which does not have a history of commercial production utilization.

Detailed statistics are reported through a series of graphical views, through which the user can navigate. These views vary across different financial pictures of the facility performance over time, down through detailed information on delivery times of specific products to different customers at given times of the year. An ani mated view of the facility is also available, which is useful in communicating the dynamics of the facility, including the consequences of the operational and scheduling logic.

6 SUMMARY

The combined simulation/optimization approach described here has proven to be a valuable tool in the design of process plant facilities. We feel that there is a great deal of potential in this relatively new area, as we gain additional experience. The feasibility of this type of analysis is also being greatly aided by continued algorithmic development, together with the continued explosion of fast, low-cost computers.

REFERENCES

Goldberg, D. E. and M. P. Samtani. 1986. "Electronic Optimization via Genetic Algorithm." In *Proceedings of the Ninth Conference on Electronic Computation*, 471.

Grefenstette, J.J. and J.M. Fitzpatrick. 1985. "Genetic Search with Approximate Function Evaluations." In *Proceedings of an International Conference on Genetic Algorithms and Their Applications*, (Pittsburgh, PA, July 24–26). 112–120/

Rhoads, D.G. 1986. "Genetic Algorithm Project." Unpublished Technical Report. David G. Rhoads Associates, Inc., Wilmington, DE. (Dec.)

AUTHOR BIOGRAPHIES

JOE FACCENDA is a Senior Consulting Engineer in the Operations Research Group of the Du Pont Company, and is located in Newark, Delaware. Joe has over 12 years of experience consulting with internal Du Pont businesses in the areas of operations analysis, optimization, and simulation. He has presented numerous papers on applications including facilities design, energy management, and production scheduling.

BOB TENGA has been with the Operations Research Group of Du Pont for the past eight years. He has worked on problems of plant design and optimization. Prior to Du Pont, Bob worked in the Applied Mathematics Group of the Navel Weapons Center in China Lake, CA.