

Industrial Strength Analytics:

A Brief Overview of SAS/OR®

Garland D. Maddox, Jr., Regions Bank, Birmingham, AL

ABSTRACT

SAS/OR® has been a part of the SAS System since 1983. However, over the last several years, it has been transformed with the addition of a number of new features that include interactive model building, its own modeling language, and state-of-the-art optimization routines. This paper presents the basic categories of **SAS/OR** procedures and facilities with a closer look at the new SAS Simulation Studio® and PROC OPTMODEL.

INTRODUCTION

I have observed over the years that the SAS® System has such a large scope that a user could become practically an expert in one part of the system and know very little about other parts. My purpose here is to provide a very basic and brief outline of one of the less well known SAS analytical facilities - **SAS/OR**. Even though the concepts and routines of **SAS/OR** are generally thought to be in the domain of career operations researchers in much the same way that **SAS/STAT®** is associated with statisticians, this analytical area has been enjoying increased interest. Ever-increasing computer power and the natural front-end interface facilities of the internet have certainly made a large contribution. The "OR" in the name, of course, refers to "Operations Research". A number of good definitions for Operations Research exist, but the definition from one of the better known OR textbooks of the 1950's, defined it as "OR in the most general sense can be characterized as the application of scientific methods, techniques, and tools to problems involving the operations of systems so as to provide those in control of the operations with optimum solutions to the problems." (Churchman, Ackoff, and Arnoff, 1957) So, in essence, OR methods are designed to help decision-makers make optimum decisions through the application of scientific methods and tools, in particular modeling and experimentation. **SAS/OR** provides a number of procedures to enable the building and solution of a wide variety of models to help decision makers make optimum decisions. The procedures generally fall into one of three categories which we will explore separately in the following sections.

BASIC CATEGORIES OF SAS/OR

There are a number of types of decision problems that are addressed in the field of operations research. **SAS/OR** includes procedures and facilities to model and solve most of these types. There are three basic categories of **SAS/OR** procedures. They are:

Mathematical Optimization

Project and Resource Scheduling

Discrete Event Simulation

We will explore these categories individually, look at some of the types of decision problems that these procedures address, and how SAS has implemented its solution software.

MATHEMATICAL OPTIMIZATION

The optimization routines in **SAS/OR** can be applied in the formulation and solution of a wide array of OR problems. These routines are used to build mathematical models that describe basic characteristics of systems and are used to develop optimum solutions in areas such as resource allocation, production planning, distribution, and scheduling. There are basically three elements of any optimization problem:

Decision variables – a quantified action or choice, such as production levels or capital allocations.

An Objective – that which is to be achieved, such as maximizing profit or minimizing travel distance.

Constraints – limiting factors in the pursuit of the goal, such as machine capacity and budgetary restrictions.

The type of mathematical expressions that are developed can be categorized into several groupings depending on how the decision variables, objective, and constraints are constructed.

Linear programming is a widely known area of OR. LP models have only linear expressions in the constraints, and the objective function is assumed to be linear. Decision variables can be any value within the allowable range. Other types of optimization models include:

Mixed integer and Integer programs
Nonlinear programs
Quadratic Programs.

The solution methodology for each type of model differs, but it generally is based on an algorithm where there is an initial solution and subsequent iterations to improve the value of the objective function. When the solution can no longer be improved, then it is deemed to be optimal. If a solution cannot be reached then the methodology will detect it.

SAS/OR includes a number of PROC's that have been included either from the beginning or for a long period of time, including:

PROC LP
PROC INTPOINT
PROC NETFLOW
PROC NLP

These procedures basically require that the DATA step be used to build a SAS data set with a specific structure to describe the optimization problem. The data set is then input into the PROC to develop the solution. PROC LP is solved using the Simplex Method, and there are no constraints as to the size of the problem. A large scale LP problem can have many decision variables and a large number of constraints. These PROC's have been used widely and will continue to be a part of **SAS/OR**.

One of the many recent changes to **SAS/OR** is the addition of the "OPT" family of optimization routines. The OPT procedures incorporate new state-of-the-art optimization routines or "solvers". The focal point of this new set of PROC's is PROC OPTMODEL. This PROC incorporates a new optimization modeling language that makes it much easier to construct a model. PROC OPTMODEL also provides access to all new optimization solvers. The other OPT PROCs are:

PROC OPTLP – specifically designed for Linear Programming
PROC OPTQP – quadratic programming
PROC OPTMILP – mixed integer LP
PROC GA – Genetic routines
PROC CLP – Constraint programming, which is still experimental.

Consider the following linear programming formulation (Thierauf and Grosse, 1970):

Maximize $Z = 10A + 12B$ (Objective Function)
Subject to:
 $2A + 3B \leq 1500$
 $3A + 2B \leq 1500$
 $A + B \leq 600$
 $A \geq 0$
 $B \geq 0$

In this case, for example, A and B could refer to two products and we would like to maximize our profit based on the constraints of various machine capacities. The intuitive approach would lead one to concentrate on producing more of product B, since it is the most profitable. But, the constraints have to be considered in optimizing the product mix.

The PROC OPTMODEL code to solve this particular problem would be as follows:

```
proc optmodel;

    /* declare variables */

    var a, b;

    /* maximize objective function (profit) */

    maximize profit = 10*a + 12*b;

    /* subject to constraints */

    con process1: 2*A + 3*B    <= 1500;
    con process2: 3*A + 2*B    <= 1500;
    con process3: A + B <= 600;

    /* solve LP using primal simplex solver */

    solve with lp / solver = primal_spx;

    /* display solution */

    print a b;
quit;
```

When executed, the Primal Simplex solver is invoked and the solution is calculated. In this case, the optimal values are A=300 and B=300 which yields a Z value of 6600. This is the optimal solution for this problem. The modeling language closely follows the mathematical notation of the problem and is much easier to set up and maintain than the older optimization PROCs. It is also possible to load models into the OPT PROCs through input data sets. A new feature in **SAS/OR** is that the OPT PROCs use industry standard input data set formats. This significantly helps users who are already familiar with optimization routines but not with SAS. The MPS format is used for linear and mixed integer linear optimization and QPS is used for quadratic optimization. MPS is an old but widely used format that is column oriented and all model components receive names. To facilitate migration from the older PROC's, the MPSOUT= option now available will stop the optimization solver and write out the problem in MPS format. The subsequent MPS formatted data set can then be used as input to one of the OPT PROC's.

PROJECT AND RESOURCE SCHEDULING

There are a number of PROC's and facilities that fall into this area of **SAS/OR**. Most of these PROC's and facilities are designed to aid in managing large projects. An important feature of project management methodology is the establishment of the **critical path** through a directed network of tasks. This methodology enables a project manager to know which tasks must be expedited in order to reduce the overall span of the project. Project management PROC's include:

PROC CPM
PROC PM
PROC GANTT
PROC NETDRAW .

The project management PROC's provide detailed project schedules and graphics such as Gantt charts and network diagrams. In addition, the PROJMAN application is a long-standing SAS/AF® application which provides a user-friendly interface to the capabilities of PROC PM. Licensees of SAS/OR can access PROJMAN by clicking on Solutions ->Analysis -> Project Management.

Decision analysis under uncertainty is the focus of PROC DTREE. It allows a user to examine and compare all possible outcomes in a sequential decision problem and to identify an optimal strategy. Input data sets provide the problem structure and the probabilities of various outcomes.

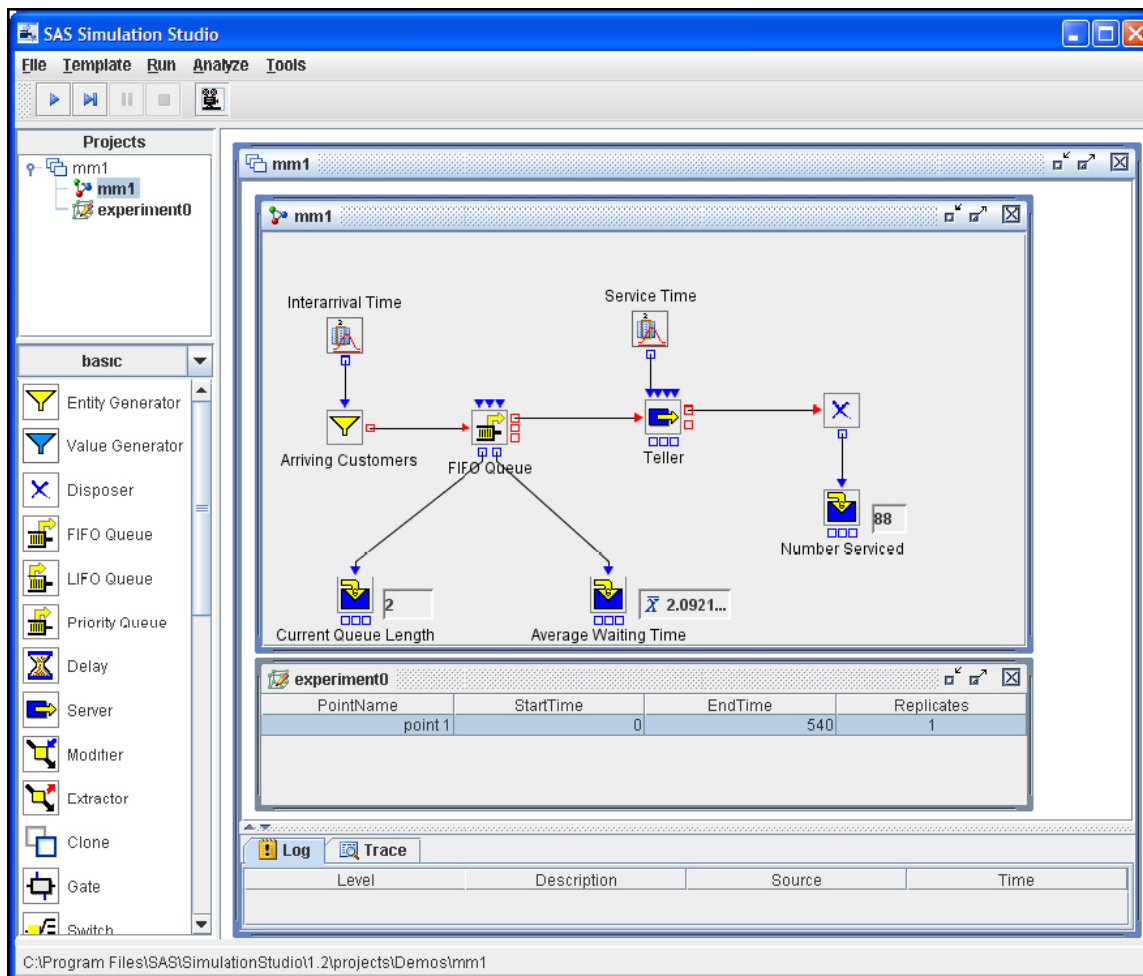
PROC BOM performs bill of materials processing in conjunction with the project scheduling software.

DISCRETE EVENT SIMULATION

Discrete Event Simulation is a widely used OR technique in which software is used to build logical models that can be used to study the characteristics of real world systems. In Discrete Event Simulation, the state of the system being modeled changes only when discrete events occur. A key element of simulation methods is that certain aspects of the system to be modeled are considered to be random. Items such as the rate of arrival or the length of time that is required to “serve” a transaction are considered to be random variables that are distributed in a known or assumed distribution. In essence, simulation is a software representation of a system. In many cases, simulation methods can provide a solution to a systems problem that is difficult to derive analytically or through direct observation.

The simulation facility in **SAS/OR** is the new SAS® Simulation Studio® which is the successor to the longstanding QSIM application. SAS Simulation Studio is offered on an experimental basis in SAS Version 9.2. It is a Java-based application which provides a graphical user interface and batch facility. Simulation models are built by drag and drop operations to select the various systems elements. Then, the model is run to represent a set period of time. The results are analyzed. After validating the model by comparing the simulation results with real world results, various assumptions can be tested. SAS Simulation Studio provides integration with the statistical analysis methods available in both SAS and JMP.

The figure below illustrates a very simple queuing model – a single teller bank. This queuing model is denoted as an M/M/1 queuing scenario in which customers arrive at an interarrival time based on an exponential distribution. The middle “M” in the notation represents the service time which is also drawn from an exponential distribution. The number “1” denotes that this is a single server system. The facility is similar to SAS® Enterprise Guide® in that it is project oriented. There may be multiple model windows in a project. The model is built by dragging the entity generator block, FIFO Queue block, Server block into the model window and making the appropriate connections. (Hughes, Chen, Lada, and Meanor, 2009)



Numeric source blocks are added to generate the interarrival time for the arriving customers and to select the service time that the teller spends with each specific customer. The interarrival service time also based on an exponential distribution. Number holder blocks and an entity disposer complete the model. When the model is run, statistics are collected pertaining to queue length and average wait time. Once a baseline is established, then different scenarios can be tested, such as a change in the interarrival rate or the addition of a server. This facility make model construction and use much easier than the code based methods of times past.

DEPLOYMENT AND VERSIONS OF SAS/OR

In order to run **SAS/OR** routines, the basic requirement is to have **Base SAS®** and **SAS/GRAPH®** in addition to **SAS/OR**. **SAS/GRAPH** is needed for network diagrams and decision trees. In SAS version 9.1.3, the newer PROCs and facilities were available in two special releases, versions 3.1 and 3.2, which were available to **SAS/OR** licensees on request. SAS Version 9.2 continues the improvements begun in the earlier releases. Certain parts of **SAS/OR** are still considered to be experimental, including PROC CLP and SAS® Simulation Studio.

CONCLUSION

As you can see, **SAS/OR** covers a lot of analytical ground. Hopefully, through this very brief overview, **SAS/OR** will not seem as mysterious as before. It will be interesting to see how this software evolves in the future and what types of problems that it will solve in coming years. My guess, however, is that it will continue to be "industrial strength".

REFERENCES

Churchman, C. W., R. L. Ackoff, and E. L. Arnoff. 1957. *Introduction to Operation Research*. New York, NY: John Wiley & Sons, Inc.

Hughes, Ed, Hong Chen, Emily Lada, and Phil Meanor. 2009. "Exploring System Performance with SAS® Simulation Studio." *Proceedings of SAS Global Forum 2009*, Washington, DC.

Hughes, Ed, and Trevor Kearney. 2009. "New Features in Optimization with SAS/OR® Software." *Proceedings of SAS Global Forum 2009*, Washington, DC.

SAS Institute, Inc. 2009. SAS Institute Fact Sheet. "SAS/OR®, Software."
http://www.sas.com/resources/factsheet/factsheet_or.pdf.

SAS Online Doc® Version 9.2. Copyright © 2008. SAS Institute Inc, Cary, NC, USA.

SAS System Help for Version 9.2. Copyright SAS Institute Inc., Cary, NC, USA.

SAS Institute, Inc. 2007. SAS Institute White Paper. "Optimization with SAS/OR®, What it is, What's New and How it Adds Value". <http://www.sas.com/apps/whitepapers/whitepaper.jsp>.

Thierauf, Robert J., and Richard A. Grosse. 1970. *Decision Making Through Operations Research*. New York, NY: John Wiley & Sons, Inc.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Garland D. (David) Maddox, Jr
Regions Bank
RCS-3
250 Riverchase Parkway East
Birmingham, AL 35244
Work Phone: 205-560-6339
E-mail: Garland.MaddoxJr@Regions.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.
Other brand and product names are trademarks of their respective companies.