The SIMRAND Methodology

Theory and Application for the Simulation of Research and Development Projects

R.F. Miles, Jr.

February 15, 1986

Prepared for
U.S. Department of Energy
Through an Agreement with National Aeronautics and Space Administration
by
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL Publication 85-98
The SIMRAND Methodology

Theory and Application for the Simulation of Research and Development Projects

R.F. Miles, Jr.

February 15, 1986

Prepared for
U.S. Department of Energy
Through an Agreement with
National Aeronautics and Space Administration
by
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

JPL Publication 85-98
ABSTRACT

A research and development (R&D) project often involves a number of decisions that must be made concerning which subset of systems or tasks are to be undertaken to achieve the goal of the R&D project. To help in this decision making, SIMRAND (SIMulation of Research AND Development Projects) is a methodology for the selection of the optimal subset of systems or tasks to be undertaken on an R&D project. Using alternative networks, the SIMRAND methodology models the alternative subsets of systems or tasks under consideration. Each path through an alternative network represents one way of satisfying the project goals. Equations are developed that relate the system or task variables to the measure of preference. Uncertainty is incorporated by treating the variables of the equations probabilistically as random variables, with cumulative distribution functions assessed by technical experts. Analytical techniques of probability theory are used to reduce the complexity of the alternative networks. Cardinal utility functions over the measure of preference are assessed for the decision makers. A run of the SIMRAND Computer I Program combines, in a Monte Carlo simulation model, the network structure, the equations, the cumulative distribution functions, and the utility functions. Repeated runs of the SIMRAND I Computer Program are made for each of the alternative networks to determine the optimal subset of systems or tasks. The SIMRAND methodology has been used by the Jet Propulsion Laboratory, the Solar Energy Research Institute, and the U.S. Department of Energy for the evaluation of solar-energy R&D projects.
SECTION 1
INTRODUCTION

A commonly occurring management decision in research and development (R&D) projects is the optimal allocation of R&D resources to achieve the project goals, within a set of resource constraints such as budget, personnel, and facilities. Because of these resource constraints, not all of the proposed R&D systems or tasks can be undertaken. The management decision then should be: "What subset of the set of proposed systems or tasks should be undertaken?" SIMRAND (SIMulation of Research ANd Development Projects) (References 1 and 2) is a methodology developed at the Jet Propulsion Laboratory (JPL) of the California Institute of Technology to aid in this management decision. There is an extensive body of literature on the subject of the allocation of resources for R&D projects (References 1 through 57).

Through personal interviews with R&D administrators and management scientists, Souder (Reference 49) has identified realism, flexibility, capability, ease of use, and cost as five major criteria for evaluating the suitability of management science models for R&D project selection. Because all modeling is of necessity an approximation to reality, and the reality of management decision-making inevitably includes aspects too complicated or unquantifiable to be modeled, the number of articles proposing models greatly exceeds the number of reported successful applications. Models invariably are forced to make simplifying assumptions that can produce either an analytically tractable model that is of strictly theoretical interest, or one that is too idiosyncratic to be of general interest. Thus, general models tend to be too general for real world application, and specific models are too specific to generalize beyond the problem, the project, the management philosophy of the decision makers, the modeling philosophy of the systems analysts, or the organizational process by which the decisions are ultimately made. Maher and Rubenstein (Reference 58) have investigated the reasons "...why a group of managers such as those in R&D refuse to use such management innovations as a computer-based project selection technique." As of 1975, Baker and Freeland (Reference 11) suggest eight limitations inherent in such models that could explain "...why few quantitative models of the R&D project selection and resource allocation decision have been implemented and used by R&D managers":

1. Inadequate treatment of risk and uncertainty.
2. Inadequate treatment of multiple criteria.
3. Inadequate treatment of project interrelationships.
4. No explicit incorporation of the knowledge of the R&D manager.
5. The inability to treat nonmonetary aspects.
6. Perceptions held by the R&D managers that the models are unnecessarily difficult to understand and use.
(7) Inadequate treatment of the time-variant property of data and criteria.

(8) Problems of consistency in the R&D project and the supporting staff.

The SIMRAND methodology originally was developed as a model to solve a specific R&D allocation problem on a specific project, at a time when the project's goals and staff were stable. The management of this specific project was both knowledgeable in and supportive of the methodology. Most of the remaining limitations listed above are directly addressed by the methodology. Thus, it was possible to develop the SIMRAND methodology and to apply it not only to the original project, but to other projects as well. The general form of the SIMRAND methodology should permit it to be applied to a wide class of R&D allocation decisions. Miles has published a brief introduction to the SIMRAND methodology (Reference 1), a more complete description of the methodology (Reference 2), and the computer code for the SIMRAND I Computer Program (Reference 59).

Section 2 discusses the class of R&D decisions to which the SIMRAND methodology is applicable. Section 3 describes the model formulation of the SIMRAND class of R&D decisions. Section 4 shows how uncertainty is incorporated in the model with the use of cumulative distribution functions (CDFs) for random variables. Section 5 is concerned with the use of utility functions to incorporate the preferences of the decision makers. Section 6 shows how the solution to the model formulation is obtained, using analytical techniques from probability theory and computer simulation techniques. Section 7 discusses the SIMRAND I Computer Program. Section 8 presents a representative analysis used in the validation of the runs of the SIMRAND I Computer Program. Section 9 discusses actual applications of the SIMRAND methodology. Section 10 proposes future enhancements to the SIMRAND methodology that could be of practical value. Section 11 contains the references.

The methods by which CDFs and utility functions are assessed are examined in detail in Sections 4 and 5. It has been JPL's experience that the exact phrasing of the assessment questions can directly affect the quality of the assessments. The phrasing can influence the understanding and cooperation of the technical experts and the decision makers.
SECTION 2

PROBLEM DESCRIPTION

The SIMRAND methodology is applicable to a class of R&D project decisions where the constraints of a project's resources only permit a subset of the set of systems or sets of tasks under consideration to be undertaken. For the full capability of the SIMRAND methodology to be appropriate, the following criteria must be satisfied:

1. **Goals**: It must be possible to relate the goals of the project to the hardware and manpower elements of the various systems or sets of tasks. It must be possible to estimate such variables as cost and performance of both the hardware and manpower elements.

2. **Common Measure of Preference**: The various systems or sets of tasks for achieving the project goals must have a common measure of preference.

3. **Multiple Alternatives**: More than one system or set of tasks must be considered to satisfy the project goals. The systems or sets of tasks may be fundamentally different, such as windmills versus solar cells for generating electricity. Or, they may be discrete parametric variations in a generic design, such as payload versus range for an aircraft.

4. **Uncertainty**: Uncertainty must be present with respect to the variables that describe the systems or sets of tasks. The raison d'être of R&D funding is to remove or at least reduce uncertainty. In doing so, the "best" system or set of tasks is identified. If this is not required, the project would enter directly into the implementation phase. It must be possible to describe this uncertainty probabilistically in terms of random variables, and to assess CDFs for the random variables.

5. **System and Preference Equations**: It must be possible to develop equations that describe the systems or sets of tasks and relate them to the measure of preference in terms of measurable random variables.

6. **Decision Makers**: It must be possible to identify decision makers and to assess their preferences on a cardinal utility scale. If there is more than one decision maker, then a group decision rule must be invoked.
The SIMRAND methodology models the R&D project-decision process by means of alternative networks that represent the feasible subsets of systems or tasks that are to be considered. An alternative network is constructed manually from the set of all systems or tasks under consideration such that:

(1) The selected subset of systems or tasks does not violate any of the project constraints.
(2) Each path through the alternative network represents one distinct system or subset of tasks capable of satisfying the project goals.

The present version of the SIMRAND I Computer Program (discussed in Section 7) does no optimization in the sense of attempting to construct an optimal alternative network.

Figure 3-1 is a simplified illustration of a task network for solar-cell production. The task network of Figure 3-1 has four stages:

(1) Silicon purification.
(2) Silicon crystallization.
(3) Ingot sawing.
(4) Cell assembly.

The task network has one task for silicon purification, and three parallel tasks for silicon crystallization, of which Task 2A requires sawing of ingots into wafer-thin sheets in Task 3A. In addition, there are two parallel tasks for cell assembly. There is a total of six paths through the task network of Figure 3-1. Consistent with the requirement that there be at least one task for each of Stages 1, 2, and 4, and that task networks with Task 2A require Task 3A, there are 20 proper subsets of the task network of Figure 3-1 that could be considered for lower levels of resource allocation.

A measure of preference is defined, and equations are developed that define the systems or tasks and relate them through the network paths to the measure of preference. Uncertainty is incorporated by treating the variables of the equations as random variables, for which CDFs can be assessed. Cardinal utility functions over the measure of preference are assessed for each of the decision makers. Either a single decision maker must be identified as the final arbitrator, or a group-decision rule must be invoked.

This completes the model formulation. The model is defined by alternative networks, and system or task equations that permit an evaluation of the alternative networks in terms of the measure of preference. The input data for the SIMRAND I Computer Program are given by a two-dimensional array that defines the task network, the equations and their random variables that define the systems or tasks, the CDFs for the random variables of the equations, the equation for the measure of preference, and the cardinal utility functions for the decision makers. The cardinal utility functions are not always required to use the SIMRAND methodology.
SECTION 4

PROBABILITY ASSESSMENT

The CDFs for the random variables that enter into the network equations must be assessed by experts in the technologies associated with the random variables. A systems analyst initially must explain the concept of probability assessment to a technical expert, preferably in a one-on-one meeting, and explain how the CDFs are to be assessed. The systems analyst should be present the first few times a technical expert assesses a random variable. Early assessments by the technical expert may later be updated by himself alone. In some cases, where competing technologies are being considered, the technical experts can be brought together in a group and asked to justify the rationales of their assessments. Because they can reveal information that is relevant to the decision-making process, these group meetings can be especially valuable to the project management. The many facets of subjective judgment that necessarily go into the probability assessment can be exposed for debate. It is extremely important that these probability assessments have the full support of the project management and also the line management of the technical experts. This will minimize organizational pressures that may bias the probability assessments.

An extensive literature (References 60 through 72) proposes several different methods of assessing probabilities. No single method of assessing probabilities has been used in conjunction with the SIMRAND methodology at JPL, and no claim is made that the probability-assessment method described here is superior to any other method. The probability-assessment method for the SIMRAND methodology has gone through minor evolutionary changes at JPL. The method described here is not the method that has been the most-used in the past, but represents the present state of the probability-assessment method. The probability-assessment method presented here is largely consistent with the recommendations of the SRI Decision Analysis Group (Reference 71), with consideration given to the assessment biases discussed by Kahneman, Slovic, and Tversky (Reference 62). All the probability-assessment methods used in the past were only minor variations of what is presented here. All have proven to be workable in that the project management, the technical experts, and the systems analysts were comfortable with them, the methods were consistent with the requirements of the model, and the methods required a minimum of explanation to be carried out.

Two different sets of questions are asked with respect to each random variable. The technical expert first is asked for the probability that the R&D activity associated with the random variable will succeed. He is then asked for the default value that the random variable would assume if the associated R&D activity were to fail. The answer to this question may be the known current value of the random variable.

The second set of questions involves assessing a CDF conditional on the success of the associated R&D activity. This permits the technical expert to disassociate the issue of the success or failure of the R&D activity from considerations of the conditional CDF. The CDF is assessed, at least initially, by the technical expert responding to a series of questions asked by the systems analyst. The questions are phrased to avoid an induced bias in the responses of the technical expert.
The probability-assessment method for the second set of questions (the conditional CDF) involves the following:

(1) Determine a lower limit for the CDF by asking an odds question.

(2) Determine an upper limit for the CDF by asking an odds question.

(3) Determine the median by an interval-estimation technique.

(4) Determine values for the CDF between the lower limit and the median by the interval-estimation technique.

(5) Determine values for the CDF between the median and the upper limit by the interval-estimation technique.

The odds question for the lower limit asks for a variable estimate such that there is only one chance in a thousand that the true value for the variable will lie below the variable estimate. The odds question for the upper limit asks for a variable estimate such that there is only one chance in a thousand that the true value for the variable will lie above the variable estimate. For the "CELL, $/m^2" CDF of Figure 4-1, the lower and upper limits would have been $12 and $25.

Figure 4-1. Cumulative Distribution Function for the Random Variable "CELL" for Task 4A
The interval-estimation technique involves asking questions as to which of two contiguous variable intervals are either more probable or about equally likely. The variable value separating the contiguous intervals is varied until the variable intervals are perceived as equally likely.

The first CDF value determined with the interval-estimation technique is for the 0.50 fractile, the 50/50 even-odds point or the median estimate. It is obtained by using the lower-limit estimate and the upper-limit estimate as the extreme end-points for the variable intervals. The questions are phrased, "Is it more likely that the <variable> lies between <lower-limit estimate> and <stated-intermediate value>, or between <stated-intermediate value> and <upper-limit estimate>, or is it about equally likely?" The words between the angle brackets are replaced by the name or values of the variable. The stated-intermediate values are initially set to extreme values, and then varied between low and high values as the technical expert "homes-in" on the equally-likely value, which is the 0.50 fractile. For the "CELL, $/m^2" CDF of Figure 4-1, the first question asked of the technical expert might be, "Is it more likely that the cost lies between $12 and $15, or between $15 and $25, or is it about equally likely?" The CDF of Figure 4-1 shows that the technical expert would have answered, "Between $15 and $25." The stated-intermediate value is varied in subsequent questions until the variable intervals are assessed by the technical expert as equally likely.

The second CDF value assessed by the technical expert is for the 0.25 fractile, and is obtained by using the lower-limit estimate and the median estimate as the extreme end-points for the variable intervals. The first question is phrased, "Is it more likely that the <variable> lies between <lower-limit estimate> and <stated-intermediate value>, or between <stated-intermediate value> and <median estimate>, or is it about equally likely?" The questions proceed as in the previous paragraph until the technical expert reaches the equally-likely point.

Subsequent sets of questions proceed in a similar manner until the CDF can be adequately defined between the lower-limit estimate and the median estimate. Then the questions switch to intervals between the median estimate and the upper-limit estimate to define the CDF between the median and the upper-limit estimate. In the SIMRAND I Computer Program runs at JPL, the one-chance-in-a-thousand odds estimates are taken as the 0.00 and the 1.00 fractile points for the CDFs.
A cardinal utility function for the measure of preference is assessed for each of the decision makers by means of von Neumann-Morgenstern lotteries (References 67, 73 through 82). Almost any of the methods proposed in these references for assessing von Neumann-Morgenstern lotteries would probably be acceptable. The method used with the present SIMRAND methodology at JPL involves only 50/50 lotteries. Let $z(u)$ be the value of the measure of preference whose utility is assessed as $u$, where $u$ is constrained to range from $u = 0.0$ for the least-preferred value of $z$ to $u = 1.0$ for the most-preferred value of $z$. A graph (Figure 5-1), placed before the decision maker, has a 50/50 lottery on the left (called a "gamble" in discussions with the decision maker) and a vertical scale (called the "sure thing") on the right. The lottery is a gamble that yields the value of $z$ at the end of the bottom path with probability 0.50, or the value of $z$ at the end of the top path also with probability 0.50. The vertical scale is marked with values of $z$ increasing in preference from the bottom up, one value of which the decision maker is told he can have "for sure." The end-points of the vertical scale are the same as the corresponding values of $z$ for the lottery.

Initially, the least-preferred value of $z$ is placed at the end of the bottom path of the lottery, and the most-preferred value of $z$ at the end of the top path of the lottery. Similarly, the "sure thing" vertical scale has the least-preferred value of $z$ placed at the bottom and the most-preferred value of $z$ placed at the top. Appropriate intermediate values of $z$ are marked along the vertical scale.

After the assessment procedure is explained to the decision maker, he is asked whether he would prefer the lottery or the least-preferred value of $z$ on the vertical scale for sure. An upward-pointing arrow then is placed beside the bottom of the vertical scale. This assumes that the decision maker answers correctly. The decision maker next is asked whether he would prefer the lottery or the most-preferred value of $z$ on the vertical scale for sure. A downward-pointing arrow then is placed beside the top of the vertical scale.

An intermediate value of $z$ on the vertical scale then is selected, and the decision maker is asked whether he would prefer the gamble or the selected value of $z$ on the vertical scale for sure. If the decision maker prefers the lottery, then an upward-pointing arrow should be placed beside the intermediate value of $z$. Contrariwise, the arrow should point down. The arrow indicates the direction from which the next intermediate value of $z$ should be selected. In this manner, the assessment procedure "homes-in" on the decision maker's indifference point. For the first set of questions—ranging between the least-preferred and the most-preferred values of the measure of preference—the indifference point yields $z(0.50)$, the value of $z$ for which the utility function value is 0.50. A new graph then is placed before the decision maker, and a new scale is drawn with $z(0.50)$ as one end-point and either the least-preferred or the most-preferred value of $z$ as the other end-point. Repeating the assessment questions will yield either $z(0.25)$ or $z(0.75)$. Continuing this procedure will generate enough points to fair in a curve for the decision maker's utility function.
This utility assessment procedure, developed by Miles, Feinberg, and Brooks (Reference 83), is not always required to use the SIMRAND methodology. Some dummy utility function, however, must be included in the input data for the SIMRAND I Computer Program to run.

Figure 5-1. Graph for Assessing the Utility Function of a Decision Maker
The SIMRAND methodology proceeds in three phases to solve the problem as modeled in Section 3—a reduction phase, a simulation phase, and an evaluation phase. In the reduction phase, analytical techniques from probability theory and simulation techniques are used to reduce the complexity of the alternative networks. In the simulation phase, a Monte Carlo simulation (References 84 through 86) is used to derive statistics on the variables of interest for each alternative network. In the evaluation phase, the statistics from the Monte Carlo simulation runs of the alternative networks are compared, and the alternative networks are rank-ordered in preference by the selected decision rule.

A. REDUCTION PHASE

The purpose of the reduction phase is to reduce the complexity of the alternative networks before entering the simulation phase and executing the Monte Carlo simulation runs. The need for a reduction phase may be seen from the complexity of an alternative network consisting of ten stages and four parallel tasks for each stage. The network has 1,048,576 distinct paths, where a path must traverse all stages through any of the parallel tasks. Clearly, any techniques that can be applied to reduce the number of paths can be extremely useful. The reduction techniques used to generate a reduced network may be of an analytical nature based on probability theory, or they may involve a simulation of a portion of the alternative networks. Series-parallel reduction techniques have been used by other workers (Reference 87) and are recommended as a precursor to simulation by Burt and Garman (Reference 88).

Two examples will illustrate the analytical concept for the reduction phase. If the measure of preference is to be at a minimum for a particular stage, and the CDFs are known for two parallel but statistically independent tasks of that stage, then the CDF for the tasks in parallel is obtained easily as follows:

\[ F(z) = F_1(z) + F_2(z) - F_1(z)F_2(z), \]

where \( F(\cdot), F_1(\cdot), \) and \( F_2(\cdot) \) respectively are the CDFs for the tasks in parallel, the first task, and the second task. If the measure of preference is additive for two stages with known and statistically independent CDFs, then the CDF for the two stages is a convolution integral of the form:

\[ F(z) = \int_{-\infty}^{+\infty} F_1(z-x) \, d[F_2(x)]. \]

If the tasks have similar technologies, the justification of the independence condition may be difficult.
Whether or not reduction techniques are applied, the problem of statistical dependence must always be examined whenever a number of random variables are present in the model. Several techniques may be considered at this point. It may be possible, through a sensitivity analysis, to demonstrate that the statistical dependence does not alter the optimal solution. It may be possible to restructure the model to separate out the statistically dependent portion. It also may be possible to assess a CDF for one random variable that is conditional on the realized value of another statistically dependent random variable. A crude approximation of assessing only two or three conditional CDFs for a random variable may suffice. For example, yields and thicknesses of solar cells are statistically correlated, since thin cells are more prone to breakage. It may suffice for the yield of solar cells to assess only two CDFs, conditional on whether the cells are thick or thin.

For the alternative networks and their associated equations to which the SIMRAND methodology has been applied, it has not been possible to assume normal distributions or to apply the central limit theorem. Random variables often associated with efficiencies or manufacturing yields will appear in the denominator of equations that have random variables for unit costs in the numerator. The ratio of two normally distributed random variables has a Cauchy distribution, having neither a mean nor a variance. The distributions also often have highly skewed shapes in addition to the displaced values for task-failure default conditions. In addition, the central limit theorem is not applicable for summations over a limited number of stages. Simulation results have shown that the resulting CDFs are often highly skewed.

Reduction techniques are important because they can make the difference whether the resulting reduced network is computationally feasible. But they must be applied with care if the results are to have sufficient accuracy. The reduced task network of Figure 6-1 shows the network that might result from applying reduction techniques to the task network of Figure 3-1 for Task 2B and Task 2C.

B. SIMULATION PHASE

The simulation phase begins by expanding the reduced task networks into simulation task networks, in which all paths through the reduced task networks are explicitly represented. The simulation task network of Figure 6-2 results from the expansion of the reduced task network of Figure 6-1. The SIMRAND I Computer Program then performs the simulation by carrying out a series of Monte Carlo trials on each of the simulation task networks.

A Monte Carlo trial is performed by:

1. Assigning a different random value to each of the different random variables in the equations associated with the simulation task network.

2. Calculating the measure of preference for each path through the simulation task network by substituting the random values into the appropriate path equations.
(3) Selecting the optimal path through the simulation task network as determined by the measure of preference.

(4) Incrementing the number of times that path has been selected as the optimal path.

(5) Updating histograms and statistics of selected random variables, the measure of preference, and other equations of interest based on values determined by the optimal path.

(6) Updating the utility function values and certainty equivalents of the decision makers.

![Diagram of Task Network](image)

Figure 6-1. Reduced Task Network for Solar-Cell Production

![Diagram of Task Network](image)

Figure 6-2. Simulation Task Network for Solar-Cell Production
Uniformly distributed pseudorandom numbers for generating realized values of the random variables are obtained by the linear congruential method (Reference 89). Miles (Reference 90) has published the computer code along with several statistical tests for the random number generator in the SIMRAND I Computer Program. A typical CDF for a random variable is shown in Figure 4-1. The technique for generating realized values of the random variables from the uniformly distributed pseudorandom numbers is described in Chapter 3 of Reference 85.

The simulation phase concludes by using the histograms to calculate CDFs of the selected random variables, the measure of preference, and other equations of interest. Figure 6-3 shows the output of the SIMRAND I Computer Program for both the CDF of the measure of preference for the example given in Reference 2, as well as the utility function used in Reference 2.

![Figure 6-3. Cumulative Distribution Function and Utility Function for the Random Variable "Total Price"](image-url)
A Monte Carlo simulation computer program has also been used in development projects by Abernathy and Rosenbloom (Reference 6) with CDFs assessed by the project management for the analysis of parallel strategies.

C. EVALUATION PHASE

The evaluation phase consists of rank ordering, for each decision maker, the simulation task networks by use of the utility function values, where greater utility function values imply more preferred simulation task networks. The same rank ordering can be achieved, of course, by using the certainty-equivalent values for the simulation task networks. If a group-decision rule is invoked, it then is used to determine a group ranking for the simulation task networks. A rank order for the simulation task networks is tantamount to a rank order for the original alternative task networks. The alternative task network ranked as most preferred will specify the optimal subset of systems or tasks for the R&D project.

As long as stochastic dominance holds, where the CDFs for the simulation task networks do not cross over, the simulation task networks can be rank-ordered by their CDFs, and the utility functions are not required. If stochastic dominance is violated, however, the utility functions (or some other decision rule such as min-max) are essential to rank-order the simulation task networks.
Development of the SIMRAND methodology was initiated in 1979. By the middle of 1980, questions were being posed that exceeded the analytical and heuristic capability of the SIMRAND methodology. At that time, an effort was undertaken to develop a Monte Carlo simulation model that could extend the SIMRAND methodology to more complex system or task networks. From the outset, one criterion imposed on the development of the model was to retain a general form for the computer program, so that it could easily be adapted to problems on different projects. The first SIMRAND I Computer Program was completed in November 1980, and was written in FORTRAN to be run on a Univac 1108 mainframe computer (Version 1), and subsequently on a Univac 1100/81 mainframe computer (Version 2). In 1981, Version 3 of the SIMRAND I Computer Program was written in Microsoft FORTRAN-80 (Reference 91) to run on an 8-bit microcomputer using the Intel 8080 microprocessor or the Zilog Z-80 microprocessor.

The present version of the SIMRAND I Computer Program (Version 5.0x03) is written in Version 3.30 of Microsoft FORTRAN (References 92 and 93), to be run under Version 2.10 of the Microsoft Disk Operating System (References 94 and 95). In addition to the FORTRAN "common blocks", the current version also has all equations defining the task network maintained in separate files to be included during compilation. Thus, only two small source-code files must be modified to apply the SIMRAND I Computer Program to an analysis of a different model. For the simulation task network of Figure 6-2, 10,000 Monte Carlo trials takes 28 minutes on an IBM PC-XT with an Intel 8088 microprocessor and an Intel 8087 numeric data processor (arithmetic coprocessor) (Reference 96). Only minor modification of the source code would be required to run the current version of the SIMRAND I Computer Program on a mainframe computer. This is because Version 3.3 of Microsoft FORTRAN is a superset of the standard ANSI X3.9-1978 Subset FORTRAN. See Reference 59 for a description of the current version of the SIMRAND I Computer Program.

The input data for a SIMRAND run consist of:

(1) Parameters that define the run identification, the diagnostic display level, whether the short or long form of the output is desired, the number of Monte Carlo trials, and the sizes of arrays.

(2) A two-dimensional integer array that defines the simulation network. The rows correspond to the paths, the columns correspond to the random variables, and the elements specify the appropriate CDFs. An extra 0/1 variable for each path is included for sensitivity analysis to permit paths to be easily excluded from the simulation task network.

(3) CDFs for the random variables.
(4) Utility functions for the decision makers. Utility function data are not required in order to use the SIMRAND I Computer Program. A dummy utility function can be input as a straight line defined by two points that bracket the range of the measure of preference.

(5) Two arrays that define the lower and upper bounds of the histogram ranges.

The CDFs and the utility functions are input as piecewise linear fits to the assessed data. A typical CDF is depicted in Figure 4-1. A typical utility function is shown in Figure 6-3. Both of these functions are used in the SIMRAND I Computer Program run for Reference 2. Piecewise linear fits have an advantage over polynomial regressions in that:

(1) Computation time is reduced. No regression equations need to be solved, and it is far easier to calculate values for the inverse function of a piecewise linear fit than for a polynomial fit.

(2) Weak monotonic increasing functions (as required by CDFs) can be easily constructed.

(3) The end-points (CDF fractiles of 0.00 and 1.00) can be matched exactly.

(4) Better control of the fitting error can be obtained, especially if abrupt changes in the slope of the function occur, or if the function has a long tail.
SECTION 8

REPRESENTATIVE ANALYSIS

A representative analysis is performed as part of the validation of the simulation networks, the associated equations, and the input data. The representative analysis consists of a calculation, for each simulation network path, of a representative measure of preference and representative values for any other equations of interest relevant to the network paths. The representative measure of preference and representative values are calculated by determining a single value for each random variable. This is obtained from the associated CDF by multiplying the median of the conditional CDF by the probability of success ($p_s$) and adding the result to the failure-default value times $(1 - p_s)$. If the mean of the conditional CDF were used instead of the median, the resultant quantity would be the mean of the unconditional CDF. The median is used because it can be directly obtained from the conditional CDF without an integration, and for the purpose of the representative analysis the mean offers no advantage. These representative values are then entered into the equation for the measure of preference and other equations of interest for each path. The representative values, and the values that result from their use in the equations, do not correspond to any commonly used statistical quantities. They may not even be realizable values, but in all calculations the measures of preference have had resultant values that fall within 10% of their means as determined by the computer program runs.

The representative analysis serves a second, extremely useful purpose in that for a specific path it can be done in a matter of minutes—a back-of-the-envelope type of calculation. This analysis can be used to explain why a certain path did or did not enter into the selection of the Monte Carlo trials, or why a certain path might or might not be expected to influence the measure of preference for a SIMRAND I Computer Program run. This provides credibility to the analysis in response to statements such as "There must be an error in the model formulation" or "The wrong data must have been entered into the run." Decision makers rightly need intuitive feelings about why certain decisions would be optimal. This is particularly true if the complexity of the networks makes the optimal decisions seem counterintuitive.

Of course, the representative analysis does not substitute for a thorough validation of the equations entered into the SIMRAND I Computer Program and the input data. This validation has typically been done by manually checking all intermediate and final calculations on a SIMRAND I Computer Program run of four Monte Carlo trials.
SECTION 9
APPLICATIONS

The development of the SIMRAND methodology was initiated in 1979 by the Flat-Plate Solar Array Project (FSA) (References 97 and 98). At that time, the Project was called the Low-Cost Solar Array Project. During the period when the SIMRAND methodology was used, the goal of FSA was to reduce the price of flat-plate solar-cell modules by improving manufacturing technology, adopting mass-production techniques, and promoting user acceptance. In 1979, strictly analytical techniques were used to calculate CDFs of the price of silicon material for sets of silicon-material processes. In early 1980, analytical and heuristic techniques were used to calculate CDFs of solar-cell module prices for several levels of FSA funding. The same techniques were used again in June 1980 for an FSA review.

By the middle of 1980, the questions that were being posed by FSA management exceeded the analytical and heuristic capability of the SIMRAND methodology. A computer program for a Monte Carlo simulation model was then developed that could extend the SIMRAND methodology to more complex task networks and more complex equations. The SIMRAND computer program first was used for the November 1980 FSA review. The July 1981 FSA review made use of an improved version of the SIMRAND computer program. For the production of solar-cell modules, the July 1981 FSA simulation network had a maximum of 20 paths with a maximum of eight stages in each path. A value-added price was calculated for each stage, and the value-added prices were summed for the measure of preference. A total of 17 random variables were required, selected for the various paths from 38 CDFs.

In July 1981, in a joint effort by the Jet Propulsion Laboratory and the Solar Energy Research Institute, the SIMRAND methodology also was used to make an assessment of crystalline silicon materials for solar cells. The simulation networks had a maximum of 18 paths with a maximum of seven stages in each path. A total of 17 random variables were required, selected from 75 CDFs.

In 1981, the Solar Thermal Power Systems Project of JPL (Reference 99) used the SIMRAND methodology in a probabilistic cost study of solar-dish power systems (Reference 100). The subsystems for these solar-dish power systems were divided into five stages, with a simulation network of 15 paths. A total of 16 random variables were required, selected from 54 CDFs. The measure of preference included not only the solar-hardware costs but also the balance of plant costs scaled from the hardware configuration, along with the annual operations and maintenance costs.

In 1984, the Photovoltaics Program Analysis and Integration Center of JPL used the SIMRAND Methodology to study the relative economic potentials of two-axis tracking and of concentrating flat-plate photovoltaic arrays for central-station applications (Reference 101). The task network contained 120 paths. A total of 15 random variables were required for cost and performance parameters, selected from 34 CDFs. The measure of preference was the projected cost of energy for commercial photovoltaic arrays in the mid-1990's.
In 1985, the United States Air Force funded JPL to develop a microcomputer program (AUTOVAL) for the evaluation of adding autonomy to a spacecraft design (References 102 and 103). The SIMRAND methodology was incorporated for treating the design parameters probabilistically. The measure of preference was multiattributed, using the Keeney-Raiffa methodology (Reference 77). Group decision rules were used to aggregate individual preferences. The AUTOVAL Program was applied to the evaluation of four alternatives for autonomous spacecraft for the Defense Meteorological Satellite Program.
The most significant enhancement of the current version of the SIMRAND methodology would be in the area of optimization. The alternative networks currently must be constructed and selected manually. In principle, it should be possible to use optimization techniques such that all constraints on the formulation of alternative networks would be explicitly stated. The optimization techniques would search out an optimal alternative network or a set of nondominated alternative networks.

Depending upon the mathematical complexity of the objective function and the constraint equations, the optimization problem could possibly be formulated in terms of the following various models: a linear or nonlinear mathematical programming model, an integer programming model (Reference 14), a goal programming model (References 53, 104 and 105), a dynamic programming model (Reference 33), a portfolio selection model (References 27, 29 and 106), an economic model (References 26, 41 and 46), a screening model (Reference 21), or a decision analysis model (References 15, 37, 40 and 42). Unfortunately, because of the inherently probabilistic nature of the SIMRAND methodology and because of the complexity of the applications discussed in Section 9, no straightforward formulation of an optimization model exists for those applications. Bard (Reference 13) has developed an algorithm combining dynamic programming and the SIMRAND methodology to obtain near-optimal solutions.

In the absence of a generalized analytical solution to the SIMRAND methodology, two approaches are possible. For small task networks, given the task descriptions and the constraint equations, a program could exhaustively enumerate all possible task networks. For large task networks, a heuristic possibly could be developed that would randomly select task networks. The heuristic would make the selection consistent with the task network constraints, and then would assign a weighting factor to the tasks, depending upon the calculated measure of preference. This weighting factor could be used to modify the probability that these tasks would be selected for future task networks. The objective would be to devise a weighting function that would result in more-preferred task networks being generated as the program run progresses.

A second enhancement of the SIMRAND methodology would be in the area of multiattribute decision analysis (References 30, 37, 77, 83, 107, 108, 109), with the alternative attribute states considered as random variables. This enhancement would be straightforward. Various stages of the alternative networks would correspond to attributes for which attribute-utility function values would be calculated in the Monte Carlo simulation. The measure of preference would be the multiattribute utility function. Smith, et al. (Reference 103) have modified the SIMRAND I Computer Program code to incorporate this enhancement.
A third enhancement of the SIMRAND methodology would be in the area of group decision making (References 110 through 114). Group decision-making models have been applied successfully in other contexts (References 7, 83, 108, 115 and 116), and the same theories would be relevant to the SIMRAND methodology. Smith, et al. (Reference 103) have modified the SIMRAND I Computer Program code to incorporate this enhancement.


69. Seaver, D. A., v. Winterfeldt, D., and Edwards, W., Soliciting Subjective Probability Distributions on Continuous Variables, SSRI Research Report 75-8, Social Science Research Institute, University of Southern California, Los Angeles, California, August 1975.


82. v. Winterfeldt, D., An Overview, Integration, and Evaluation of Utility Theory for Decision Analysis, SSRI Report 75-9, Social Science Research Institute, University of Southern California, Los Angeles, California, August 1975.


90. Miles, R. F., Jr., The RANDOM Computer Program: A Linear Congruential Random Number Generator, JPL Publication 85-97, Jet Propulsion Laboratory, Pasadena, California, February 15, 1986.


# The SIMRAND Methodology

**Abstract**

A research and development (R&D) project often involves a number of decisions that must be made concerning which subset of systems or tasks are to be undertaken to achieve the goal of the R&D project. To help in this decision making, SIMRAND (SIMulation of Research ANd Development Projects) is a methodology for the selection of the optimal subset of systems or tasks to be undertaken on an R&D project. Using alternative networks, the SIMRAND methodology models the alternative subsets of systems or tasks under consideration. Each path through an alternative network represents one way of satisfying the project goals. Equations are developed that relate the system or task variables to the measure of reference. Uncertainty is incorporated by treating the variables of the equations probabilistically as random variables, with cumulative distribution functions assessed by technical experts. Analytical techniques of probability theory are used to reduce the complexity of the alternative networks. Cardinal utility functions over the measure of preference are assessed for the decision makers. A run of the SIMRAND Computer I Program combines, in a Monte Carlo simulation model, the network structure, the equations, the cumulative distribution functions, and the utility functions. Repeated runs of the SIMRAND I Computer Program are made for each of the alternative networks to determine the optimal subset of systems or tasks. The SIMRAND methodology has been used by the Jet Propulsion Laboratory, the Solar Energy Research Institute, and the U.S. Department of Energy for the evaluation of solar-energy R&D projects.

**Key Words (Selected by Author(s))**

- Computer Programming and Software
- Systems Analysis

**Distribution Statement**

Unclassified—unlimited