A PRACTICAL APPROACH TO SENSITIVITY TESTING OF SYSTEM DYNAMICS MODELS

by

Andrew Ford
Economics Group
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

Jeffrey Amlin
Advanced Modeling and Simulation Group
Control Data Corporation
Dayton, Ohio 45439

George Backus
Advanced Modeling and Simulation Group
Control Data Corporation
Bloomington, Minnesota 55437

Submitted to:

The 1983 International System Dynamics Conference
Chestnut Hill, Massachusetts

July 27-30, 1983

A PRACTICAL APPROACH TO SENSITIVITY TESTING OF SYSTEM DYNAMICS MODELS

I. INTRODUCTION

Sensitivity testing, according to the glossary of terms in a Congressional manual on simulation modeling, is defined as the "running of a simulation model by successively changing the states of the system...and comparing the model outputs to determine the effects of these changes" (Congress 1975, p. 129). Sensitivity testing is generally viewed as an important part of the modeling process because it helps researchers narrow down those areas where more data gathering would be most useful. In our introductory remarks, we argue that detailed sensitivity testing is particularly important in system dynamics modeling efforts, and we list several obstacles that make detailed sensitivity testing difficult. We introduce a set of testing procedures developed at the Los Alamos National Laboratory and verified by the Control Data Corporation that can help system dynamicists perform detailed sensitivity testing on a routine basis.

In the body of the paper, we present an illustrative application of the testing procedures, and we list six specific uses of the procedures. We describe the availability of the testing package, and we conclude with a set of practical guidelines for investigators wishing to make use of this unique set of procedures.

A. Importance of Sensitivity Testing

Sensitivity testing is an important part of all modeling projects, but it is especially valuable in system dynamics projects because system dynamicists tend to:

1. close feedback loops in their models;
2. rely on less precise information in estimating parameters;
3. expect model behavior to be insensitive to changes in most parameters;
4. and are reasonably successful in achieving model implementation.

PREFACE

In July 1983, system dynamicists will meet at the 1983 International System Dynamics Conference to discuss model validation. The conference focus on model validity is appropriate since the question most frequently asked about models of social systems is "Has the validity of the model been proved?" Our view is that scientific proof of model validity is impossible. No model has been or ever will be thoroughly validated since models are designed as simplifications of the simulated system. Rather than aspiring for a proof of validity, one should look for simple, pragmatic steps to bolster confidence in the model.

We expect that the majority of the participants in the 1983 conference will agree that sensitivity testing is one of several pragmatic steps that can be taken to improve one's confidence in a model. With the hope that sensitivity testing will be performed in a more thorough and more detailed fashion in the future, we present "A Practical Approach to Sensitivity Testing of System Dynamics Models."

ABSTRACT

Sensitivity testing is an important part of all modeling projects, but it is especially valuable in system dynamics projects. This paper presents a practical approach toward sensitivity testing that will allow system dynamics analysts to perform thorough and detailed testing on a routine basis. We illustrate the approach by calculating tolerance intervals on a system dynamics model projection of oil and natural gas consumption by the US electric utility industry. Forty-five model input parameters are considered uncertain, and twenty simulations runs are used to gain the statistical information needed to calculate confidence bounds. The paper discusses six applications of the sensitivity testing procedures, and concludes with suggestions for practical application. Ten alternative approaches to sensitivity analysis are reviewed in the Appendix.

Driven by the expectation that closure of key feedback loops will lead to better understanding of system behavior, system dynamicists are likely to include many highly uncertain parameters in their models. It is often the case that many of the causal relationships in a loop are easily quantified, but a final relationship needed to close the loop is quite difficult to represent. The natural tendency of most analysts is to omit the difficult relationship, but a system dynamicist is likely to include the difficult relationship even if he must rely on expert judgment or personal intuition to close the loop. System dynamicists proceed with this style of modeling fully aware that the inclusion of such uncertain parameters opens the model to criticism, especially from analysts more accustomed to open system models which only include the more easily estimated parameters.* System dynamicists risk such criticism because they expect model behavior to be insensitive to the vast majority of the parameter values, and they wish to concentrate their efforts on the search for the few sensitive points where small changes in parameter values may cause large changes in the pattern of model behavior. It is our opinion that the analyst who is willing to include highly uncertain parameters in a model should be prepared to perform detailed sensitivity testing to confirm or reject his expectation that there are only a few sensitive points in the model.

The fourth reason for the special importance of detailed sensitivity testing in system dynamics projects is the past success that system dynamicists have achieved in model implementation. With successful implementation, however, one often finds multiple layers of structural additions motivated by the client's interest in new problems. With each new layer of structure, the analyst's ability to understand the essential workings of the model is diminished. Over time, a model may grow to such proportions that it resembles an artichoke with so many leaves that only the most persistent analyst can get

---

*These aspects of system dynamics projects are aptly described by Greenberger as follows:

> The users of system dynamics do not shy away from applying their models to complex social problems. They strive to compensate for the limited supply of reliable data by drawing on the opinions of experts and on their own intuitions. They seek to identify causal structures and set parameter values, not by traditional data analyses and correlation studies, but by what some consider "armchair speculation" and economist Lawrence Klein apprehensively refers to as "stylizing the facts."
>
> (Greenberger 1976, p. 126)

to the heart of the model.* We feel that the sensitivity procedures described in this paper are particularly applicable to system dynamicists wishing to distinguish between the leaves and the heart of an "artichoke model."

B. Difficulties in Sensitivity Testing

Our views on the importance of sensitivity testing are not unique. We doubt that there is a single participant at this conference who does not view sensitivity testing as a crucial step in the modeling process. One must wonder, therefore, why detailed sensitivity testing is not reported on a routine basis as part of the customary model documentation and summary reports. We suspect that the following attributes of simulation models have made detailed sensitivity testing difficult:

1. There are a large number of model parameters that require testing.
2. There are a large number of output variables that might be monitored as a measure of sensitivity.
3. The output generated is dynamic and consists of patterns that vary with time.
4. Many models are constructed without a clear statement of purpose. The decision to be made as a result of the information gained from the model is not specified.

Our purpose in this paper is to present a set of statistical procedures that can help the system dynamicist overcome the first of the four obstacles. For suggestions on dealing with the remaining three obstacles, we refer conference participants to "A New Measure of Sensitivity for Social System Simulation Models" developed by Ford and Gardiner (1979). We refer participants to Appendix A for a brief description of related research on sensitivity methods for system dynamics models. Participants are referred to the paper by Tank-Nielsen (1980) for a discussion of the overall role of sensitivity testing in the model construction process.

---

*The "artichoke effect" is a term coined by Walter Carlson to help describe the complexity of computer systems:

> We have a proclivity to add features, add functions, and add interfaces--layer upon layer--onto existing systems. Each succeeding layer has less and less useful or tasty substance on it, until the outside layers merely add weight, complexity, and a prickly hindrance to reaching the core of the problem.
>
> (Greenberger 1976, p. 73)

C. Background on the Sensitivity Testing Procedures

The procedures described here were originally developed by Michael McKay of the Statistics Group of the Los Alamos National Laboratory. McKay was asked to develop a sampling procedure that would allow Los Alamos scientists to learn the most important inputs to complex computer models of nuclear reactor performance during a simulated loss of coolant accident. The computer code solved three dimensional partial differential equations to find the temperature and pressure changes in the simulated core of a commercial nuclear reactor. The code required substantial computational time for one simulation experiment. In research performed for the Nuclear Regulatory Commission, McKay used the sampling procedure known as Latin Hypercube Sampling (LHS) which would allow the Los Alamos scientists to gain the most information on key inputs within a budget constraint on the number of computer runs. To apply the LHS procedure for selecting proper values of the model inputs, the user-specified range of plausibility on each input is divided into N equal probability intervals, where N is the number of computer runs allowed with the model. A value is selected from each interval according to the user specified conditional distribution, and the values for each input are assigned at random to the N model runs. The sampling procedures and properties of the estimators obtained from LHS are described in previous publications (McKay, Conover, and Whiteman 1976; McKay, Conover and Beckman 1979).

Later in research performed for the Energy Information Administration, Los Alamos scientists applied the LHS procedures to COAL2, a medium-sized model of the US energy system.* The COAL2 case study indicated that McKay's procedures could be easily applied once the ranges of plausibility of each model input were specified. The procedures were later applied to a complex simulation model of oil resource exploration on government lands in Alaska by Abbey and Bivins (1982).

*COAL2 is a system dynamics model of the US energy system developed by Dr. Roger Naill (1976, 1977) as part of a Dartmouth College research project on the US coal industry. A revised version of the COAL2 model was used to test the effects of President Carter's National Energy Plan at the request of the House Subcommittee on Energy and Power (Naill and Backus 1977). Extensions and improvements in COAL2 led to the FOSSIL2 model now used at the Department of Energy in preparing the department's annual forecasts (EEA 1980, NEP II 1979). In the Los Alamos sensitivity test, 72 COAL2 input parameters were considered as uncertain. Results of the Los Alamos test are described in two technical reports from the laboratory (Ford, Moore, and McKay 1979; McKay 1978).

By 1981, Los Alamos researchers had applied the LHS testing procedures to six different models which employed a variety of techniques including linear programming, numerical analysis of partial differential equations, algebraic equations, and system dynamics. The procedures function in the same way regardless of the particular modeling approach because the model is treated as a "black box" whose properties are to be tested statistically. Although McKay's procedures are applicable to any kind of model, the package of computer programs were limited to application on the Los Alamos computer system in 1981.

In 1982, analysts from Los Alamos and the Control Data Corporation set out to verify McKay's procedures. Our purpose was to enhance confidence in the LHS testing package by demonstrating that results of past tests could be reproduced on an independent computer system.* A second objective was to make the package of programs available to a wider group of modelers than those with access to just the Los Alamos computer system. The Los Alamos/Control Data Corporation project successfully implemented the LHS testing procedures for specific application to system dynamics models (Amlin 1982), and detailed sensitivity testing can now be performed on a routine basis by any system dynamicist with access to the Dartmouth College computer or the Control Data Corporation's CYBERNET system. In the remainder of this paper, we demonstrate through illustrative examples the type of information to be gained when applying the sensitivity testing procedures to system dynamics models.

II. TOLERANCE INTERVALS ON US OIL AND GAS CONSUMPTION BY ELECTRIC UTILITIES

A. The Illustrative Example

Our demonstration makes use of a system dynamics model designed to simulate the operations of a hypothetical, investor-owned electric utility company subject to rate-of-return regulation as practiced by state public service commissions. The model was developed to serve as the new electric

*Transferring a model to an independent computer system and reproducing previously published results is a good test. This process, sometimes called "model verification" (House and McLeod 1976), is equally useful in bolstering confidence in a package of programs such as the LHS sensitivity testing package.

utility sector of the FOSSIL2 modeling system at the Department of Energy. FOSSIL2 is a system dynamics model of the nation's supply and demand for energy used in policy analysis at the Department of Energy (NEP II 1979; EEA 1980). The new electric utility sector was constructed by adapting a model of an individual utility company to perform the nationwide calculations needed in FOSSIL2. Full technical details of the model of a single electric utility company are given in a technical report from SRI International (Yabroff and Ford 1980).

Figure 1 shows the new electric utility sector's projection of the oil and gas used by the nation's electric utility companies. The model is initialized in 1950 and simulates thirty years of historical behavior before moving to the projections for the 1980s. Figure 1 provides a comparison of model and industry behavior during this 30-year period. Cases A, B, and C shown in Fig. 1 differ in the reserve margins projected by the model for the 1980s. Case A represents a vigorous building program in which utility companies maintain the high reserve margins of the late 1970s. In Case B, reserve margins decline from the high levels, but not all the way to the 20% level of the late 1960s. In Case C, construction programs are severely

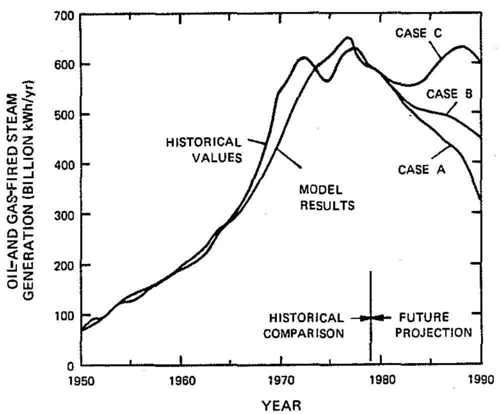


Fig. 1. Three projections of nationwide oil- and gas-fired steam generation.

limited, and reserve margins decline to the 20% level by the end of the decade. Depending on which of these cases is viewed as most representative of the nation's oil and gas burning utilities, nationwide oil and gas usage could be cut approximately in half by the end of the decade, or it could be no lower than it is today.

The three projections shown in Fig. 1 are somewhat typical of the informal sensitivity testing that is performed with system dynamics (and other) models. The analysts uses his or her own judgment to find the most important inputs and presents several simulations to demonstrate the importance of a particular parameter or policy. In this paper, we demonstrate what can be done to move beyond these informal procedures. We assume that one is interested in obtaining a tolerance interval on the oil and gas projections like those shown in Fig. 1.* This may be done through iterative application of the sensitivity testing procedures.

B. Iterative Application of LHS Sensitivity Testing

The application of LHS procedures yields some information on the range of possible values of a given output variable, but one cannot interpret the range probabilistically unless the many inputs to the model are independent. Given the complexity of energy systems, one must expect that there will be hundreds of interdependencies among the numerous inputs to any moderate-sized energy model. The key question, therefore, is whether the many interdependencies among model inputs are important impediments to obtaining a probabilistic interpretation of the range of values on a given output variable. Figure 2 gives an overview of the procedures we recommend for answering this question.

The analysis begins with the application of the LHS procedure to obtain an initial estimate of the confidence bounds on model output and a list of the most important parameters of the model. The model user then determines whether the most important parameters are truly independent. If they are, we assume that one may ignore whatever interdependencies may exist between less important input parameters and proceed to interpret the confidence bounds in probabilistic terms. If they are not, the model user alters the model to

*As Mass and Senge (1980) explain, at least three criteria are possible in a model testing process: (1) changes in the predicted numerical values of the model, (2) changes in the behavior mode of the model, and (3) changes in the policy recommendations drawn from the model. The illustrative calculations presented in this paper adopt the first criterion.

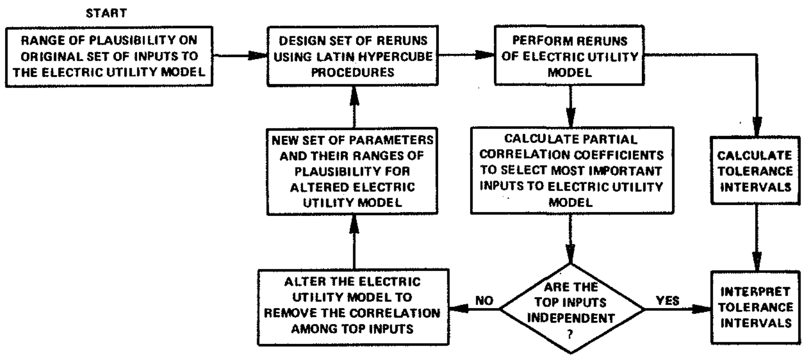


Fig. 2. Overview of the iterative application of Latin Hypercube Sampling to obtain interpretable tolerance intervals on model output.

remove the correlation among the top inputs. With the alterations, one obtains a new model, a new set of inputs, and the sensitivity testing must begin again with a new application of the LHS procedures. The iterations in Fig. 2 are repeated until confidence bounds are obtained for a model whose key input parameters are judged to be uncorrelated with one another.

C. Results from the First Iteration

We begin with a list of 45 parameters and their associated ranges of plausibility listed in App. B. The appendix gives the name of each input, the nominal values, their definitions, and our estimate of their ranges of plausibility. The first five parameters characterize the growth and shape of the demand for electricity, and the next two influence the way in which the price of electricity is regulated by state commissions. Parameters 8-12 influence the outcome for the utility company choice of the amount and kind of power plants to build in the future. Parameters 13-40 give the specific attributes for each of the generating technologies used in the model. The final set of parameters characterizes the transmission, distribution, and hydroelectric components of the electric utility model.

With the parameter ranges as a starting point, a set of twenty simulation experiments were designed using the LHS rules to ensure full

coverage of the 45-dimensional input space. The final result of the sampling analysis is a set of instructions for twenty* computer simulations with different parameter values for each of the 45 parameters considered uncertain. The information obtained from these twenty simulations is summarized in Fig. 3A.

Figure 3A reports the summary statistics for the first iteration analysis of the Oil and Gas Used in Electricity Generation (OUEG). Figure 3 shows the mean, maximum, and minimum results from the twenty simulation experiments. The variability among the different simulations is apparent from comparing the minimum and maximum values and also from the behavior of the standard deviation over time. The "nominal" values shown in Fig. 3A result when the model is run with all parameters taking on the "nominal values" reported in App. B. These summary statistics show that the nominal and mean results are quite close and that the maximum value is almost twice as large as the mean in the year 1990. Notice that the Fig. 3A information begins in the year 1980--the first year of the model projections into the future. Thus, the ranges of plausibility on input parameters (Appendix B) must be expressed in terms of an uncertain estimate of parameters in future years. (We do not necessarily agree with Carsten Tank-Nielsen (1980, p. 195) that a parameter change that "destroys the history fit of the model should not be viewed as a reasonable change.")

Figure 3B shows the tolerance intervals obtained from the first iteration analysis of OUEG. These limits encompass the range of values that could be expected in either 75% or 90% of the simulation runs of the model. The 90% tolerance interval in 1990, for example, ranges from a low of around 150 billions kWh/yr to a high of around 950 billion kWh/yr. The interval is largest around 1987 and decreases in size thereafter.

*To learn how many simulations are required, one may simply repeat the sensitivity analysis with a larger number of runs. If the new analysis yields the same set of tolerance intervals and the same set of partial correlation coefficients, one need not worry about the sample size. In a more detailed description of the tolerance interval calculations, Ford and McKay (1982) show that sensitivity analyses with 100 runs yields the same general results as the analyses with 20 runs shown here.

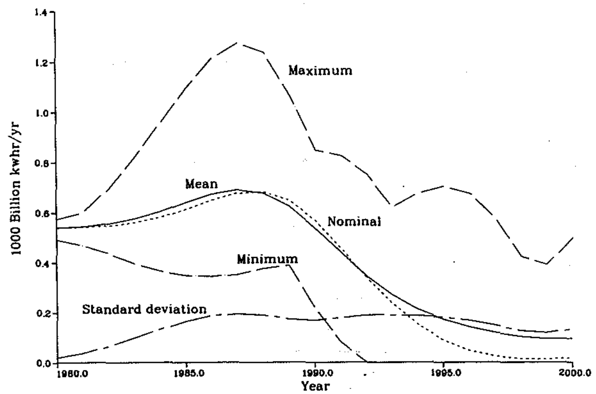


Fig. 3A. Summary statistics from the first iteration analysis.

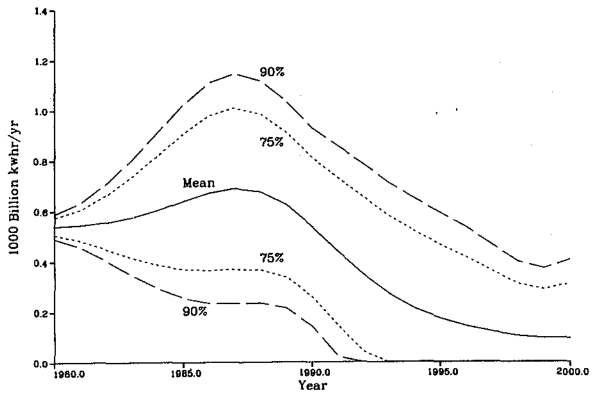


Fig. 3B. Tolerance intervals from the first iteration alaysis.

Figure 3C gives the partial correlation coefficients* between the value of OUEG in a given year and the values assigned to the more important input parameters. Strong positive or negative correlation indicates that the particular input parameter is especially influencial during that time period. Figure 3C shows that the Indicated Demand Growth Rate Constant (IDCRC) is positively correlated with OUEG in the 1980s and negatively in the 1990s. Higher growth rates in electricity demand lead to higher oil and gas usage during the 1980s because the model is limited to the number of new coal and nuclear power plants that will come on-line. By the 1990s, however, faster growth in electricity demand leads to less dependence on oil and gas power plants because faster growth prompts the model to invest more heavily in new coal or nuclear plants. Once these plants come on line in the 1990s, the older oil and gas burning plants are phased out of service. Figure 3C shows that the inflation rate (INFLR) is also highly correlated with OUEG, but in a pattern the opposite of the demand growth rate constant. A third input which the exhibits strong influence on OUEG is the desired reserve margin constant. The effect of changes in this input were revealed previously in Fig. 1. Higher reserve margin targets correspond to an overbuilding program designed to bring larger numbers of coal and nuclear power plants on-line to displace oil and gas. Thus, a larger DRMC leads to less OUEG in the 1990s once the extra coal and nuclear plants are operating. The lower portion of Fig. 3C shows three additional inputs found to have strong influence on OUEG during the 1980s: the availability factor for coal plants (NCAFC), for nuclear plants (LWAFC), and the coal plant operating lifetime (NCCL). Each of these inputs is negatively correlated with OUEG during the 1980s but shows little influence

---

*Conference participants should not confuse the partial correlation coefficients shown in Fig. 3C with the coefficients obtained in standard statistical tests. As explained by Mass and Senge (1980), the partial correlation coefficient is normally used to provide the modeler with a "measure of the incremental contribution of a single right-hand side ("explanatory") variable in accounting for variation in a dependent variable." Mass and Senge criticise the use of the partial correlation coefficient in such "single equation tests" as unreliable because "single-equation statistical tests focus on hypothesized relationships in isolation from the context of feedback relationships in which they are embedded." They advocate "full model behavior tests" as a more reliable indicator of the importance of a particular input parameter. The partial correlation coefficients shown in Fig. 3B provide a statistical summary of the apparent influence of a particular input in numerous "full model behavior tests."

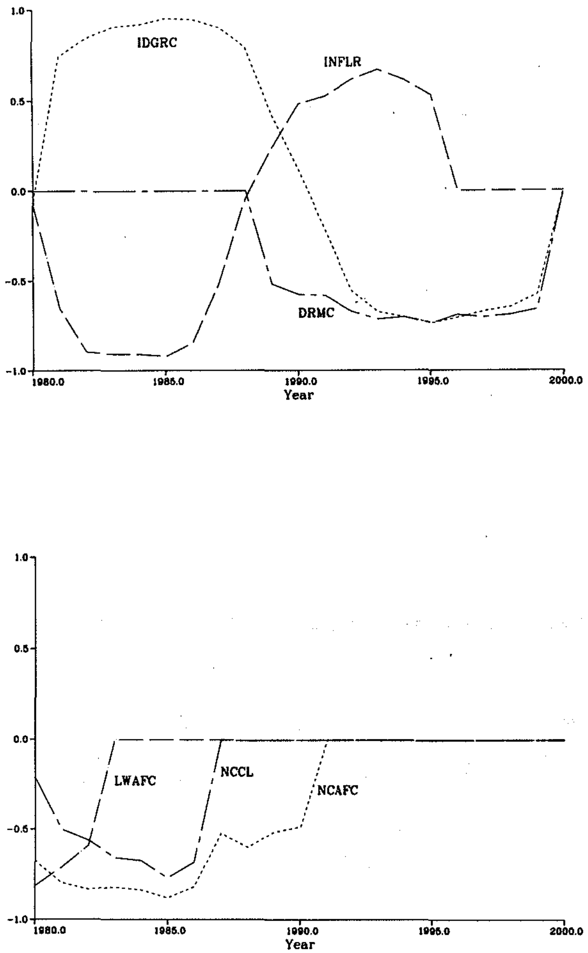


Fig. 3C. Partial correlation from the first iteration analysis.

after 1990 because of the model's internally generated capacity expansion plants that adjust construction to account for different lifetimes and availabilities.

The Fig. 3 results make good sense. No spurious tendencies have revealed themselves in this collection of runs. One can only interpret the tolerance intervals in Fig. 3B in probabilistic terms, however, if the most important inputs to the model are uncorrelated. This is not the case. Two collinearities exist between the top six inputs identified in Fig. 3C. First, the desired reserve margin cannot be specified independently from the availability factors for the nuclear power plants and the coal-fired power plants. Should the availability factors decline, for example, the utility company would compensate by increasing the desired reserve margin target used in capacity expansion planning. The second collinearity involves the availability factors for the coal and nuclear power plants which should be positively correlated as both types of plants have certain components in common. Following the approach diagrammed in Fig. 2, the next step is to remove these correlations through alterations in the electric utility simulation model.

To remove the collinearity between DRMC and the two availability factors, we changed the model to calculate the desired reserve margin as the sum of a Minimum Reserve Margin from Availability Factor (MRMAF) and a Reserve Margin Over Building Increment (RMOBI). The portion of the desired reserve margin which is dependent on the availability factors of the new coal and nuclear plants is calculated internally. The overbuilding increment is a new parameter which is varied to reflect the inclination of utility companies to overbuild to displace oil and gas. This new parameter, RMOBI, is not correlated with the availability factor for the new coal and nuclear power plants. To remove the collinearity between the two availability factors, we have introduced three new parameters--a steam power plant availability factor, an incremental difference between coal plant availability and the steam plant availability, and an incremental difference between nuclear plant availabiltiy and the steam plant availability. These three parameters are now inputs to the electric utility model, and while the actual availability factor for a new coal or nuclear power plant is calculated internally.

D. Results from the Second Iteration

We begin the second iteration with a somewhat different list of input parameters than is shown in App. B. The DRMC input is replaced by RMOBI, for example. Also, the two availability factors, LWAFC and NCAFC, no longer appear as inputs to the model. Instead, we have three input parameters needed for the model's calculation of the availability factors for new coal and nuclear power plants. All told, these changes result in a list of 46 input parameters, the majority of which are the same as those listed in App. B. LHS was used to design a set of twenty simulation experiments with the model that would cover the 46-dimensional input space. The results from the new set of twenty simulations are shown in Fig. 4.

Figure 4A reports the summary statistics for the second iteration analysis of OUEG. A comparison of Figs. 3A and 4A shows that the maximum value of OUEG around 1987 is lower in the second iteration. Also, the standard deviation of OUEG over the twenty simulations is generally smaller in the second iteration. Thus, one would expect the tolerance intervals to be somewhat narrower in the second iteration analysis. Figure 4B shows that the tolerance intervals do become narrower with the altered model. The 90% coverage in 1987, for example, runs from around 300 to 950 billion kWh/yr in the second iteration (versus 250 to 1150 billion kWh/yr in the first iteration). The reduction in the size of the tolerance interval from one iteration to the next may be attributed to the removal of the collinearities between the most important inputs to the model. With the collinearities removed, the twenty simulation experiments are less likely to be defined with extreme sets of inputs that would lead to unusually high dependence on oil and gas for electric power generation. The twenty simulations in the second iteration will not encounter a situation where the desired reserve margin is set very low even though, for example, the model is specifying poor availability factors for new coal and nuclear power plants.

Figure 4C gives the partial correlation coefficients between the value of OUEG in a given year and the values assigned to the more important input parameters. Several of the inputs that were selected in the first iteration analysis appear again in the second iteration (IDGRC, INFLR, NCCL, for example). Figure 4C also shows that the new parameters (RMOBI, NCAFD, LWAFD) created in the alteration of the electric utility model now appear in the list of more important inputs. An important result from Fig. 4C is that the six

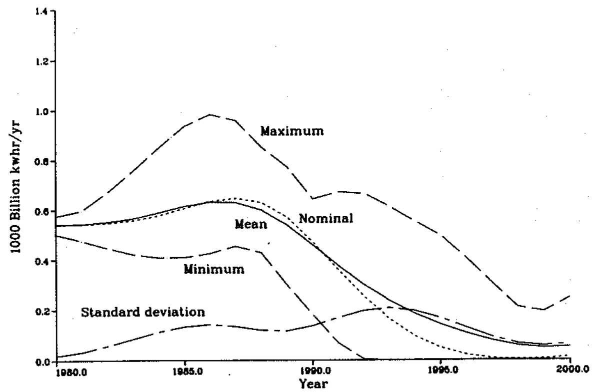


Fig. 4A. Summary statistics from the second iteration analysis.

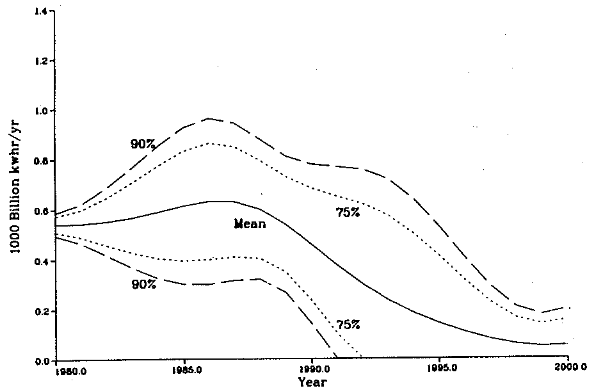
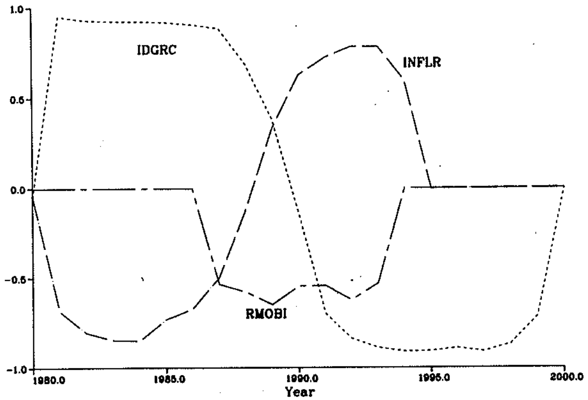


Fig. 4B. Tolerance intervals from the second iteration analysis.



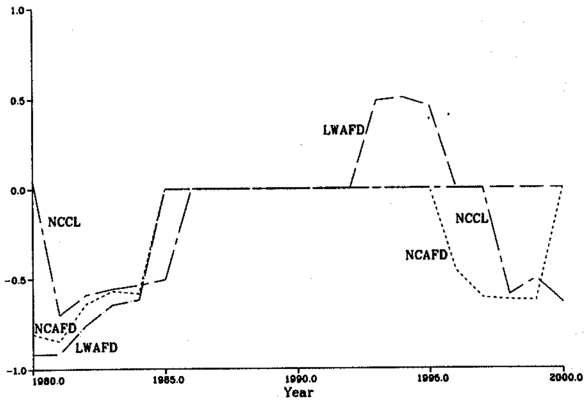


Fig. 4C. Partial correlation from the second iteration analysis.

variables selected as having most influence on OUEG are not correlated with one another in an important manner. Thus, we are free at this point to interpret the tolerance intervals in Fig. 4B in probabilistic terms.

E. A Measure of Parameter Uncertainty

The range of variation in the model projections of OUEG is represented by the tolerance intervals in Fig. 4B. The mean value of the forecasts is bordered by two sets of curves representing 75% and 90% coverages. Thus, one can readily see the uncertainty in OUEG forecasts due to parameter uncertainty. An examination of the graphs in the year 1990, for example, shows the mean value to be about 480 billion kWh/yr. We expect 75% of the OUEG forecasts to lie between 240 and 690 billion kWh/yr and 90% of the forecasts to lie between 150 and 780 billion kWh/yr. These intervals are calculated at the 95% confidence level. Thus, the probability that they are not sufficiently large is 5%.

We reemphasize that the tolerance intervals in Fig. 4B represent only the parameter uncertainty in the model forecast. That is, they represent the uncertainty in OUEG forecasts given that one accepts the structure* of the electric utility model as an accurate representation of the nation's electric utility industry.

III. USES OF SENSITIVITY TESTING RESULTS

The iterative application of these sensitivity testing procedures to obtain tolerance intervals is one of six useful applications that should

*The distinction between parameter uncertainty and structural uncertainty is not clear cut, especially when certain key parameters have "structural implications" (Tank-Nielsen 1980) such as the possible removal or inclusion of feedback loops. To our knowledge, the only systematic attempt to gauge the importance of structural uncertainty in energy models is the series of "forum" exercises sponsored by the Electric Power Research Institute. In the forum format, several models of the same system are operated with a commonly specified question and set of parameters. Differences in model performance are interpreted by the forum participants to gain an understanding of the effect of different model designs. In the Utility Modeling Forum's "Case Study Comparison of Utility Corporate Models," for example, a dozen corporate models were compared in terms of their analysis of the effects of company investment in customer conservation. Differences in the model calculations were attributed to differences in model structure because all models were exercised with a common set of parameter values such as the cost and savings from conservation investments (Shaw 1981).

bolster confidence in a model. Depending on the model purpose, the investigator may be interested in any or all of the six applications which we discuss below.

A. Model Shakedown

The process of running a model twenty times, or fifty times, or one hundred times with input parameters set at positions far removed from the typical "base case" conditions is a stiff test for any model to pass. Model builders are used to anticipating and eliminating bugs* that appear when the model is operated with the vast majority of the inputs set at their nomial value. Subjecting a model to the sensitivity testing procedures described here can reveal bugs that would otherwise go undetected in informal sensitivity testing.

B. New Behavior Modes

By simply displaying the results of numerous runs, one can sometimes discover a new mode of behavior. In the sensitivity test of the COAL2 model, for example, we discovered that the average price of electricity could decline in the future. This unexpected pattern was revealed in the display of results shown in Fig. 5. Here, the average price of energy from the first 45 of 100 runs of the COAL2 model are plotted over time. To pin down the reasons for the decline in price, we examined the input values for those parameters whose partial correlation coefficients stood out in the COAL2 test. We were interested in whether the particular simulations with a decline in price had any one input feature in common. It turned out that the common elements were a low capital cost for synthetic fuels facilities and a high propensity for oil companies to invest in synthetic fuels. Under these less likely (but plausible) conditions, the model exhibited a decline in average price due to the substantial production of low cost synthetic fuels.

C. Important Inputs

A third application of the sensitivity procedures is to isolate the inputs that have the most influence on a particular output variable--the motivating application for McKay's original research on sensitivity analyses. As illustrated in Figs. 3C and 4C, indications of the relative importance of

*"Bugs" may include division by zero, uncontrolled oscilations, the "DT problem" or a variable becoming negative when only positive values make sense.

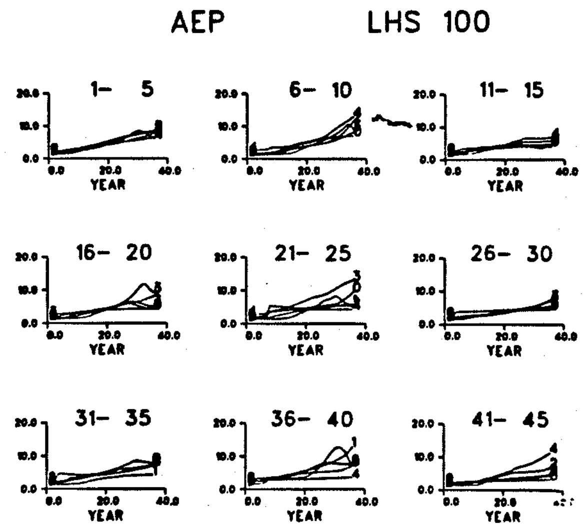


Fig. 5. Time series plots showing Average Energy Price in runs 1-45 of the Latin Hypercube Sample of runs of the COAL2 model.

different inputs are found by using a procedure similar to step-up regression with rank transformed data. We use the partial rank correlation coefficient (PRCC) with critical values from the ordinary correlation coefficient from normal theory to select potentially important inputs at each time point of the dynamic simulation. The sets of selected inputs in neighboring time points are compared, and inputs are added or deleted at a specific time point depending on their occurrence in the sets of neighboring time points. In this way, the analysis depicts those inputs that exhibit a strong influence on a particular output variable that persists over several time periods. Time plots of the PRCCs allow the analyst to select the most influencial inputs during different parts of the simulation and to determine the polarity of influence. From Fig. 4C, for example, one learns that the indicated demand growth rate constant IDGRC and the inflation rate INFLR appear to have the most influence on the oil and gas use by utilities in the mid 1990s. To put the size of the partial correlation coefficients shown in Fig. 4C into perspective, one can obtain cross plots for a particular year. In Fig. 6, for example, the value of OUEG in year 1994 is displayed relative to the input value selected for four of the inputs selected in Fig. 4C. The cross plots show examples of strong negative correlation (IDGRC), strong positive correlation (INFLRC), and examples of relatively little correlation (RMOBI and NCAFD).

D. Tolerance Intervals on Forecasts

A fourth application is the one illustrated in this paper--to present tolerance intervals on the parameter uncertainty in model projections. To

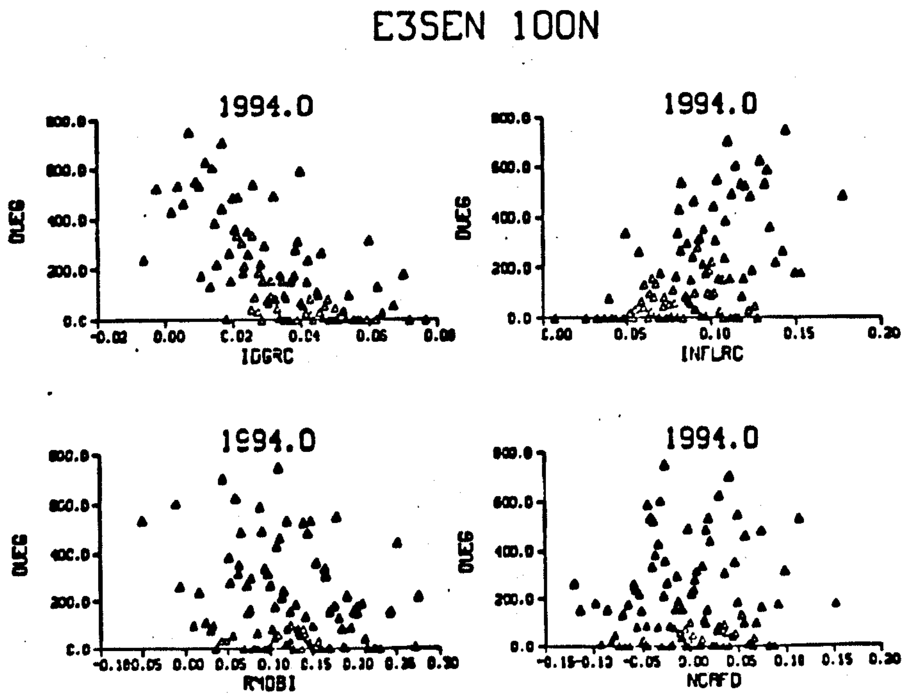


Fig. 6. Cross plots of oil and gas use in electricity generation OUEG with four input parameters whose partial correlation coefficients are shown in Fig. 4C. Cross plots refer to results in the year 1994 in 100 runs of the electric utility model in the second iteration analysis.

those participants who feel that system dynamics models are more appropriately used for policy analysis (rather than forecasting), we would note that the distinction between policy analysis and forecasting becomes extremely clouded when models are used in the day-to-day planning of large agencies or corporations. For more details on the iterative procedures needed to obtain tolerance intervals on model projections, conference participants are referred to the paper by Ford and McKay (1982).

E. Tolerance Intervals on Policy Tests

Policy relevant simulation results are usually obtained by comparing two model projections. One projection is obtained with "base case" or "business as usual" conditions while the second is generated with a different value for a set of parameters that describe the policy of interest. In reports and papers with which we are familiar, such policy results are often presented with the reassuring statement that the policy results are "robust" (they do not vary with changes in the many uncertain parameters). To verify the "robustness" of important policy results, one may repeat the sensitivity testing procedures described here with the provision that the model "output" of interest is simply the difference between the two relevant simulation runs.

F. Hitting Preselected Targets

A final application of the sensitivity testing procedures to be mentioned here is the examination of time plots of numerous runs to see if a particular output will hit a preselected target or follow a certain trajectory. The target may be the result of model projections obtained from a different department where more detailed calculations are available. Targets are sometimes imposed on the short-term portion of long-term models when agencies or companies maintain entirely separate models to address different issues.

In some cases, hitting a preselected target is an absurd exercise that should not be attempted. In other cases, it is merely a time consuming aspect of performing policy relevant analysis in such a way as to gain impact on the agency's or company's deliberations. In those cases where one wishes to learn what set of parameter assumptions could be employed to hit a preselected target, the display of numerous runs (like those shown in Fig. 5) can be quite helpful. Results of numerous simulations from a previous sensitivity test are simply displayed for visual examination, and the collection of runs that could hit the desired target are known by inspection. This particular application of the sensitivity testing procedures is especially valuable when the preselected target is impossible to hit with simple parameter changes.

IV. SUGGESTIONS FOR PRACTICAL APPLICATION

We feel that the combined research efforts of the Los Alamos National Laboratory and the Control Data Corporation have led to a practical approach to sensitivity testing that all system dynamicists would usefully apply in their major modeling projects. For those who wish to gain the advantages of detailed sensitivity testing, we offer three suggestions for practical application:

1. Input Ranges:

   Expect that the task of specifying the ranges of plausibility on model inputs will be a difficult initial obstacle, especially for large models that may have outgrown their original documentation.

2. Model Shakedown:

   Be prepared to observe spurious behavior when the model is run many times with Latin Hypercube Sample design on the input parameters. Accept the needed changes in the model as a problem with the model structure and not a problem with the sensitivity testing procedures.

3. Sensitivity of the Sensitivity Testing Results:

   Be prepared to test the results of the sensitivity analysis to changes in the starting assumptions (like those shown in App. B) if you are unsure about the description of input parameter uncertainties. You may be unsure, for example, of whether a uniform or a normal distribution best characterizes the uncertainty in a given input. Rather than devoting scarse resources to a detailed examination of such a question, one can simply repeat the sensitivity analysis to see if the tolerance intervals or the partial correlation coefficients are affected with a change in the probability distribution.*

   Another question that is easily answered by repeated applications of the sensitivity testing procedures is how many simulation runs are required to cover the input space. Rather than wrestle with this question analytically, we suggest that one simply repeat the analysis with a larger sample size to see if there are any important changes in the findings.

*In a previous paper, Ford and McKay (1982) show that all the uniform distributions noted in App. B could be changed to an equivalent normal distribution without changing the tolerance intervals or partial correlation coefficients shown here.

Investigators willing to follow these practical suggestions should be able to perform the type of analysis shown here without incurring significant computer costs. The computer related costs of the Fig. 3 calculations with 20 runs of the electric utility model cost about $50, for example. Half of the cost is due to the single compilation and twenty runs of the model. The remaining cost components include telephone linkage to the Dartmouth Time Sharing System ($6), connect time ($4), compilation of the HYPERSENS code ($2), Latin Hypercube sample design for 20 runs with 45 uncertain input ($7), and statistical analysis and display of the output from the 20 runs ($6).

We note in conclusion that the procedures described here are particularly valuable for system dynamics modeling projects because of their inclusion of many highly uncertain parameters needed to close feedback loops. The procedures are totally statistical in nature, however, and they do not rely on any apriori knowledge of the structure of the model. Conference participants interested in other sensitivity methods specifically tailored to system dynamics models are referred to the summary description in App. A.

APPENDIX A

SUMMARY OF RELATED RESEARCH

The statistical procedures described here are but one of several methods to assist in the sensitivity testing of system dynamics models. This appendix provides a brief review of ten alternative approaches. We begin with five approaches developed in the United States and listed in Table A-1.

Table A-1 begins with the statistical approach described in the body of this paper. Because of the use of Latin Hypercube Sampling procedures, Control Data Corporation analysts have chosen to refer to their computer code as HYPERSENS (Amlin 1982).

The second method listed in Table A-1 was developed by Joseph Talavage (1980, 1981) to allow calculation of the eigenvalues of a system dynamics model. Talavage refers to this approach as MODSENS (for MODal SENSitivity) because of the emphasis on the dominating modes of behavior that are more easily discovered through application of his computer code. The approach is "to judiciously select a small set of points in the state space at which the system can be linearized, and then to use efficient procedures of analysis at those points to gain insights into system behavior" (Talavage 1982, p. 2). Once a "base case run" is available, Talavage suggests that "one small set of state space points useful for analysis can be obtained from the values of the system state at, say, five-year intervals along the nominal trajectory." Once a linear system of equations is obtained, MODSENS performs the matrix manipulations needed to obtain the eigenvalues for the model. In a case study application to the electric utility model discussed in the body of this paper, for example, Talavage finds 64 eigenvalues associated with the system state in 1980--the initial year of the simulation. Talavage argues that these eigenvalues lend insight into the likely behavior modes of the model. Talavage noted, for example, that only 2 of the 64 eigenvalues had positive real parts indicating that "there is little opportunity for growth in the electric utility model." Talavage also noted that "there are several possibilities present for system instability (represented by oscillatory modes)...but that, in most cases, the time constant of the real part of these complex modes is so small that they would have little or no effect on longer-term behavior."

TABLE A-1. Five Projects on System Dynamics Sensitivity Testing in the United States

Name or Acronym	Research Group	Case Studies	References
1. HYPERSENS	Los Alamos National Laboratory Control Data Corporation	Electric Utility Planning Model	This Paper
2. MODSENS	Purdue University	Electric Utility Planning Model	(Talavage 1980, 1981)
3. Zero Stability Sensitivity	University of Minnesota	8th order, linearized production-inventory model and a 12th order nonlinear urban model	(Starr and Pouplard 1981)
4. Probabilistic System Dynamics	The Futures Group	Electric Utility Planning Model	(Stover 1978)
5. GPSIE/FIMLOF	Massachusetts Institute of Technology	9th order, nonlinear market growth model	(Peterson 1980)

Talavage identifies a separate "mode" of behavior with each of the 64 eigenvalues and then seeks to determine which of the modes dominate the overall model behavior. To distinguish between important and unimportant modes, Talavage calculates a "mode-magnitude" based on the contribution of an individual mode to the value of a particular state (level) variable of interest. In the case study, Talavage concentrated on the ten most important modes influencing two or three level variables of interest. To determine which of the input parameters have most influence on model behavior, Talavage repeated the calculation of eigenvalues, mode magnitudes, and dominant modes with 1% perturbations in the inputs.

The case study application of MODSENS to the electric utility planning model (which was also the subject of the statistical analysis reported in the body of this paper) allowed researchers from Los Alamos, Control Data Corporation, and Purdue University to compare the findings from two quite different approaches to sensitivity testing. We were particularly interested in learning whether input parameters judged most important from the statistical approach would also prove to be most influencial in altering the mode magnitudes calculated from MODSENS. Unfortunately, we were not able to draw strong conclusions from the comparison because MODSENS results were only available for 1980--the first year of the simulation.

The third method in Table A-1 by Starr and Pouplard (1981) is similar to Talavage's approach in that the eigenvalues of a linearized system dynamics model are the focus of the sensitivity study. In the third approach, however, the investigator is interested in the input parameters that have NO influence on the eigenvalues. Starr and Pouplard refer to this property as "Zero

Stability Sensitivity" and demonstrate that the unimportant parameters can be located by inspection using graphical methods involving "first order cuts" in the diagraphs of the linearized model. Starr and Pouplard illustrate with two examples how the analytically based graphical procedure can help one locate unimportant input parameters without performing any model simulations.

The first illustration involves an 8th order linearized version of a production-inventory model. The search for parameters with no influence on the eigenvalues begins by putting the equations into reduced form and drawing the diagraph representation of the equations. Starr and Pouplard define "type one cuts" which separate the diagraph into appropriate subsections. Arcs in the diagraph that are intersected by "type one cuts" indicate coefficients in the reduced form equations that can be arbitrarily altered without affecting the eigenvalues. In the second illustration, Starr and Pouplard examine a 12th order nonlinear model of urban interactions between affuent and poor population groups. The investigators compared the conclusions drawn from their graphical inspection procedures with so-called "elasticity coefficients" obtained by simply testing the non-linear model through one-at-a-time changes in 14 of the input parameters. After comparing the results of these two tests, the authors concluded that "the inspection procedure not only yielded conclusions which corresponded to those found through successive simulations, but it also identified the source of the growth mode and traced its effects through the system" (Starr and Pouplard 1981, p. 380).

Probabilistic System Dynamics, the fourth method listed in Table A-1, allows the analyst to investigate the effect of uncertainty in parameter estimates and the uncertain timing of discrete events. The method has been used in several studies by The Futures Group of Glastonbury, Connecticut. In the illustrative application discussed here, the so-called ELECTRIC3 model (a predecessor to the electric utility planning model used in the body of this paper) was examined to determine the variability in model output due to both parameter uncertainty and event uncertainty (Stover 1978). The Futures Group used cross impact matrices to represent the conditional probabilities of each of 21 key events thought to be important in affecting the electric utility industry. These included a nuclear moratorium, discontinuation of the breeder program, and a moratorium on strip mining in some western states. The probability of each of the 21 events was dependent on the performance of variables in the system dynamics model (such as the amount of installed nuclear capacity) and on whether one of the other 20 events had occurred. The cross impact

matrix was linked to the deterministic ELECTRIC3 model in such a way that the electric utility model would react properly to the occurrence of an event. Should a nuclear moratorium be called, for example, the electric utility model would prohibit any new construction of nuclear power plants. To ascertain the effect of uncertainty in parameters and events on model output, the expanded model was run 40 times in Monte Carlo fashion. The Futures Group was interested in the variability of certain model projections over time. One interesting, but unexplained, result of this test application was that the deterministic projections of the original system dynamics model lay completely outside the interquartile range of the probabilistic runs.

The GPSIE/FIMLOF approach listed in the final row of Table A-1 refers to the General Purpose System Identifier and Evaluator computer program described by Peterson (1974) to implement the Full-Information Maximum Likelihood via Optimal Filtering method of parameter estmation. In a test application to a ninth order nonlinear system dynamics model of market growth, Peterson (1980) illustrates the improvement in parameter estimates obtained from GPSIE relative to those obtained from standard econometric tools such as ordinary least squares (OLS) and generalized least squares (GLS). Peterson choose this particular model to facilitate comparison with Senge's (1974) parameter estimates obtained from OLS and GLS using "synthetic data" generated by the market growth model itself. Peterson argues that the GPSIE/FIMLOF approach can not only be used to obtain better parameter estimates, but it can provide confidence bounds on the model projections. In this application, the GPSIE/FIMLOF approach provides for system dynamics models what Fair's (1980) approach provides for econometric models. Both approaches generate confidence bounds and both require calculations with the "raw data" used in parameter estimation. Thus, these two methods differ from the confidence bounds calculations shown in the body of this paper in which the user specified ranges of plausibility of each input are the starting point.

Table A-2 lists a second group of five studies which have been conducted outside the United States.* In the first study by Schreiber (1981), the search

*System dynamics research projects are sometimes characterized as belonging to the "classical school" based on Forrester's original concepts or the "European school" where the original ideas are extended to incorporate such diverse elements as catastrophe theory and thermodynamics (Wolstenholme and Holmes 1982). Our grouping of the two sets of projects in Tables A-1 and A-2 is merely for convenience and does not imply that the ten projects fall naturally into a "classical school" or a "European school" of thought on sensitivity testing.

TABLE A-2. Five Projects on System Dynamics Sensitivity Testing Outside the United States

Name or Acronym	Research Group	Case Studies	References
1. nonlinear, n-dimensional optimization through the evolution strategy	Technical University of Berlin	WORLD2	(Schreiber 1981)
2. structural stability	University of Sevilla	low order models of urban systems	(Aracil 1981A,B)
3. local parameter sensitivity through perturbation methods	University of Bradford	7th order inventory-production model and a "pseudo-model" ten times larger	(Sharp updated, Sharp 1976)
4. decomposition and linearization	University of Eindhoven	WORLD3	(Thissen 1978)
5. sensitivity functions	University of Pretoria	WORLD3	(Vermeulen and De Jongh 1977)

for important inputs is translated into a nonlinear optimization problem. From Schreiber's point of view, "sensitivity is above all a question of defining a metric function" which indicates when a change in an input parameter has produced an important change in the model behavior. Schreiber views sensitivity testing as an optimization problem in which various techniques for nonlinear, n-dimensional optimization are applicable. In his test application to the WORLD2 model, Schreiber used an evolution search strategy. The idea was to apply Darwin's theory of biological evolution as a powerful search algorithm based on the hypothesis that a carefully copied principle of mutation and selection is a basic element of a fast and stable search algorithm. Schreiber argues that the optimization problem can be solved to maximize the change in model output if one is looking for the most important inputs.

Alternatively, the procedures can be reversed if one is interested in the control of model output. If, for example, one is looking for the set of parameter values which cause the model output to closely follow a certain trajectory, one can specify the objective function as the difference between the model output and the trajectory. Running the optimization algorithm to minimize the objective function leads to insights as to which inputs provide the most control. In his test application to the WORLD2 model, for example, Schreiber found the parameter changes needed to ensure that world population would closely follow a trajectory with a smooth approach to a stable equilibrium (as opposed to the overshoot and collapse mode characteristic of many of the WORLD2 runs).

In the second approach listed in Table A-2, Javier Aracil of the University of Sevilla is interested in the equilibrium surfaces of a system dynamics model (Aracil 1981A,B). Aracil uses a method supplied by the theory of qualitative analysis of differential equations and the mathematical tools from bifurcation theory and catastrophe theory. Aracil argues that the application of these theories to study the structural stability of system dynamics models warrants further work--a point which he demonstrates through illustrative examples. In the first illustration with a simple model of business formations, Aracil finds an extreme divergence of behaviour due to small variations in the initial conditions. In the second illustration with a second order model of population and business interactions in an urban area, Aracil finds the equilibrium surface and analyses the type of stability at the equilibrium points. In a more recent application to a third order model of population/business/housing interactions, Aracil (1981) obtains results from equilibrium curves which are the same as those developed by the original investigators. Aracil emphasizes, however, that "the equilibrium curves have the limitation of showing only what happens in equilibrium disregarding the transient evolution."

The third row of Table A-2 refers to John Sharp's (undated, 1976) work at the University of Bradford. Sharp distinguishes between "local sensitivity theory" and "global sensitivity theory" and suggest that "the estimates of the local sensitivity coefficients can be used with a hill-climbing program to drive the system as far as possible from its initial position while the parameters and the initial values remain within the bounds prescribed" (Sharp 1976, p. 8). Sharp used perturbation methods to find the sensitivity of a simple production-inventory model with seven levels and 16 uncertain parameters. Sharp compared the results from the perturbation method with those obtained from Monte Carlo methods and concluded that the perturbation method gave generally accurate indications of "system robustness."

The general approach of decomposition and linearization suggested by Thissen (1978) has been applied by investigators from the University of Eindhoven in their analysis of the WORLD3 model. Although many of the tests of the WORLD3 model were specific to that particular model evaluation, some techniques were identified as being of generic value. First among these is decomposition. Here, the Eindhoven group has a mind breaking the model down into functional sectors, each one of which is examined separately. In their discussion of the WORLD3 model, for example, separate analyses of the population sector, the agricultural sector, and the capital accumulation sector are presented. A second technique that may be useful in sensitivity studies is linearization. This technique requires the investigator to replace the nonlinear model with a simpler model whose interrelationships are linearized in the neighborhood where the model is most likely to operate. A third procedure is to introduce major shocks in certain portions of the model and monitor the model's response. The purpose of this shock testing (called "falsification of state variables" by the Eindhoven group) is to uncover the general dynamic principles by which model behavior is governed (Thissen 1978, p. 189).

The final approach listed in Table A-2 is to calculate sensitivity functions based on the expected rate of change of the output variable with respect to each of the many input parameters. An example of this approach is the analysis of the sensitivity of the WORLD3 model by two mathematicians from the University of Pretoria (Vermeulen and De Jongh 1977). The sensitivity functions calculated in their approach are somewhat similar to the partial correlation coefficients described in the body of this paper. Both of these indicators provide a measure of the expected rate of change due to a particular input. The sensitivity functions, however, indicate sensitivity to a particular input when all other inputs are at their base case values. The partial correlation coefficients provide a measure of sensitivity when all other inputs are allowed to vary throughout their range of plausibility. Although the sensitivity functions provide only a limited feeling for the sensitivity, their calculation does not require the investigator to design a sample and to generate numerous simulations with the model.

APPENDIX B

UNCERTAINTY IN MODEL INPUT PARAMETERS

The 45 input parameters that are considered uncertain in the illustrative example are listed below. Nominal values are used to generate the "nominal" results in Fig. 3A. The range of uncertainty on each input is described as either a normal (N) or a uniform (U) distribution.

Sector	No.	Nominal Value	Line No.	Definition	Range	Comments
Demand	1.	IDGRC=.035	720	Indicated Demand Growth Rate Constant	N, m=.035 st. d.=.0175	Not Time Dependent
	2.	PEED=-1.0	780	Price Elasticity of Electricity Demand	U(-.5,-1.5)	Restricted Sample Area Depending on LRDAD (#5)
	3.	DLFC=.61	940	Demand Load Factor Constant	U(.59,.69)	Not Time Dependent
	4.	MILDT=960/ 1300	2400	Minimum Intermediate Load Duration	U(1000,1500)	
	5.	LRDAD=6	820	Long-Range Demand Adjustment Delay	N,m=6,σ=2	Restricted Sample Area Depending on PEED(#2)
Financial/ Regulatory	6.	INFLRC=.09	1025	Inflation Rate, Future	N, m=.09, St. d.=.03	
	7.	LRLC=1	4330	Length of the Regulatory Lag	N, m=1, St. d.=.25	
Planning Parameters	8.	LOSP=-10	3735	Logit Function Shaping Parameter	U(-7,-12)	
	9.	DRMC=.35	4190	Desired Reserve Margin Constant	U(.20,.45)	
	10.	ICFP=.55	3400	Intermediate Cap. Factor for Planning the Cap. Mix	U(.45,.60)	
	11.	BCFP=.80	3410	Base Capacity Factor for Planning the Cap. Mix	U(.65,.85)	
	12.	FAYER=5	3550	Forecast Averaging Period	U(3,7)	
Availability Factor	13.	NCAFC=.70	1670		U(.6,.8)	Not Time Dependent
	14.	LWAFC=.63	1630		U(.55,.75)	Not Time Dependent
	15.	OUAFC=.7	1690		U(.6,.8)	Should Not Be Important
Fixed Charge Rate	16.	NCFCR=.128	6000		N, m=.128, st. d.=.02	
	17.	LWFCRC=.128	6900		N, m=.105, st. d.=.04	

Direct Construction Costs	18.	NCCAC=500	6040		U(400,600)	Not Time Dependent, Direct Cost Only
	19.	LWCAC=550	6920		U(400,700)	Not Time Dependent Direct Cost Only
	20.	GTCDCC=100	7245		U(75,150)	
Delivered Price of Fuel Increase Rate	21.	NCFIR=.01	6060		U(0,.02)	
	22.	OUFIR=.03	6500		U(.01,06)	
	23.	LWFIR=.01	6940		U(-.01,.04)	
	24.	TFPF=2	7260		U(1.8,2.2)	
Plant Efficiency	25.	NCEFC=.32	6020		U(.3,.34)	
	26.	OUEFC=.32	6460		U(.3,.34)	
	27.	LWEF=.32	6880		U(.3,.34)	
	28.	GTEF=.33	7230		U(.3,.35)	
Plant Operating & Maintenance Costs	29.	NCOMCH=.82	6000		U(.7,1)	
	30.	OUOMCH=.20	6440		U(.1,.3)	
	31.	LWOMCH=.38	6880		U(.25,.5)	
	32.	GTOMCH=3.0	7230		U(2,4)	
Plant Lifetime	33.	NCCL=35	5840		U(30,45)	
	34.	OUCL=35	6230		U(30,45)	Not Time Dependent
	35.	LWCL=35	6710		U(30,45)	
	36.	GTCL=25	7090		U(20,35)	
Plant Construction Delay	37.	NCCD=4.5	5840		U(3,6)	
	38.	LWCD=6	6720		U(4,8)	
Pre-Construction Planning Delay	39.	NCPD=1.5	5840		U(1,2)	
	40.	LWPD=2	6730		U(1.5,3)	
Transmission and Distribution	41.	CDCCC=310	7710	Capital Cost of Trans. Capacity	U(250,500)	
	42.	CDCCD=34.11	7730	Capital Cost of Distribution	U(30,40)	
	43.	NTDL=35	7780	Nominal Trans. & Distribution Lifetime	U(30,45)	
	44.	ELTF=1.11	840	Electricity Losses Transmission Factor	U(1.09,1.13)	
Hydro Generation	45.	NYEGT=285/325	3290	Hydro Generation in 1990	U(285,385)	Ramp Increase from 80 to 90

REFERENCES

1. D. Abbey and R. Bivins, "Leasing Policy and the Rate of Petroleum Development: Analysis with a Monte Carlo Simulation Model," report No. LA-9272-MS of the Los Alamos National Laboratory, Los Alamos, New Mexico 87545, March 1982.

2. J. Amlin, "HYPERSENS: Sensitivity Analysis Using Latin Hypercube Sampling," report of the Control Data Corporation, Dayton, Ohio prepared for Andrew Ford, Los Alamos National Laboratory, November 1982.

3. Javier Aracil, "Further Results on Structural Stability of Urban Dynamics Models," System Dynamics and Analysis of Change, B. E. Paulre, Editor, North-Holland 1981.

4. Javier Aracil, "Structural Stability of Low-Order System Dynamics Models," International Journal of Systems Science, Vol. 12, No. 4, 1981.

5. "Computer Simulation Methods to Aid National Growth Policy," report of the US House of Representatives, Appendix A, Glossary of Terms, p. 129 (July 30, 1975).

6. Energy and Environmental Analysis, Inc. "FOSSIL2 Energy Policy Model Documentation," prepared by the US Department of Energy, Office of Analytical Services, DOE/PE/70143-01, October 1980.

7. R. C. Fair, "Estimating the Expected Predictive Accuracy of Econometric Models," International Economic Review, Vol. 21, No. 2, June 1980.

8. Andrew Ford and Peter Gardiner, "A New Measure of Sensitivity for Social System Simulation Models," IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-9, No. 3, March 1979.

9. A. Ford and M. McKay, "A Method of Quantifying Uncertainty in Energy Model Forecasts," report No. LA-UR-82-1423 of the Los Alamos National Labotatory, November 1982.

10. A. Ford, G. Moore, and M. McKay, "Sensitivity Analysis of Large Computer Models: A Case Study of the COAL2 National Energy Model," report No. LA-7772-MS of the Los Alamos National Labotatory, April 1979.

11. M. Greenberger, M. Crenson, and B. Crissey, Models in the Policy Process, Russel Sage Foundation, New York, 1976.

12. Peter W. House and Jon McLeod, Large Scale Models for Policy Evaluation, (Wiley-Interscience, New York, 1976), p. 66.

13. Nathaniel Mass and Peter Senge, "Alternative Tests for Selecting Model Variables," Elements of the System Dynamics Method, Jorgen Randers, Editor, MIT Press, Cambridge, Mass. 1980.

14. M. D. McKay, W. J. Conover, and R. J. Beckman, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," Technometrics, Vol. 21, pp. 239-245, 1979.

15. M. D. McKay, W. J. Conover, and D. E. Whiteman, "Report on the Applications of Statistical Techniques to the Analysis of Computer Codes," report No. LA-NUREG-6526-MS of the Los Alamos National Laboratory, Los Alamos, New Mexico 87545, 1976.

16. M. D. McKay, Andrew Ford, G. H. Moore, and Kathleen H. Witte, "Sensitivity Analysis Techniques: A Case Study," Los Alamos Scientific Laboratory report No. LA-7549-MS (November 1978).

17. Roger F. Naill, "COAL1: A Dynamic Model for the Analysis of United States Energy Policy," Thayer School of Engineering, Dartmouth College report DSD-54 (February 1976).

18. Roger F. Naill, Managing the Energy Transition: A System Dynamics Search for Alternatives to Oil and Gas (Balinger Press, Cambridge, 1977).

19. Roger F. Naill and George A. Backus, "Evaluating the National Energy Plan," Technology Review, July/August 1977.

20. "National Energy Plan II," prepared by the Office of Policy and Evaluation, US Department of Energy (1979)--available from the Superintendent of Documents, US Government Printing Office, Washington, DC 20402.

21. David Peterson, "Statistical Tools for System Dynamics," Elements of the System Dynamics Method, Jorgen Randers, Editor, MIT Press, Cambridge, Mass. 1980.

22. David Peterson and Fred Schweppe, "Code for a General Purpose System Identifier and Evaluator: GPSIE," IEEE Transactions on Automatic Control, Vol. Ac-19, December 1974.

23. Bernd Schreiber, "Sensitivity Analysis and Control by Simulation," System Dynamics and Analysis of Change, B. E. Paulre, Editor, North-Holland, 1981.

24. Peter Senge, "Evaluating the Validity of Econometric Methods for Estimating and Testing of Dynamic Systems," System Dynamics Group memo D-1944-2, Alfred P. Sloan School of Management, MIT, Cambridge, Mass., February 12, 1974.

25. John Sharp, "Sensitivity Analysis Methods for System Dynamics Models," System Dynamics Research Group, University of Bradford, Bradford, England, no date.

26. John Sharp, "Some Comparisons of Sensitivity Testing Methods for System Dynamics Models," working memo of the System Dynamics Research Group, University of Bradford, Bradford, England, April 1976.

27. R. W. Shaw, Jr., Principal Investigator, "Case Study Comparison of Utility Corporate Models," report No. EA-2065 of the Electric Power Research Institute, October 1981.

28. Patrick Starr and Jean-Pierre Pouplard, "Identifying Critical Parameters in System Dynamics Models," System Dynamics and Analysis of Change, B. E. Paulre, Editor, North-Holland, 1981.

29. John Stover, Winslow Hayward, and Harold Becker, "Incorporating Uncertainty in Energy Supply Models," report EA-703 prepared for the Electric Power Research Institute, Palo Alto, California, February 1978.

30. Joseph Talavage, "Model Analysis to Aid in System Dynamics Simulation," to appear in a special issue of Management Science, 1980.

31. Joseph Talavage, "MODSENS, A Modal Control Approach to Determining System Sensitivity," report prepared for Andrew Ford, Los Alamos National Laboratory, Los Alamos, New Mexico, 1981.

32. Carsten Tank-Nielsen, "Sensitivity Analysis in System Dynamics," Elements of the System Dynamics Method, Jorgen Randers, Editor, MIT Press, Cambridge, Mass. 1980.

33. Wil Thissen, "Investigations into the World3 Model: Lessons for Understanding Complicated Models," IEEE Transactions on Systems, Man, and Cybernetics, March 1978.

34. P. J. Vermeulen and D. C. J. De Jongh, "Growth in a Finite World--A Comprehensive Sensitivity Analysis," Automatika 13, 77-84 (1977).

35. E. F. Wolstenholme and R. K. Homes, Editors, Dynamica, editorial in Vol. 8, part II, Winter 1982.

36. I. W. Yabroff and A. Ford, "An Electric Utility Policy and Planning Analysis Model," Final Report, SRI International, Menlo Park, California, December 1980.