REAL-WORLD VALIDATION OF SHAC MODELS

LOUISE MORRISON

JANUARY 1980

TO BE PRESENTED AT THE SECOND ANNUAL SYSTEMS SIMULATION AND ECONOMICS ANALYSIS CONFERENCE, BAHIA HOTEL, SAN DIEGO, CALIFORNIA, JANUARY 23-25, 1980

PREPARED UNDER TASK NO. 3525.40

Solar Energy Research Institute
1536 Cole Boulevard
Golden, Colorado 80401

A Division of Midwest Research Institute

Prepared for the
U.S. Department of Energy
Contract No. EG·77·C·01·4042

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.
NOTICE

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use or the results of such use of any information, apparatus, product, or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.
REAL-WORLD VALIDATION OF SHAC MODELS

Louise S. Morrison
Solar Energy Research Institute
Golden, Colorado 80401

ABSTRACT
This paper proposes a statistical approach to validation of SHAC models. It includes a definition of validation, an explanation of its purposes, and a description of the statistical aspects of experimental design. It proposes a study to validate design codes with statistical samples of real-world systems. Also included is a summary of present SHAC validation methodologies and studies as well as recommendations for future activity.

INTRODUCTION
In designing a validation study it is essential to consider carefully the following: (a) the definition of validation being used, (b) the user community and purposes of the validated model, (c) the method of selecting the test systems/components, (d) the instrumentation of the test systems/components, (e) the statistical aspects of the study including interpretation of results, (f) the generalization of results, and (g) the expected impact of the study.

PRELIMINARY CONCEPTS
A major decision in the experimental design of SHAC validation studies is the intended degree of control of the test system/component environment.

Controlled Data and Environment
In a controlled environment, design and performance data is obtained from a system/component that fits the model assumptions. The system/component is well designed, well installed, and fine-tuned for optimal performance. Load and weather data are collected at the site together with the performance data so that the driving forces of the model are known. A modeling expert using the actual weather and load data obtains the model prediction "after-the-fact." The input data to the model is referred to as controlled data.

Uncontrolled Data and Environment
In an uncontrolled environment, design and performance data is obtained from a system operating in a "real-world" or field environment. A model user of uncertain expertise predicts performance for a system that may only minimally fit the model assumption, whose quality of design and installation is uncertain, and whose existing performance level is not altered for the study. Load and weather data are not collected at the site; instead user-estimated or model-supplied data is used. The model prediction is usually made before the performance data is collected. The input data to the model is referred to as uncontrolled data.

The intent of most software/hardware studies has been to test a model with highly controlled data; however, systems malfunctions and instrumentation problems have limited the degree of control obtained. The intended degree of control has great impact, since the experimental design is radically different for a controlled approach than for an uncontrolled approach.

"After-the-Fact" Prediction
In validation studies using controlled data, "model prediction" is a misleading expression because the prediction is made "after-the-fact" using load and weather data that was collected at the site together with the performance data. This type of study verifies that the modeled performance agrees reasonably well with the real performance.

"In-Advance" Prediction
In validation studies using uncontrolled data, "model prediction" really means prediction. The model is used with estimates of load and weather data to forecast "in advance" how a system will perform for the next few years.

DEFINITION OF VALIDATION
A major problem in SHAC validation has been a lack of consensus about its definition and purposes. The definition of validation which will be used in this paper is as follows. A model is defined as validated for a group of systems/components operating in a particular environment if performance predictions are not significantly different from measured performance values for a sample of systems/components that has been selected from the set in a statistical manner. The degree of significance of results required for validation is a decision of the researcher and is addressed later in a section on experimental design.

*This work was supported by Systems Development Division, Office of Solar Applications, DOE.
According to this definition, no SHAC model has been validated with systems/components in general. However, each of the various SHAC models has been validated with a few systems/components operating in a controlled environment. (This is the default case, in which the set of test systems/components is a statistical sample of itself but is not selected to be a statistical sample of a larger set.)

VALIDATION PURPOSES

Models Validation with Controlled Data

Researchers, engineers, economists, and others desire systems and component models that have been validated with controlled data for purposes which include the following:

- to make comparative performance predictions similar to the EPA mileage ratings for cars;
- to compare system/component performance with standards such as BEPS (Building Energy Performance Standard);
- to perform sensitivity studies to optimize physical systems; and
- to examine system/component effects of shading, snow cover, etc.

In all these cases, generalization of results applied to uncontrolled data is uncertain, but some confidence is obtained.

Models Validated With Uncontrolled Data

Installers, distributors, architects, homeowners and others desire techniques which have been validated with uncontrolled data for purposes which include the following:

- to estimate the economic feasibility of an installed system; and
- to optimally size the system collector array and other components.

A STATISTICAL APPROACH TO EXPERIMENTAL DESIGN

A very brief discussion of this topic is presented here; for a detailed description see Cohen [11].

In a statistical approach an attempt is made to select data that is representative of some larger set so that one knows how much confidence can be placed in the results and so that generalization of results can be justified. No matter what data is used, there is an emphasis on careful experimental design.

Data Selection and Generalization of Results

Ideally, a model validator would identify a population of systems/components and select a rigorous statistical sample from the population according to random, stratified, or other sampling methods. For example, one sampling strategy would select randomly from all existing SHAC systems, another would weight the sample to be representative of a potential SHAC systems distribution. A decision would be made whether to collect data in a controlled or uncontrolled environment. Any restrictions imposed on the sample would be identified, such as systems which are working, systems that fit the model assumptions, etc. Results would generalize only to those systems/components in the population that meet the limitations imposed on the sample and that are operating in a similar environment.

Most validation studies do not state the method of selecting the test systems/components. Occasionally, one reads that systems were selected "randomly," but random selection has a precise mathematical meaning and should not be used unless actually employed. It is often assumed that results generalize even though it is not known how the test systems/components came to the attention of the validators.

Interpretation of Results

The results of a validation study can be interpreted statistically in a number of ways. For each system/component, a calculation of (predicted-measured)/predicted could be made for a parameter such as solar fraction. The mean and standard deviation of these values could be used to establish a confidence interval on model output. The size of the sample and the value of the confidence limits are interrelated. A discussion of these aspects of experimental design can be found in books on sampling theory.

It is often desirable to consult a statistician for assistance in experimental design and interpretation of results. Research today involves expertise in many disciplines in order to solve the really difficult problems.

Error Sources

Many sources of error contribute to differences between model prediction and measured performance, including error due to unmodelled parameters, inaccurate estimates for modelled parameters, inadequate measurement of performance, etc.

Sensitivity studies have been done with controlled data to determine confidence intervals for a model prediction due to individual sources of error. However, because of the magnitude and number of sources of error, it appears impossible to combine the contributions of individual sources of error to obtain a confidence interval for all error.

In a validation study with uncontrolled data, all error is combined in a single confidence interval because sources of error are not eliminated or minimized in the experimental design. This is desirable for "real-world" validation purposes since the homeowner or installer is primarily interested in the bottom line, which is the range of system performance that can be expected based on a model prediction.

STATISTICAL APPROACHES IN OTHER DISCIPLINES

Other disciplines recognize the need for a statistical approach for problems similar to validation. Opinion sampling in politics and market research is done with rigorous statistical techniques. Even so, results often do not generalize as expected; in a classic example, in the 1948 Presidential election Dewey was predicted as the victor...
over Truman when a telephone survey underrepresented the many farmers who had no phones.

Medical experimenters use double-blind experiments to minimize the unconscious biases of a researcher. In such an approach, neither the researcher nor the test subject knows whether the subject is in the experimental group or the control group. Solar researchers are naive if they think they can carry out validation studies in an unbiased manner without taking specific measures to minimize personal bias.

INSTRUMENTATION

Obtaining adequate data has been a significant problem in validation. Performance and data acquisition difficulties have occurred during data collection, and measurement error has been large. The cost of presently used instrumentation schemes has dictated that only a few types of systems/components be monitored. The average cost of installed instrumentation for a system monitored by the National Solar Data Network is approximately $40,000, and data processing of the information is prohibitively expensive.

Existing data may be able to serve many of the needs of statistically sound validation studies. However, new data may be required to permit the use of relatively large samples of systems/components. There should be an investigation of inexpensive instrumentation plans (often devised by those without federal funding) and a study of the requirements which can be met by data of reduced precision.

Examples of inexpensive instrumentation plans for validating design codes with SDHW systems are given in Cohen [1]. In one scheme, Btu meters are proposed that cost less than $530 and that have a claimed accuracy of ±3%. Drawbacks include potential difficulties in using inexpensive Btu meters, such as imprecise measurements of temperatures and flow rates, degradation due to time buildup, circuit failures, etc.

PERFORMANCE MEASUREMENTS

Solar fraction is the most common performance measurement utilized in validation studies. This indicator can be defined in many different ways for active systems, and its interpretation is especially difficult for passive systems. Because of the ambiguity of the expression, quantities such as solar contribution, auxiliary fuel contribution, or displaced fuel should be used in validation studies.

PROPOSED VALIDATION STUDY

A needed validation study would test the ability of design codes to predict "in-advance" for a typical solar space heating system operating under "real-world" conditions. An example outlining how this might be done is given in the following sections.

Data Selection

The population of residential SHAC systems in a region would be estimated by the utility companies. Meter readers would record addresses of domestic solar energy systems, and mailings would be sent to homeowners requesting their participation in the study. A firm experienced in statistical sampling would select a representative sample of the identified systems. Site visits would be made to determine systems performance according to model-independent criteria, and restrictions for excluding systems would be established. Results would generalize to all systems meeting the chosen criteria. Design and load data would be estimated using information available to the average model user. Weather data would be obtained for the year of the study from an appropriate design code weather data city.

Performance Predictors

For each design code, performance would be predicted "in-advance" using design weather data and "after-the-fact" using the weather data that occurred during the year of the study.

A group of engineers, architects, and others who have had SHAC experience in the region would be selected. Each person would use engineering judgment (without design code usage) to estimate performance for the set of systems. It would be determined if the members of the group have previously used a design code.

Other techniques used by the industry (such as "rules-of-thumb" and nomographs) would be identified. Performance predictors used in foreign countries would be appraised, since design codes are not applied universally. These identified techniques would be utilized to predict performance for the test systems.

The sample of systems would be instrumented as inexpensively as possible to obtain essential performance values. Measurements would be taken for the test systems, and comparisons would be made with the various predictors. A statistical analysis would be made of the results.

Analysis of Results

The HVAC industry has developed assumptions about the thermal performance of residential heating and cooling of buildings, and SHAC models have for the most part incorporated these basic assumptions. Software/software validation studies often result in the modification of the models to make the models agree more closely.

If the study shows that design codes are adequate predictors, it will justify their usage for "real-world" purposes. Otherwise, the research community should do one or more of the following:

- re-examine the basic assumptions of the design codes to determine which assumptions need to be extensively reworked to increase "real-world" predictive capability;
- develop new models with different assumptions that better serve "real-world" predictive use;
- recommend that design codes be used strictly for studying systems under controlled conditions;
- recommend the use of engineering judgment or other techniques for "real-world" predictive purposes; or


- encourage "real-world" usage of design codes for learning purposes only in developing engineering judgment.

Actually, most researchers agree that design codes are poor predictors of "real-world" performance. Precise prediction may be an impossible task for a SHAC model, since weather and load data is unpredictable. Some feel that the industry should be encouraged to use SHAC models until a better alternative is developed. Unfortunately, some members of the solar industry do not employ sufficient engineering judgment in design code application. Whether or not the study is implemented, disseminators should state clearly the limitations of design codes in making absolute predictions.

The study would indicate needed improvements in systems design, manufacture, installation, maintenance, and instrumentation. A sociological survey of the system owners would give better information. The data base of residential SHAC systems would be valuable for many applications. At present, there is not even a good estimate of how many residential systems are installed. A pilot project should be implemented to give experience in creating a statistically valid experimental design. From sampling theory, it has been estimated that approximately 20 systems would suffice for a statistical sample of SHAC systems in a state such as Colorado.

SHAC VALIDATION EFFORT

Methodologies

Theoretical approaches for validation of SHAC simulations and design codes have been proposed by Cohen [1], Kennish and Knasel [2], Lantz and Winn [3], Winn et al. [4], and others. Because validation is such a difficult and many-faceted problem, it is probable no single approach suffices.

Kennish's methodology addresses the validation of models with controlled data for "after-the-fact" prediction. He stresses a practical approach by recommending use of existing data, limited numbers of test systems/components, software/software studies, and component validation. Kennish is particularly interested in performing sensitivity analyses of design code inputs. There is little emphasis on statistical experimental design.

Cohen enlarges the validation picture to include studies with uncontrolled data for "in-advance" prediction and emphasizes a statistical approach in studies with either controlled or uncontrolled data. Cohen's major concern is careful experimental design. He recommends using statistical samples of systems/components and stresses software/hardware studies and system validation. Cohen presents an idealized approach which he recognizes may be difficult and expensive to implement, but he feels that the researcher should be aware of the ideal and recognize the extent to which actual studies depart from it. See Cohen [1] for a detailed description of Kennish's and Cohen's methodologies.

Studies

As stated previously, SHAC models have been validated with a limited number of systems/components operating in limited kinds of controlled environments. Generally, model predictions have compared favorably to measured performance when using controlled data; at times the models were modified until agreement was reached.

In most comparisons there has been little description of the experimental design and minimal statistical interpretation of results. In few, if any, studies have the test systems/components been selected as a statistical sample of a larger set taking into account factors such as geographical location, installer, system type, etc. Therefore, it has been difficult to determine the extent to which the results of these experiments generalize. In addition, studies have attempted to generalize from systems operating in a controlled environment to those in an uncontrolled environment. Sources of error, such as typical installation problems, have been missed. This has been a very frustrating approach and has made validation appear to require an unending series of studies.

Software/software studies have been done extensively, and models have generally compared well with each other when using controlled data or have been modified to do so. When models correlate highly, undue confidence in them sometimes results, even though none of the models has been validated with performance data.

Many validation studies have incorporated a defense of the models and a defense of solar technology. Attempts have been made to interpret reality to fit the model rather than the reverse. "Real-world" validation needs have been grossly ignored because validators have anticipated the many uncertainties of field systems.

The results of the validation studies give confidence that SHAC models can predict performance for systems/components operating in a controlled environment. However, they give little evidence that SHAC models can predict performance "in-advance" for field systems.

Bibliographies

SAI has summarized some of the SHAC validation efforts in Refs. [2] and [3] and ADL has compiled a validation bibliography [4]. Results of validation studies are often difficult to obtain since they are presented in memos, working group minutes, technical reports, and performance papers.

MAJOR RECOMMENDATIONS

- Validation studies should be done with greater emphasis on careful experimental design and on statistical interpretation of results.

- The first step in a validation study should be the preparation of a study plan which includes (a) the definition of validation being used, (b) the user community and purposes for which the validated model is intended, (c) the method of selecting the test systems/components, (d) the instrumentation of the test systems/components, (e) the statistical aspects of the study including interpretation of results, (f) the generalization of results, and (g) the expected impact of the study. This information also should be stated in the study results.
• Validation studies with controlled and uncontrolled data complement each other and studies of each type are needed. More software/hardware studies with controlled data are needed because of the poor quality and limited scope of the data used in present studies. Software/hardware studies with uncontrolled data, such as the study proposed here, should be implemented.

• Proposed software/software studies should be examined to ensure their potential contribution. These studies take computer and staff time, generate reports, and at times are little more than busy work.

• Sensitivity studies should concentrate on the sources of error, such as load estimates, which are of large magnitude.

• The use of solar fraction as a performance measurement in validation studies should be de-emphasized.

• Persons operating the National Solar Data Network (NSDW) and other data collection programs should consider defining populations of solar energy systems and instrumenting a statistical sample. They should also investigate inexpensive instrumentation schemes and data requirements which can be met by data of reduced precision.

• EPRI should encourage utility companies to participate in a nationwide census of residential solar energy systems.

• Researchers should be honest with disseminators in expressing the limitations of design codes to make absolute predictions for "real-world" systems.

• There should be increased communication between the research community and the solar industry. Model developers need industry input to design relevant and useful models. The industry needs input from the researchers to understand the available tools and how to apply them.

ACKNOWLEDGEMENT

The author wishes to acknowledge the inspiration of Joel Cohen, the assistance of L. M. Murphy, and the exchange of ideas with Bill Kennish.

REFERENCES


