CONF-850996--4

MASTER

NUCLEAR POWER PLANT TRAINING SIMULATOR FIDELITY ASSESSMENT*

CONF-850996--4

TI85 017082

Richard J. Carter
K. Ronald Laughery*

Oak Ridge National Laboratory
Oak Ridge, Tennessee

Presentation To Be Made At:

29th Human Factors Society Meeting
Baltimore, Maryland
September 29 - October 3, 1985

*Micro Analysis and Design
Boulder, Colorado

---

*Research sponsored by U. S. Nuclear Regulatory Commission Office of Nuclear Regulatory Research with Oak Ridge National Laboratory operated by Martin Marietta Energy Systems, Inc. under contract #DE-AC05-84OR21400 with U. S. Department of Energy.

NUCLEAR POWER PLANT TRAINING SIMULATOR FIDELITY ASSESSMENT*

Richard J. Carter

Oak Ridge National Laboratory
Oak Ridge, Tennessee

K. Ronald Laughery

Micro Analysis and Design
Boulder, Colorado

ABSTRACT

The fidelity assessment portion of a methodology for evaluating nuclear power plant simulation facilities in regard to their appropriateness for conducting the Nuclear Regulatory Commission's operating test was described. The need for fidelity assessment, data sources, and fidelity data to be collected are addressed. Fidelity data recording, collection, and analysis are discussed. The processes for drawing conclusions from the fidelity assessment and evaluating the adequacy of the simulator control-room layout were presented.

INTRODUCTION

The U.S. Nuclear Regulatory Commission (NRC) has proposed that revisions be made to Part 55 (Operating Tests) of Title 10 to the "Code of Federal Regulations" and to Regulatory Guide 1.149 (1984). If the rule changes are enacted, the operating test would be administered in a plant walk-through and in a simulation facility, which could be the plant, a plant-referenced simulator, or another simulation device, alone or in combination. During the simulation facility part of the operating test, reactor operators would be assessed on their ability to respond to normal plant operations and malfunctions in a realistic environment. The proposed modifications would require the facility licensee for each nuclear power unit to evaluate their simulation facility as to its appropriateness for the conduct of the operating test.

NRC's Office of Nuclear Regulatory Research contracted with Oak Ridge National Laboratory to develop a methodology for performing the simulation facility evaluations. The methodology is to be utilized during two phases of the life cycle: initial simulator acceptance and recurrent analysis. Initial evaluation is aimed at ensuring that the simulation facility provides an accurate representation of the reference plant. There are two components of initial simulator evaluation: fidelity assessment and a direct determination of the simulator's adequacy for operator testing. Recurrent evaluation is aimed at ensuring that the simulation facility continues to accurately represent the reference plant throughout the life of the simulator. It involves three components: monitoring reference plant changes, monitoring the simulator's hardware, and examining the data from actual plant transients as they occur.

This paper describes the methodology's fidelity assessment portion of initial simulation facility evaluation.

THE NEED FOR FIDELITY ASSESSMENT

It has become increasingly important that each simulation facility realistically mimic the actual plant. This increased importance is attributable to at least two factors. First, simulators have become a desired medium for providing nuclear power plant (NPP) operators with the skills required for plant operation. Second, it has become increasingly apparent that some simulators may not function in the same manner as the actual plant for some plant conditions. When opportunities have arisen to compare actual plant performance during a transient to a simulator's performance, the simulation facility has sometimes behaved very differently.

Fidelity is not the "bottom line" of a simulator's performance. The true determination of a simulation facility's effectiveness is how the operators trained and tested in it can perform in the plant. However, since direct measurement of operator performance is difficult, impractical, or even impossible for many of the NPP tasks which are tested, the measurement of simulator fidelity is often the best measure that can reasonably be taken.

SOURCES OF FIDELITY DATA

To assess fidelity, two types of data have to be collected for each task for which the simulation facility will be used for operator testing, namely, simulator performance and baseline plant data. Generation of simulator performance data will involve setting up a set of simulation facility starting conditions, developing a scenario of events which will occur, and then collecting data on the changes in values of selected operator display parameters over some period of time. Those parameters which are monitored will depend upon the particular task.

There are three primary sources of reference data on which simulation facility evaluations are based. First, there is actual plant data from the reference plant for which the simulator is being designed. This is clearly the best measure since it represents the ultimate goal of simulation facility performance. However, actual plant data cannot be obtained to represent all operator tasks which will be tested on the simulator. The situation may never have occurred in the plant, particularly for the relatively severe transients for which the simulator will be used. Power plants which are under construction will obviously not have any operational data. Even if the plant exists and the situations have occurred, the data collected on plant performance during the occurrence may not be of sufficient precision, accuracy, or completeness to use for simulation facility evaluation. If good plant data had been collected, it may have been used for developing the simulator mathematical models, in which case the data could not be used for simulation facility evaluation.

The second source of baseline data is from similar plants. The definition of what constitutes "similar" is not a simple issue. Some of the characteristics of the plant which must be considered are:

1. The nuclear steam supply system including reactor type, number of coolant loops, and the power rating.

2. The emergency core cooling system including system types, number of pumps, and automatic initiation conditions.

3. The arrangement of reactor auxiliaries.

4. The secondary plant.

If a plant exists that is sufficiently similar to the reference plant, then one should consider collecting baseline data from the similar plant for those tasks for which data are available. Of course if reference plant data exist for an operator task, then similar plant data need not be collected.

The same problems may arise for similar plant data as for reference plant data; the situation may never have arisen, the data may be inadequate, or the data may have been used during simulator design. In addition, there are obvious logistical problems in locating and securing the data from other NPPs.

The third potential source of baseline data that is considered is plant performance data generated by the use of best estimate engineering models. These models are generally more sophisticated than the mathematical models that are used in the simulator, primarily because the constraint for real-time model execution does not exist. Because these models can include more variables and interactions among variables and can operate in smaller time increments, they are generally better predictors of plant performance than the simulator models. Therefore, if no actual plant data exist, the engineering models can provide a baseline for comparison. There are, however, drawbacks of these models including the time required to set up the model for any scenario, the computer costs of running the models, and some doubt as to these models' ability to predict plant operation during a scenario.

The selection of a baseline data source should be made individually for each operator task. The baseline selected should be the best possible. As previously stated, reference plant data is far and away the preferred alternative, with similar plant and engineering model data as acceptable alternatives for the situations in which plant data do not exist and cannot be obtained. When deciding whether to use similar plant vs. engineering-model data, one should consider first, the degree of similarity to the plant as outlined above and, second, the expected accuracy of the engineering models which would be used.

THE FIDELITY DATA WHICH NEED TO BE COLLECTED

The assessment of the fidelity of the simulation is considered to be the same as assessment of the accuracy of the simulator's mathematical models. To do this assessment of mathematical models, two approaches might be taken. One would be to examine the simulator models directly. This would involve examining the lines of computer code and determining whether (1) the correct variables were included in each of the submodels, (2) the variable update frequency was sufficient, (3) the numerical approximation techniques were adequate, and (4) the functional relationships were correct. Unfortunately, the state-of-the-art in NPP modeling is not sufficiently well advanced to permit a determination of what constitutes "adequate," "sufficient," or "correct."

The second approach, and the one used in our methodology, is simply to observe the outputs of the simulator models at a level where they can be directly compared to the baseline data, i.e., display parameters. It can be argued that this is a better approach since we are directly measuring the simulator's ability to correctly simulate the baseline plant performance data. The disadvantage of this approach is that it is far less clear when one has collected a sufficient amount of data to support the contention that the simulation facility does predict plant performance under conditions which differ from those under which the baseline data were collected.

Two factors are considered in determining the display parameters for each operator task: (1) those displays which operators rely on most in performing the task and (2) those display parameters for which data have been or can be collected.

Determining the Display Parameters On Which Operators Rely in Performing the Task

What must be developed to answer this question is a "critical-display-parameters-by-operator-tasks" matrix. The operator tasks are the evolutions and malfunctions listed in American National Standard (ANS) 3.5 (American Nuclear Society, 1981). The critical displays for each task can include any of the displays which are available to the operator. The design of this matrix proceeds according to the following steps.

Develop a list of operator displays. To determine which displays are used by operators, one must first develop a list of all displays that are available. This includes all analog, digital, and binary displays. To facilitate development and use of this display listing, the displays should be grouped hierarchically first by the plant system to which they pertain and second, within each system, by the types of information which are presented on the physical status of the plant.

Obtain opinions from at least two plant operators on the ten most critical displays for each task. The first set of data is collected from experienced plant operators regarding which of the displays are used for each of the tasks. Each of the operators is given a series of forms, one for each operator task. On these forms is the following:

1. The title of the operator task.
2. A brief description of the task.
3. A set of instructions to the plant operator who will fill out the form.
4. A list of all operator displays.

The brief description of the task serves to clarify any ambiguity about the task which is not clarified in the task title. The instructions indicate to the plant operator filling out the form that he is to allocate a total of 100 points to no more than ten of the listed displays. The number of points allocated to any display should reflect the importance of that display to the operator when he performs the task. If less than ten displays are truly important, then points should only be allocated to those that are important. Then, the operator is to mark the appropriate number of points in the space provided by each of the displays listed.

<u>Obtain opinions from at least two nuclear engineers/designers on the ten most critical displays for each task.</u> Data on the most critical displays is also collected from at least two nuclear engineers or designers; the method of collection is identical to that used for plant operators. This step is intended to provide a check on the data obtained from the operators. The engineers/designers should also know what parameters the operators should be considering in performing the task. Preferably, the engineers/designers should be familiar with the reference plant.

<u>Reconcile differences between plant operators and nuclear engineers, if necessary.</u> Once these data are collected, the opinions from the individuals (operators and engineers/designers) are compiled onto one sheet for each operator task. Ideally, the two lists would be identical. However, since this is a subjective rating of relative importance, differences should be expected. If the operators and the engineers agree on less than 50% of the critical displays, then that may be

an indication that the operators and engineers have perceived the definition of the task differently. A reconciliation of views can be attempted which simply involves a meeting between the two groups to mutually determine which displays are truly important. At worst, more than ten parameters can be measured for a task during fidelity assessment, thereby reflecting the combination of the two groups' opinions on the most critical displays. If the two groups agree on more than 50% of the displays, then the total points for each display are summed. The ten displays receiving the highest total score are deemed as those for which measurements should be taken during fidelity assessment of the simulator for that task.

Determining the Display Parameters on Which Data Have Been or Can Be Collected

At this point, one has determined what variables should be measured during the fidelity portion of initial simulator evaluation. Now practical considerations must come into play. The question is whether what one has is what one wants and, if not, what can be done to rectify the shortcomings?

Operator tasks for which reference or similar plant data are to be used as a baseline. The rule for determining whether the plant data are satisfactory is if data were collected on 40% of the critical display parameters, then these data are sufficient. If one samples 40% of the critical displays and they are found to be satisfactory, then it can be assumed with some level of confidence that the other displays will function properly. However, if data for less than 40% of the critical displays can be assessed with plant data, then engineering models should be used instead to collect the baseline data.

Operator tasks for which engineering-model data are to be used as a baseline. The rule for determining whether the engineering-model data are satisfactory is if data were collected on 60% of the critical display parameters, then these data are sufficient. The reasons for setting a higher requirement for engineering-model data than for plant data are two-fold. First, the engineering-model data are probably less valid than plant data and, therefore, one should be more cautious in accepting the simulator when using engineering-model data as a baseline. Second, engineering models can usually be modified to provide output about the parameters of interest.

Absolute/Trend Parameters

Before one can develop the critical-display-parameters-by-operator-tasks matrix, he needs to know, for each variable, whether it is an absolute or trend parameter. An absolute parameter is one in which the absolute value of the parameter is important to the operator during performance of the task. Trend parameters, on the other hand, are important only with respect to the rate at which they are changing, and not necessarily to the absolute value of the parameter. For every operator task, each parameter is designated as either a trend or an absolute parameter. This designation is made by a subject-matter expert. It is conceivable that display parameters may be trend parameters in some tasks and absolute parameters in others.

Once this analysis is completed, the matrix is constructed. The dimensions of the matrix are (1) the operator tasks which are to be tested in the simulation facility (rows of the matrix) and (2) the plant control-room displays (columns of the matrix). The cells of the matrix are either (1) blank indicating that the display is not critical for performance of the task, (2) a "T" indicating that the display presents a critical trend parameter for performing the operator task, or (3) an "A" indicating that the display presents a critical absolute parameter for performing the operator task.

COLLECTING AND RECORDING THE FIDELITY DATA

In order to minimize the overall data collection and analysis effort, careful attention should be given to the form and format of recording the data as well as the source and content of the data themselves. In order to conceptualize the problem, consider that during data analysis every baseline data point must have a corresponding simulator data point and the focus will be to determine whether the two numbers are nearly the same.

To facilitate this determination, the prime goal of data collection is to ensure that these pairs of points are truly comparable. This requires that the baseline and simulator data are synchronized and that any simulated operator actions or equipment malfunctions occur at the same relative time. A shift of even a few seconds can lead to the appearance of great differences between the simulator and the baseline when, in fact, the differences are simply due to a phase shift in the data collection timing.

The methods of collecting and recording the data have a significant impact upon the effort required in analyzing the data. With state-of-the-art NPP parameter recording systems and simulator performance monitoring systems, the data analysis requires little more than developing several computer programs to reformat the data. However, if all data must be collected manually, then hundreds of man-hours may be required to reduce the data. Even if the data are collected automatically, careful attention must be paid to ensure that the synchronization issues are adequately addressed.

ANALYZING THE FIDELITY DATA

Four summary descriptive statistics are computed for both absolute and trend parameters: (1) root mean squared (RMS) error, (2) percent error, (3) maximum error, and (4) error t-score. The first three statistics provide descriptive information about the simulator's fidelity. Each of these three statistics represents a different aspect of fidelity, each of which is important to human perception in a different way. The computation of an error t-score provides an inferential statistical test of the simulator's resemblance to the plant with respect to the observed parameters. The four statistics are computed for each of the critical displays on the tasks which are being evaluated.

DRAWING CONCLUSIONS

The fidelity assessment procedure does not result in a statement as to whether the simulator has adequate fidelity as a whole. Rather, the simulation facility is deemed acceptable or unacceptable for the testing of individual tasks on the basis of the simulator's performance during a scenario embodying that task. To assess the simulation facility one must first determine if the simulator sufficiently replicates each of the critical operator display parameters within each scenario. Then, based on the number of critical display parameters successfully simulated, the simulation facility's overall acceptability in simulating the scenario is decided. If the scenario can be faithfully replicated by the simulator, we assume that the simulation facility can produce other scenarios of the same task with roughly equal success, and, hence, one should consider the simulator acceptable for testing of that task.

Two levels of acceptability were defined for each individual critical display parameter: fully acceptable or marginally acceptable. The criteria for each of these levels of acceptability with respect to each of the four statistical measures computed were also specified. All criteria are such that if the observed measure is less than the criteria, it is acceptable at the appropriate level. The selection of these criteria was based upon the recommendations of ANS 3.5.

To consider the simulator fidelity for testing a particular task acceptable, at least 75% of the critical display parameters must be deemed fully acceptable with respect to all four criteria, and at least 90% of the critical display parameters measured during the performance of the scenario must be deemed either fully or marginally acceptable. If the displays are critical to task performance, then it is essential that they behave properly in the simulator.

ASSESSING THE ADEQUACY OF THE SIMULATOR CONTROL-ROOM LAYOUT

Up to this point, we have described procedures designed to investigate whether the simulation facility's representation of the reference plant's dynamics are adequate for operator testing. If these criteria are satisfied, then one must ensure that the control-room layout is a sufficient replication of the reference plant's control room.

The procedures for conducting this part of fidelity assessment are relatively straightforward. A collection sheet is prepared for each operator task for which the simulator is to be used for operator testing. On this sheet is a list of the critical displays for the task. Data is collected by comparing the simulation facility's control-room layout with respect to each critical display. Each of the displays is compared with respect to the similarity of (1) display type, (2) minimum display value, (3) maximum display value, and (4) relative display location. To assess the similarity of relative display location, the following factors are considered:

1. The display should be located on a panel which is similarly located in the control room with respect to other control panels.

2. The display should be located in a similar location on the panel with respect to other displays on the panel.

It is not essential that the relative location of each of the displays be identical to the reference plant. Rather, if an operator could identify a display strictly by its location (i.e., without referring to any display markings or other identifiers), then it is considered as having the same relative location.

For a control-room layout to be considered adequate, all of the critical displays must be rated as the same in the simulator as in the plant on all of the above four dimensions.

REFERENCES

American Nuclear Society (1981). Nuclear power plant simulators for use in operator training. American National Standard 3.5.

U. S. Nuclear Regulatory Commission (1984). 10 CFR parts 50 and 55, operator's licenses and conforming amendment. Federal Register, 49, 46428-46440.

U. S. Nuclear Regulatory Commission (1984). Nuclear power plant simulation facilities for use in operator license examinations.

Proposed Revision to Regulatory Guide 1.149.

NOTE

The methodology described in this paper has not been adopted (or rejected) by NRC. The paper in no way represents current or planned policy of NRC.