CONF-8506100--5

CONF-8506100--5

TI85 017152

NUCLEAR POWER PLANT SIMULATION FACILITY EVALUATION METHODOLOGY*

P. M. Haas
R. J. Carter
K. R. Laughery, Jr.**

Oak Ridge National Laboratory
Oak Ridge, Tennessee

Presentation Made At:

1985 IEEE Third Conference on Human Factors and Power Plants

Monterey, California
June 23-27, 1985

* Research sponsored by the U. S. Nuclear Regulatory Commission under DOE Interagency Agreement 40-550-75 with Martin Marietta Energy Systems, Inc. under Contract No. DE-AC05-84OR21400 with the U. S. Department of Energy.

**Micro Analysis and Design
Boulder, Colorado

NUCLEAR POWER PLANT SIMULATION FACILITY EVALUATION METHODOLOGY

P. M. Haas and R. J. Carter
Reliability and Human Factors Group
Engineering Physics and Mathematics Division
Oak Ridge National Laboratory
Oak Ridge, Tennessee

K. R. Laughery, Jr.
Micro Analysis and Design
Boulder, Colorado

Abstract

A methodology for evaluation of nuclear power plant simulation facilities with regard to their acceptability for use in the U.S. Nuclear Regulatory Commission (NRC) operator licensing exam is described. The evaluation is based primarily on simulator fidelity, but incorporates some aspects of direct operator/trainee performance measurement. The panel presentation and paper discuss data requirements, data collection, data analysis and criteria for conclusions regarding the fidelity evaluation, and summarize the proposed use of direct performance measurement. While field testing and refinement of the methodology are recommended, this initial effort provides a firm basis for NRC to fully develop the necessary methdology.

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 to Oak Ridge National Laboratory (ORNL) for the development of a methodology for performing the simulation facility evaluations. Drafts of a proposed handbook for evaluation [1] and of a technical report describing the technical bases for the handbook [2] have been produced which will be used as part of NRC's basis to fully develop, test, and implement a simulation facility evaluation process. This presentation summarizes the essential contents of the proposed methodology/handbook.

Summary of Methodology

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 (i.e., evaluation of operator/trainee performance).

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.

An essential premise of the methodology is that the simulation facility is not evaluated on a "go or no-go" basis. That is, the decision as to a facility's acceptability for licensing is not binary. Rather, the evaluation is performed on a "task-by-task" basis. The simulator may have acceptable fidelity for many tasks and marginal or unacceptable fidelity for others.

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 the facility can perform in the plant. However, since direct measurement of operator performance is difficult, impractical, or even impossible for many of the nuclear power plant (NPP) tasks which are tested, the measurement of simulator fidelity is often the best measure that can reasonably be taken. The methdology proposed relies heavily on fidelity assessment, but includes some simple elements of direct operator/trainee performance measurement based on basic concepts and experimental paradigms for "transfer of training" studies.

Fidelity Assessment

Sources of Fidelity Data. For fidelity assessment, 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" or reference data. The simulator performance data will involve setting up a set of simulation facility initial conditions, developing a scenario of events which will occur, and then collecting data on the changes in values of selected operator display parameters for 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, it is recognized that many practical factors make it impractical or impossible to have sufficient plant data, particularly when the simulator is often available before the plant is constructed.

The second acceptable source of reference data is from similar plants. The definition of what constitutes "similar" is not a simple issue. Some of the constituents to 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.

Again, there are practical constraints which mitigate against having sufficient data from "similar" plants.

The third potential source of data that is considered is plant performance data generated by the use of best-estimate engineering models. These models are generally more sophisticated than simulator models.

The selection of a reference data source should be made individually for each operator task. As was previously stated, reference plant data is by far the preferred alternative with similar plant and engineering-model data being acceptable alternatives for the situations in which plant data do not exist and cannot be obtained.

Data To Be Collected. The approach taken for fidelity assessment is simply to observe the outputs of the simulator models at a level where they can be directly compared to the reference data, i.e, display parameters. 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. The latter decision takes into account the three acceptable sources, their order of preference, and practical constraints of time, cost, equipment, etc. The determination of critical displays is accomplished in four steps:

1) Develop a list of operator displays.

2) Obtain opinions from at least two experienced plant operators on the ten most critical displays for each task to be tested.

3) Obtain opinions from at least two "nuclear engineers/designers" on the ten most critical displays for each task.

4) Reconcile any differences between (2) and (3), and finalize list.

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 for recording the data as well as the sources and content of the data themselves. To conceptualize the problem, consider that during data analysis every reference data point must have a corresponding simulator data point, and the focus will be to determine whether the two values are nearly the same.

Therefore, it is of primary concern in preparing for data collection to ensure that these pairs of points are truly comparable. This requires that the reference 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 reference 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: (a) root mean squared (RMS) error, (b) percent error, (c) maximum error, and (d) 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 on Fidelity. As noted previously, 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 as is evidenced by 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 simulate other scenarios of the same task with roughly equal success, and hence, one should consider the simulator acceptable for testing of operators on 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. [3]

In order for the simulator fidelity for testing a particular task to be considered 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 must be deemed either fully or marginally acceptable as they were measured during the performance of the scenario. If the displays are critical to task performance, then it is essential that they behave properly in the simulator.

Evaluation of Operator/Trainee Performance

To support the fidelity assessment, it is desirable to have some direct measure of simulator acceptability that addresses the end product, i.e., operator performance on the job. Numerous research approaches involving empirical evidence for "transfer of training" in simulators are available in the literature. Ten different approaches summarized by Caro [4] are:

- Transfer of training (forward transfer)
- Self-control transfer
- Pre-existing control transfer
- Uncontrolled transfer
- Simulator-to-simulator transfer
- Backward transfer
- Simulator performance improvement
- Simulator fidelity analysis
- Simulator training program analysis
- Opinion survey

Assessment of these approaches and other literature in the field with regard to practical constraints in the operational setting of nuclear power plants led to the conclusion that implementation of most of research paradigms involving direct performance measurement was not practical. However, the essential concepts of forward and backward transfer-of-training could be retained, and useful information based on trainee/operator performance could be obtained. It was proposed that this be accomplished in two ways:

(1) Backward transfer - Reactor operators experienced in plant operation will perform operations on the simulator which they have experienced in the plant (normal evolutions and/or transients), and their performance will be observed/recorded. Observed difficulties or differences in their performance (based on subjective data on their in-plant experience) will be identified as candidates for investigation of discrepancies between simulator and plant performance.

(2) Forward transfer - Trainees who have recently attained certification and have little or no direct plant operating experience will have their initial in-plant performance of specific tasks observed/recorded. Difficulties encountered will be identified as candidates for investigation of potential simulator/plant discrepancies.

Obviously, in either of these cases, there are many potential reasons for differences in human performance. Observed differences are suggestive only, not definitive.

Conclusion

A practical approach to evaluation of nuclear power plant simulation facilities has been developed. It has not been tested in an operational environment. It is fully expected that such testing would lead to considerable refinement and modification. However, the methodology was designed to incorporate the essential concepts derived from a substantial base of experience and research reported on evaluation of nuclear and non-nuclear simulators with utmost emphasis on practical constraints within the nuclear industry and the NRC. Given the appropriate testing and refinement, which apparently is planned by NRC, it should provide the nuclear industry and NRC with an effective process for assessing acceptability of a simulation facility for use in the NRC licensing process.

References

[1] K. R. Laughery, Jr., R. J. Carter, and P. M. Haas, "Nuclear Power Plant Simulation Facility Evaluation Methodology: Handbook," NUREG/CR-4188, Vol. 1, ORNL/TM-9570/V1 (Publication expected 08/85).

[2] K. R. Laughery, Jr. and R. J. Carter, "Nuclear Power Plant Simulation Facility Evaluation Methodology: Technical Bases," NUREG/CR-4188, Vol. 2, ORNL/TM-9570/V2 (Publication expected 09/85).

[3] ANSI/ANS 3.5-1981, American National Standard Nuclear Power Plant Simulators for Use in Operator Training (Draft revisions, May 1985, designated ANSI/ANS 3.5-1985).

[4] P. W. Caro, "Some Factors Influencing Air Force Simulator Training Effectiveness," Technical Report HumRRO-TR-77-2, 1977.