ANALYSES METHODS AND PILOT APPLICATIONS OF SACADA DATABASE

Mohamad Ali Azarm [a], Inn Seock Kim [b], Clifford Marks [a], and Faramarz Azarm [a]
[a] Innovative Engineering and Safety Solutions, Germantown, MD, USA
[b] ISSA Technology Inc., Germantown, MD, USA

Abstract: The U.S. Nuclear Regulatory Commission (NRC) is collecting the licensed operator performance information from simulator exercise of nuclear power plants (NPPs) to develop empirical basis to better understand the elements affecting operator performance and to help with estimating the associated human error probabilities (HEPs). USNRC is developing Scenario Authoring, Characterization, and Debriefing Application (SACADA) system to collect the licensed operator data in a consistent manner. This study uses a selected SACADA Data Set provided by the NRC. This SACADA database contains a total of 25,784 data points (individual crew performances) for a total of 2014 actions; called Training Objective Elements (TOEs).

A program was sponsored by the NRC at the Innovative Engineering and Safety Solutions (IESS), in August of 2017, with the objective of developing methods for SACADA data analysis and to show their utilities using example applications. A final goal of this program was to develop an empirical basis to estimate the HEP values and identify the major situational factors (SFs) contributing to those HEP estimates.

The approach taken in this study uses the concept of context similarity. The underlying premise of context similarity is based on the assumption that the HEP values associated with two actions are close as long as all or the majority of their SFs are the same. The human action for which the HEP value is to be estimated (identified as the input action) is first characterized by the associated SFs similar to other SACADA entries. The entries from SACADA database are then examined to identify those which closely match the SFs associated with the input action. The operator performance in SACADA database then is used to estimate the HEP associated with the input action using evidential data obtained from SACADA database.

Data from context similarity were considered applicable for consideration only if the critical SFs are matched. Critical SFs are those SFs and their combinations that are expected to have major impacts on the estimated HEP values. Critical SFs; i.e., the most important SFs, are identified for each macro-cognitive functions (MCFs). Statistical significance tests are performed to examine whether the performance data of an MCF that includes a specific SF shows strong deviation from the overall performance data for that MCF. The impact of critical SFs on the HEP values was evaluated by an MCF tree and with the use of Bayes estimation method.

Feed and Bleed (FB) operation in response to loss of feedwater transient, in a four loop Westinghouse plant, was used as a comprehensive pilot application of the methodology. The HEP values for all subtasks were estimated using Bayes method and were aggregated to get the feed and bleed initiation error probability. The result shows the feasibility and the reasonableness of the methods developed in this study.

Keywords: Human Factors, Human errors, HRA, SACADA

1. INTRODUCTION

The Scenario Authoring, Characterization, and Debriefing Application (SACADA) methodology (1) significantly enhances and formalizes the previous Human Reliability Analysis (HRA) approaches. It defines specific Training Objective Elements (TOEs) and Macro-Cognitive Functions (MCFs). There are currently four macro-cognitive functions defined in SACADA data base. These are: monitoring/detection (information), diagnosis/response planning (understanding and choosing response), manipulation (action) and external communication. This study similar to earlier SACADA studies(2) differentiates between the diagnosis (choosing right procedure) and response planning (understanding the situation). As a result; we define five MCFs rather than the four identified in the SACADA database. Each key operator response is explicitly identified within a TOE, and it is associated with one of the five MCFs.

Every TOE in the SACADA database contains context information and performance results. The context represents the performance challenges, where the performance results include success, failure, or poor (i.e., a degraded performance that did not result in failure). The end results of the human performance are evaluated in terms of excellent (SAT+), satisfactory (SAT), satisfactory with deficiencies (SATΔ), and unsatisfactory (UNSAT). If a TOE's performance is rated as SATΔ or UNSAT, additional information regarding the details of the performance issues is collected. This includes error types, error specifics, error causes, team error recovery, and overall effects on the scenario.

When developing SACADA taxonomy, the developers envisioned using the **context-similarity** approach to use SACADA data to inform HRA. Defining human performance subtask in terms of TOE with its associated MCF and specific context (a set of SF states), may allow development of generic HEP values that can be used as table look-up. One could envision that a PRA analyst may search a large SACADA database to identify all similar cases to his/her input action for which a HEP value is to be estimated. To be able to do this we need to define what is meant by similar. If similarity is defined as all SFs for an input action to be the same as SACADA entries, the database will not have sufficient scope to yield performance data for most cases. We anticipate this will be the case in the foreseeable future. One option is to use performance data from SACADA entries that partially match the SFs associated with the input action (rather matching all SFs). This can be achieved by first allowing differences in SFs that are not considered critical to be neglected, and then identifying the minimum number of matches that can yield statistically homogenous performance data (i.e., performance data that can be pooled).

The above approach necessitates that a set of critical SFs be first identified. This study focusses on all actions that are categorized under one MCF, regardless of their TOEs. As an example, a manipulation action with its specific SFs is expected to have the same HEP values for different TOEs (not differentiated for various TOEs).

Within the same MCF, statistical significance tests are used to identify the individual SFs that can have major impact on the resulting HEP values (i.e., detecting strong deviation of performance data of the SFs and the overall MCF). The critical SFs per each MCF are currently limited to five to ensure that the approach can be supported by the existing scope of SACADA database. Estimates of HEP values based on critical SFs and their combinations, are made using Bayesian approach and with the aid of MCF tree which will be discussed later. This study similar to previous studies use Bayes method for HEP estimation(3).The HEP values estimated for the combination of critical SFs are used as a prior distribution for the input action HEP value, with evidence provided by the performance data resulting from the context similarity evaluation.

1.2 Objectives

The objective of this study is to develop methods and show their utilities in empirically estimating human error probabilities (HEPs) using SACADA database. A secondary objective is to identify the specific SFs in SACADA database which contribute the most to HEP values and quantitatively estimating their contributions. Finally, the study demonstrates the feasibility and the reasonableness of implementing all the methods in an integrated fashion through a comprehensive pilot application.

2. CONTEXT SIMILARITY METHOD

In context similarity approach, the input action is characterized by a specific context; a set of relevant Performance Influencing Factors (PIFs) and specific SFs. The context of the input action is used to identify the SACADA entries with similar context. The degree to which two actions are similar in context is measured by the number of SFs that are common (or matched) between them.

Separating out the scenario specific portion of the SACADA database would leave 31 PIFs with a total of 134 SFs for context similarity evaluation. The degree of context similarity is therefore defined by the number of PIFs with the matching SFs. For example context similarity of 27 indicates that 27 out of 31 PIFs have matching SFs.

The basic premise of using context similarity approach is based on a hypothesis that if two actions have exactly the same context, they should also have the same HEP values. The validity of this hypothesis was the main reason behind developing the SACADA approach. However, the current SACADA database is rarely expected to yield a data point with a complete match to the SFs of the input action. To alleviate this issue; the original hypothesis is relaxed such that: "if two actions have approximately the same context, they should also have closely estimated HEP values." This new definition of the hypothesis requires that two new terms be technically formulated. These are:

1. What is meant by "approximately the same context"?
2. What is meant by "closely estimated HEP values"?

This study defines the term "approximately the same context" as two human actions that possess the following characteristics:

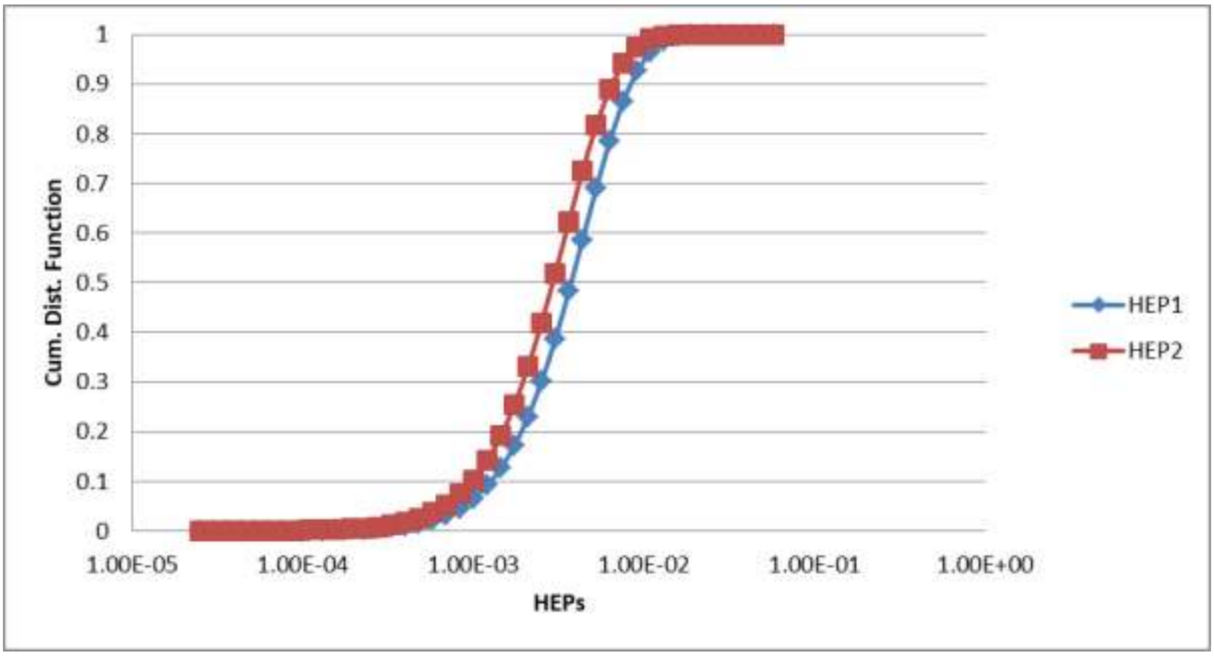
- The most dominating SFs have to be matched. This basically means that the two actions should possess matching critical SFs. Methods for identifying critical SFs are discussed in section 2.3.
- All partially matched data that meet the first criterion will be included if they do not significantly deviate from the performance data obtained for the higher number of matches. This is done via statistical significance test. The statistical significance test examines the hypothesis if the performance data generated from a lower-level matched data from context similarity supports the performance data generated from the higher number of matches. The same statistical tests are also used in support of developing critical SFs. They are discussed in Section 2.4 of this paper.

This study initially defines the term "closely estimated HEP values" by comparing the two distributions of the HEP values that are considered to have "approximately the same context". The actual numerical criteria used for defining "closely estimated HEP values" is exploratory in nature and could be subject to change after further use. The two estimated HEP distributions are considered close, if the two distributions have a large overlap and their distribution means are close. The large overlap is satisfied if the 90 percent of the 90 percentile interval (i.e., five and ninety five percentile interval) of the updated distribution (i.e., selected distribution) is covered by the 90 percentile interval of the original distribution. This implies that there is 81% probability (0.9*0.9) that a sample taken from the updated distribution which resides within its ninety percentile will be covered by the ninety percentile interval of the original distribution. This criteria, although similar to 90/90 criteria which is generally used in deterministic analysis[4], should not be confused with that concept. The mean of the updated distribution is considered close to the mean of the original distribution if it is within the twenty five – to seventy five percentile interval of the original distribution.

For these analyses, each of the HEP distributions is estimated using the Bayes method, with an appropriate prior and evidence from SACADA data search using SACADA Data Mining Software (SDMS) code for number of matches greater than a specified value.

As an example, consider that the performance data for the highest number of matches (say 28 matches out of 31) shows occurrence of one UNSAT condition in 103 trials, typically shown as (1, 103). At one level lower, i.e. 27 matches, the performance data is (0, 141). Statistical significance test, as will be discussed later, would show that the two data can be supported by one underlying failure rate (HEP value). The Bayes method now can be used to estimate the two HEP distributions; one based on performance data for 28 matches (1, 103), and the other based on cumulative performance data for 27 and 28 matches (1, 244). The Cumulative Distribution Functions (CDFs) for the posteriors are shown in Figure 1. The ninety percentile interval of HEP2 is first estimated (from 6.09E-4 to 7.6E-3). The corresponding probability interval of HEP1 lies between 6.09E-4 to 7.6E-3 is about 0.83 (0.86 (for 7.60E-3) - 0.03 (for 6.09E-4)). Considering the proposed criteria for ninety percent probability for the ninety percentile coverage (i.e., 0.9*0.9=0.81), the combined distribution would be considered to be close estimates and combining data is valid. Similarly, the two means are very close; estimated at 4.4E-3 and 3.4E-3 for HEP1 and HEP2 respectively. The mean of HEP2 is well within the twenty five to seventy five percentile interval of the original distribution, i.e. HEP1 (2.4E-3 to 6.3E-3).

Figure 1: Posterior CDFs for the highest number of matches (HEP2) and combined matches (HEP1)



2.2 SCADA Data Mining Software (SDMS)

Working with the SACADA database without automation is quite time consuming. It could also be susceptible to various types of errors. Large number of data queries was initially made manually for developing the critical SFs. The experience showed that for the continuation of work and specially for the purpose of developing the context similarity vectors, it would be beneficial to develop in-house software.

SDMS has been developed in FORTRAN 90 for window environment. It was written in a manner that is scalable. It currently can handle a database which is more than 10 times larger than the current SACADA database. The software is developed as a standalone program and is optimized for both time and memory management.

SDMS input file contains the name and specific SFs associated with input actions. Up to twenty human actions (i.e., human failure event; HFE) currently can be described in a single input file. The information associated with each HFE, i.e., the specific assignment of 31 SFs to that HFE, is included in the input file. SDMS performs all database queries and identifies all SACADA data which meet the critical SFs and identifies the number of SF matches for each SACADA entry.

SACADA data entries in database may have recorded “Null” for some SFs. In some cases, a “Null” value was assigned to an SF in SACADA database since the exact assignment was not known. In other cases, the analyst used “Null” when facing with multiple choices. Null assignment is problematic, especially for matching critical SFs. A null assignment to a critical SF could cause valid entries to be eliminated from the analysis. For this reason SDMS offers the option to treat Null for critical SFs as missing data, hence consider them to be a match.

For some analyses, the user may specify certain PIFs to be critical for a specific HFE. All SFs under those critical PIFs then are considered critical SFs. As an example, the analyst may require that the overarching PIFs (for example for time criticality) be critical. The use of critical PIFs rather than critical SFs is very restrictive and could result in a large number of SACADA data entries to be screened out from further analysis. There is an option in the SDMS, if invoked by the user, it relaxes this restriction. This is done by accepting the SACADA entries even if not all SFs of critical PIFs are matched (for example allowing one or two mismatch), as long as all critical SFs in the input HFE are matched.

SDMS software is a work in progress. The development so far is for demonstrating the feasibility of implementing context similarity as a means to estimate empirical HEP values. The ultimate goal however is to develop integrated software for SACADA data mining, statistical tests, and Bayes estimation routines.

2.3 Methods for Identifying Critical SF States

A necessary input for successful implementation of context similarity approach is the identification of the critical SF states. The SF associated with an input action could include various combinations of critical SFs. Therefore, it is important to estimate the probability distributions of HEP values that contain any number of critical SFs. These distributions may be used as the prior distributions in Bayes routine for larger performance data that could be generated from context similarity approach. The HEP estimates for various combinations of critical SFs are accomplished by use of MCF tree, statistical tests, and Bayes estimation technique. A process flow chart depicting the proposed method is provided in Figure 2.

The process starts with selecting one MCF, such as manipulation or monitoring and detection. There are five MCFs for SACADA database provided that diagnosis and response planning are differentiated. For the selected MCF, one starts with developing the list of the relevant SFs.

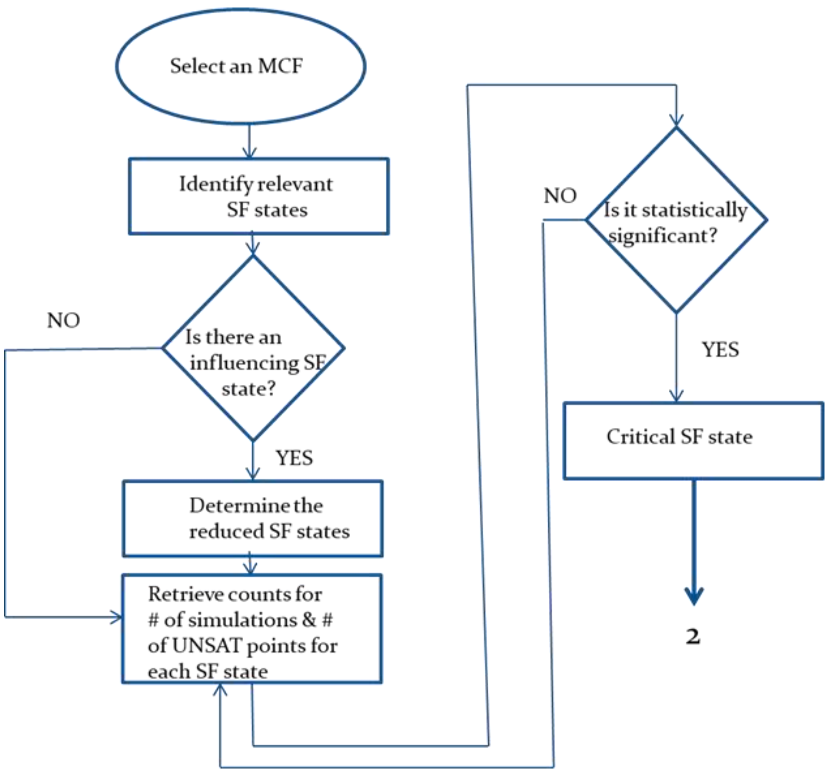
For some MCFs, selection of specific SF states can significantly reduce the number of relevant SFs. This is referred to as the influencing SFs. For example, if alarm is selected for monitoring and detection MCF (for PIF associated with the detection type), all SFs in other PIFs that relate to meters/lights/flags could be eliminated.

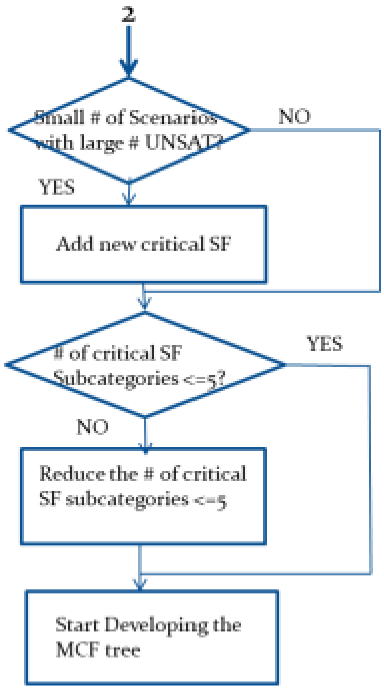
In the next step of the process, the data associated with the number of simulations and the numbers of UNSAT cases are retrieved for each of the SF states. Statistical significance tests are performed to examine if the data for the SF state shows strong deviation from the overall data for the MCF. If there is no data or there is insufficient data for a SF state, they are identified as data gap. All SF states that are considered statistically significant are considered critical SFs. The statistical significance tests for the exploratory analysis are discussed later in Section2.4.

For some MCFs there could be one or two scenarios with a small number of training simulations but relatively a large number of UNSAT cases which may not have been picked up by the statistical test. The SACADA database is also examined for cases with a large number of UNSAT points in one row. This could result in the selection of additional critical SF state (or a combination of SF states that could be important).

The number of SF states that are considered for the exploratory data analysis in this study is limited to five (5). This is done to assure that the evaluation can be done within the Microsoft Excel program, without the need to develop an external program for automation. To ensure that important SF states have not been missed since the study is limited to five SF states, expert judgment is also utilized. Expert judgment was used to examine if there is any additional SF states that may be important but not captured by data evaluation. A small team composed of two project staff was assembled for expert review of the data. The team included a senior reactor operator with 15 years of experience in plant operation, and a senior HRA/PRA analyst with over 25 years of experience. The experts worked in close coordination with the data analyst to identify the important SF states that might have to be added.

Figure 2 Process flow chart for the exploratory data analysis





2.4 Statistical Significance Test

The statistical significance test is used to test the hypothesis that two sets of observations are drawn through a binomial process with the same binomial parameter (p) (5). If the test accepts the hypothesis with 90% confidence then the two sets of observations are similar; they can be represented with the same probability. The test may reject the hypothesis; i.e. based on the lower and upper confidence tails at 5% and 95%. When the test is rejected, there would be no justification to assume that the two observations share the same probabilities, and furthermore their data cannot be pooled together.

The statistical significance test is used for two different purposes:

(1) Identify the critical SF states by determining when the performance data associated with a specific SF significantly deviates from the overall performance data for that MCF.
(2) Decide if the performance data of a parent can be used in support of developing the prior distribution for Bayesian estimation of the child probability. When the parent data is not significantly different than the child data, the parent's estimated probability distribution is used as a prior for the child within the structure of MCF tree which will be discussed next.

2.5 MCF Tree

The objective of developing the MCF tree is to develop a systematic method for estimating the probability distribution of the combinations of the critical SFs using a Bayes method. The MCF tree along with the statistical significance tests used for this purpose. Let us consider two critical SFs; A and B (as parents), and a combined SF, AB (as the child). There could be four possible outcomes when the statistical significance test (ST) is applied.

Table 1. Statistical significance for two SF states

Cases	ST(AB, A), ST(AB, B)	Comments
1	NS*,NS	AB cannot be differentiated from neither A nor B
2	NS, S	AB cannot be differentiated from A, but can be from B

3	S, NS	AB cannot be differentiated from B, but can be from A
4	S,S	AB can be differentiated from both A and B

*NS and S stand for Not Significant and Significant, respectively.

Case 1 indicates that the existing evidence cannot differentiate between AB and A, or AB and B. Therefore the prior distribution for AB can be assigned by prior distribution of either A or B. It is generally expected that for Case 1, the prior distributions for A and B would be the same. In rare occasion that the two prior distributions are not the same, the prior distribution with a higher mean is assigned. Cases 2 and 3 indicate that the prior distribution could be assigned the same as the prior distribution of A or B, respectively. Case 4 basically shows that the data from AB is different from both A and B. When the statistical test is rejected at the upper tail of the confidence level, the mean of the prior distribution of the child is multiplied by a factor of 10, and assuming a Log-Normal (LN) distribution, an error factor of 10 is applied. If the statistical test is rejected at the lower tail of the confidence level, the mean of the prior distribution of the child is divided by a factor of 10, and an LN distribution with an error factor of 10 is applied.

For individual critical SFs, a subjective prior was used. The subjective prior was considered to be in the form of a lognormal distribution. The mean was subjectively determined based on previous estimates from the results of HRA studies for the human errors associated with the MCF. An error factor of 10 (ten) was subjectively assigned to cover the range of uncertainties. Some sensitivity analysis was also performed using an error factor of 30 (thirty).

2.6 Bayes Estimates

Bayes estimation method was used to estimate the distributions of the HEP values for combination of critical SFs. All prior distributions were developed based on the MCF tree and the application of statistical significance tests. Some of the prior distributions for HEP values were expressed in the form of lognormal distributions, others in the form of beta distributions. For uniform application, all lognormal priors were converted to the equivalent beta distributions. The application of the Bayes method was based on the overall SACADA data, and therefore, it averages out the effect of uncertainties resulting from variabilities amongst crews (such as training level), plant to plant variability, and variabilities as a result of different scenarios (such as stress level, etc.). The current scope of the SACADA data does not offer sufficient data for proper treatment of all sources of uncertainties.

2.7 Example OF A MCF Tree

An example of MCF tree is presented here. This example is done for the Manipulation MCF. Initial exploratory analyses of data including the statistical significance tests identified five critical SFs as shown in Table 2.

Table 2. Significant SF states for Manipulation MCF

Col. & Row Number	Column and Row Headings	Data Points, # of UNSAT
AD2	Manipulation Location = Back Panel	(84, 4)
AF3	Manipulation Recoverability= Unrecoverable	(672, 16)
AJ3	Overarching Issues: workload= Multiple Concurrent Demand	(778, 20)
AM1	Overarching Issues: Others = Non-Standard	(274, 13)
AK3	Column AK: Overarching Issues: Time criticality = Barely Adequate Time	(451,8)

The MCF tree was developed for all combinations of these five critical SFs based on the discussion given in section 2.5. The final MCF tree, showing all the combinations of critical SFs that are represented with different HEP probability distributions, is presented in Figure 3. Table 3 shows the posterior distributions and the point estimates for each SF combinations shown in the MCF tree. Note that the uncertainties in the HEP estimates do not account (i.e. average out) for the uncertainties resulting from variabilities amongst crews (such as training level), plant to plant variability, and variabilities as a result of different scenarios (such as stress level, etc.).

Figure 3 Example MCF Tree for Manipulation

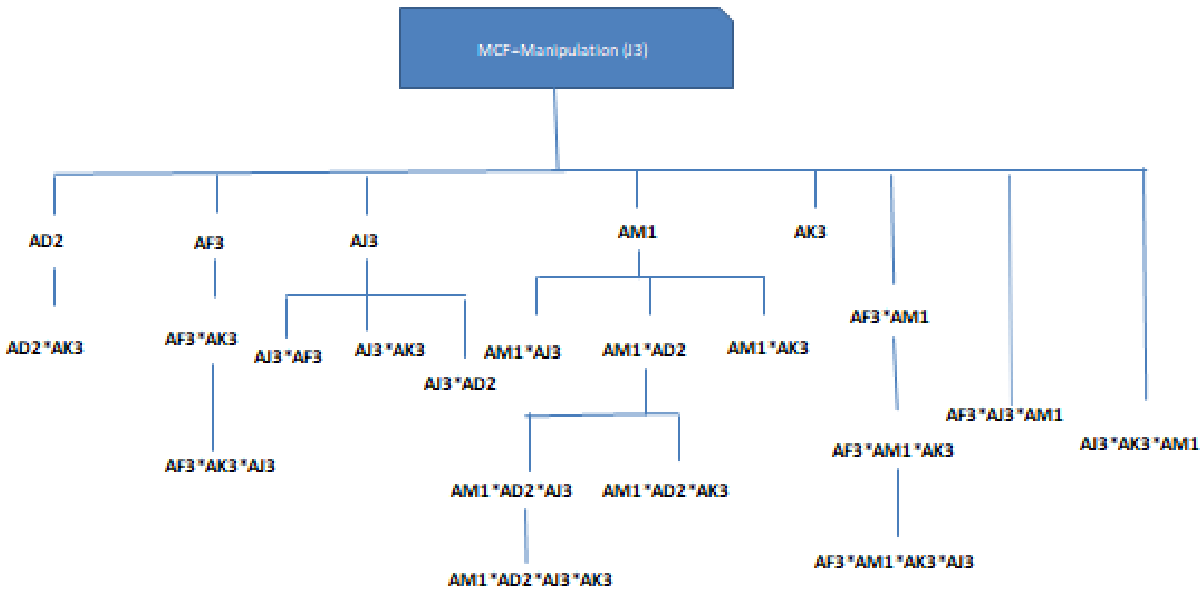


Table 3. HEP estimates for Manipulation MCF based on Combination of Critical SFs

SF States combination	Observation	Prior LN(mean, EF)-Beta (α,β)	Beta Posterior (α ,β)	Posterior 5% Lower Bound	Posterior 95% Upper Bound	Posterior Mean	Point Estimate
AD2	(84, 4)	LN(0.1, 10)	(5.1,91.7)	2.2E-2	1.0E-1	5.5E-2	4.8E-2
AD2*AK3	(15,0)+		(1.3, 26.8)	4.5E-3	1.3E-1	4.7E-2	<6.6E-2*
AD2*AJ3	(15,0)		(1.3, 26.8)	4.5E-3	1.3E-1	4.7E-2	<6.6E-2
AD2*AK3*AM1	(15,0)		(1.3, 26.8)	4.5E-3	1.3E-1	4.7E-2	<6.6E-2
AD2*AJ3*AM1	(15,0)		(1.3, 26.8)	4.5E-3	1.3E-1	4.7E-2	<6.6E-2
AD2*AK3* AM1* AJ3	(15,0)		(1.3, 26.8)	4.5E-3	1.3E-1	4.7E-2	<6.6E-2
AF3	(672,16)	LN(0.1, 10)	(17.3, 667.8)	1.6E-2	3.5E-2	2.5E-2	2.4E-2
AK3	(451,8)	LN(0.1, 10)	(9.3, 454.8)	1.0E-2	3.2E-2	2.0E-2	1.8E-2
AJ3	(778,20)	LN(0.1, 10)	(21.3, 769.8)	1.8E-2	3.7E-2	2.7E-2	2.6E-2
AJ3*AF3	(210,7)		(8.3, 214.8)	1.9E-2	6.0E-2	3.7E2	3.3E-2
AJ3*AK3	(320,7)		(8.3, 324.8)	1.3E-2	4.0E-2	2.5E-2	2.2E-2
AM1	(274,13)	LN(0.1, 10)	(14.3, 272.8)	3.1E-2	7.3E-2	5.0E-2	4.7E-2
AM1*AJ3	(76,5)		(6.3,82.8)	3.3E-2	1.2E-1	7.1E-2	6.6E-2
AM1*AD2	(15,0)		(1.3, 26.8)	4.5E-3	1.3E-1	4.7E-2	<6.6E-2

AM1*AK3	(84,6)		(7.3, 89.8)	3.8E-2	1.3E-1	7.5E-2	7.1E-2
AF3*AM1 AF3*AM1*AK3	(96, 13) (48, 6)	LN(0.1, 10)	(14.3,94.8) (7.3, 53.8)	8.3E-2 6.0E-2	1.9E-1 1.9E-1	1.3E-1 1.2E-1	1.4E-1 1.3E-1
AF3*AK3 AF3*AK3*AJ3	(166, 7) (107, 6)	LN(0.1, 10)	(8.3, 170.8) (7.3, 112.8)	2.4E-2 3.0E-2	7.5E-2 1.0E-1	4.6E-2 6.1E-2	4.2E-2 5.6E-2
AF3*AJ3*AM1 AF3*AJ3*AM1*AK3	(16,5) (16,5)	LN(0.1, 10)	(6.3, 22.8) (1.3, 26.8)	6.5E-2 6.5E-2	3.6E-1 3.6E-1	2.2E-1 2.2E-1	0.312 0.312
Other Combinations	(1970, 14)	LN(0.03, 10)	(15.2, 1995)	4.8E-3	1.1E-2	7.6E-3	7.1E-3

* When there is zero UNSAT observation, a maximum point estimate error is calculated by considering one failure; (i.e. instead of zero out of 15, a maximum based on 1 out of 15 is calculated).
+ The data for this entry was 15 trials with zero UNSAT observation. Both the grouping and point estimates are affected by insufficient data. This data is consistent with the overall data for manipulation so the prior used is the same as the prior for the manipulation MCF.

3. COMPREHENSIVE PILOT APPLICATION

Feed and Bleed (F&B) operation at a Westinghouse four loop plants was used for the pilot application. The following steps were taken for performing the pilot application.

- Define an accident scenario consistent with the PRA accident sequence such that specific accident sequence timing and additional demands on operators can be clearly defined.

- Describe the evolution of the scenario and clearly define the entry condition and cues/symptoms that help the operator to decide on the specific objective and the transition to appropriate procedures (e.g., emergency operating procedures (EOPs), abnormal operating procedures (AOPs), and annunciator/Alarm response procedures).

- Determine the expected path through the procedure, and identify the critical transition nodes. For some complex HFEs, there could be a possibility that multiple procedures are activated at the same time (e.g. fire scenarios or loss of support systems). Defining the HFEs and identifying the critical transition nodes requires an understanding of the crew's priorities in addressing the relevant responses as instilled by training and plant operations practices. Discussion of these scenarios with experienced reactor operators could help identify the most common priorities considered by operators.

- Develop a crew response diagram [CRD] (see example Figure 4-1 in NUREG-2199(6)). The objectives of the crew response diagram are two-folds; 1) graphically represent the transition nodes in the procedure, and 2) help identify the critical tasks. To develop the CRDs, the following elements should be considered. The crew response diagram for the purpose of using SACADA database has to meet certain requirements. These are:

    1. The level of resolution for critical tasks should be consistent with SACADA framework. For each critical task we have to define the associated PIFs within the SACADA framework. Whenever multiple PIFs apply to a critical task, the breakdown of that critical task is recommended. A critical task has to be associated with at least one of the five MCFs.
    2. Super nodes may be considered for non –critical tasks as long as their impact on accident progression timing and additional demand for operator actions associated with critical tasks can be captured.
    3. For complex situations when multiple procedures and priorities are involved, the CRD should be developed for the most commonly practiced

procedural steps (preferably by the way of feedback from those with reactor operating experience).

- Determine possible recovery actions for critical tasks, and evaluate the required time. SACADA framework has an explicit structure for the recovery actions associated with the manipulation MCF. The recovery is categorized as immediately recoverable, recoverable with significant effort and not recoverable. Recovery credit is currently implicit for other MCFs such as monitoring/detection or diagnosis/decision-making. It is tied to the time criticality. If expansive time is available for a critical task and there are independent checks and verifications, the task will have much higher probability for recovery and success as reflected by SACADA statistics. The main determinations to be made by the analyst are as follow:

  - For manipulation tasks determine if the action is recoverable and what category within manipulation recovery should be assigned to.
  - For all other MCFs, determine if there is sufficient time available for recovery and if there are independent checks and verifications before considering that the task is recoverable. For example if the task is not recoverable, the analyst may decide to assign nominal time for time criticality even if expansive time may be available.

Once the critical HFE subtasks are identified and characterized by assigning the appropriate SFs to SACADA PIFs, the HEP values can be estimated using SDMS software as discussed in Sections 2.

F&B operation for the specific scenario used for this pilot application consisted of eleven (11) critical HFEs and four (4) non-critical HFEs. Non-Critical tasks were considered for their possible impact on the time available for critical tasks.

The analysis result is shown in Table 4 for the point estimate results for two cases; 1) based on all SFs from context similarity approach, and 2) based on critical SFs only.

Table 4. HEP estimates for F&B operation using SDMS program

Actions	Applicable # of data points, # UNSAT	Number of Matches: # of data points, # of UNSATs							HEP point estimate	Mean HEP estimate
		31	30	29	28	27	26	25	All SFs	Critical SFs
0POP05-EO-E000-1	1018,3	0,0	0,0	14,0	147,1	282,1	323, 1	252,0	2.9E-3	2.5E-3
0POP05-EO-E000-2	1018,3	0,0	0,0	28,1	201,1	461,1	297,0	31,0	2.9E-3	2.5E-3
0POP05-EO-E000-3	23,0	0,0	0,0	0,0	14,0	9,0	0,0	0,0	2.4E-3	2.5E-3
0POP05-EO-E000-4	23,0	0,0	0,0	0,0	0,0	23,0	0,0	0,0	3.2E-3	3.4E-3
0POP05-EO-ES01/F003-1	430,0	0,0	0,0	0,0	29,0	129,0	252,2	20,0	1.4E-3	3.0E-3
0POP05-EO-FRH1-2	196,0	0,0	0,0	41,0	26,0	79,0	36,0	14,0	2.0E-3	3.0E-3
0POP05-EO-FRH1-4	196,0	0,0	0,0	41,0	26,0	79, 1	36,0	14,0	2.0E-3	3.0E-3
0POP05-EO-FRH1-1	1970,14	43,2	147,1	586,2	436,3	492,6	170,0	96,0	7.1E-3	3.0E-3
0POP05-EO-FRH1-3	1772,14	38,0	248,4	241,2	516,3	289,2	314,3	99,0	8.0E-3	3.0E-3
0POP05-EO-	2760,14	0,0	0,0	0,0	355,0	1104,4	892,5	235,5	5.4E-3	3.0E-3

E000-5										
0POP05-EO-F003-2	2760,14	0,0	0,0	0,0	0,0	14,0	1222,4	976,3	3.2E-3	3.0E-3
HEP for overall Feed and Bleed Action									~3.9E-2	~3.0E-2

The result shows that considering all SFs and using SDMS provides a significant improvement over using just the critical SFs. The former approach has much higher discrimination power in identifying the most important tasks and the major contributors to the final HEP values. It also identifies all relevant performance data that can be examined for possible root cause evaluation, and a more defensible final HEP estimate.

4. CONCLUSION

SACADA database is a comprehensive data source for empirical HEP estimation of operator actions within the control room. Developing a SACADA data analysis approach in a systematic and defensible manner is the key to its future use. This research sponsored by the USNRC is a step forward towards this goal. It demonstrates such an approach, provides the required formulation, develops a set of in-house tools for implementation, and performs a comprehensive pilot application. Maturation of these methods, enhancement of the software tools, extending the SACADA database with populating it with additional simulation data, and additional pilot applications including a direct application to a full-scope PRA will greatly boost the benefit of SACADA methodology.

5. REFERENCES

1. Y. James Chang, Dennis Bley, Lawrence Criscione, et al., "The SACADA database for human reliability and human performance," Journal of Reliability Engineering and System Safety, pp 117-133 (125), 2014.

2. Y. James Chang, Carmen Franklin, Laurence Criscione, "Example use of the SACADA data to inform HRA," PSAM 13, October 2016, Seoul, Korea.

3. Katrina M. Groth, Curtis L. Smith, and Laura P. Swirler, "A Bayesian method for using simulator data to enhance human error probabilities assigned by existing HRA methods," Journal of Reliability Engineering and System Safety, pp 32-40 (128), 2014.

4. M.A. Azarm, "A Probabilistic Decision Making based on Prediction of the Computer Models for Accident Analysis," Research Gate, 10/06/2015, http://dx.doi.org/10.13140/RG.2.1.4654.6642.

5. Norman L. Johnson and Fred C. Leone, "Statistical and Experimental Design", John Wiley & Sons, Volume 1, 2nd edition, 1977, New York, Chapter 8.

6. NUREG-2199, "An Integrated Human Event Analysis System (IDHEAS) for Nuclear Power Plant Internal Events At-Power Application," Volume 1, USNRC/EPRI, Rockville, MD, March 2017.