Methodology for Supporting the Determination of Human Error Probabilities from Simulator Sourced Data

Pamela F. Nelson[a], C.R. Grantom P.E.[b], and David Quintanar-Gago[a]
[a] Universidad Nacional Autónoma de México, Facultad de Ingeniería, Departamento de Sistemas Energéticos, Mexico City, Mexico
[b] CRG LLC, Huffman, Texas, USA

Abstract: In this paper, a methodology is presented that processes data from the U.S. Nuclear Regulatory Commission (NRC) Scenario Authoring, Characterization, and Debriefing Application (SACADA) into conditional probabilities for calculating a human error probability (HEP). The approach restructures the data based on SACADA Situational Factors (SFs). These SFs represent the working conditions and performance challenges (i.e., Context) presented to an operating crew. Once the data is restructured a Bayesian network approach is performed that provides the necessary relationships of the Context to specific error modes of each macrocognitive function contained in the SACADA method. A separate Bayesian network model is used for each macrocognitive function that produces the conditional probabilities of a human error based on a specific Contextual situation. The results from a case study indicate that the approach provided in this paper is a reasonable and repeatable method for calculating HEPs from SACADA data. The paper also provides insights and recommendations for further improvements in SACADA relative to accuracy of data input and in the Bayesian modeling approach. Also, uncertainty and other sources of inaccuracy are discussed relative to their impacts on the HEP results.

Keywords: PRA, HRA, HEP, SACADA, Control Room Simulator, Bayesian network.

1. INTRODUCTION

For many years there has been a lack of human performance data from nuclear power plants both in terms of data quality and appropriateness. Human actions in Control Room environments are of keen interest since they directly affect safety and generation risks, as well as being included in plant specific PRAs. Control Room human actions are also highly varied across many plant systems and components and thus require different actions, skills, and knowledge. Estimating human error probabilities (HEPs), under control room conditions is an area where realistic experience and data are necessary for improved HEP estimates to be obtained. In recent years significant progress has been made in developing a process and method for obtaining human performance data from station simulators. This process has been documented in the U.S. Nuclear Regulatory Commission (NRC) Scenario Authoring, Characterization, and Debriefing Application (SACADA) methodology [1]. The data collected in this process is based on actual simulator drills and scenarios performed by licensed operator crews. The scenarios performed by the operating crews are developed by licensed simulator instructors and represent both normal and off-normal plant transients and events. The simulator scenarios can also be structured to cover severe accidents and, in so doing, represent empirical data for human operator actions that are credited in plant specific PRAs.

In the SACADA process, each simulator drill or exercise is divided into basic human actions referred to as Training Objective Elements (TOEs). These TOEs are those human action elements associated with a specific Control Room simulator scenario. Since the SACADA program provides the method for characterizing a scenario into TOEs, the further decomposition of the TOEs into human error macrocognitive functions allows the SACADA program to encode operator actions into the specific human actions necessary to accomplish the TOE's intent. The macrocognitive functions currently included in SACADA are Detection and Monitoring, Diagnosis and Response Planning, Manipulation, and Communication. Once the TOEs are classified into their proper macrocognitive function group, SACADA then allows simulator instructors to assign a set of situational factors (SFs). These SFs characterize the equipment, environment, and working conditions by which Control Room operators must perform their human actions. The set of SFs that constitutes a given environment and working condition by which the Control Room operators perform their tasks is referred to as the “Context” (i.e., the set of SFs under which a TOE is performed). The inclusion of the SFs or Context provides much valuable information; however, the characterization of the human error elements with data taxonomy, as provided by the SACADA, does not in and of itself produce an HEP.

In this paper, a methodology is presented that processes the raw SACADA data into conditional probabilities for calculating an HEP. The approach used restructures the data based on the SFs. These SFs represent the performance challenges (i.e., the Context) presented to an operating crew for a specific simulator scenario. In other words, this represents the environment they must perform a human activity to succeed for a given TOE. This paper’s fundamental premise is that the crew activity and the associated Context define the specific human actions and thus, are the basis for establishing the constituents of human error probabilities. This fundamental premise resolves many issues associated with other HRA methods where the context is considered separately from the action. Those approaches inherently contain uncertainty in attempting to account for variation in context or other performance shaping factors (PSFs). With the SACADA approach this is resolved since the contextual information is now a part of the human action/error. In SACADA, any variations in SFs or PSFs are now specifically part of the human action and the human error probability. Once the SACADA data is restructured, a Bayesian network approach is performed that provides the necessary relationships of the Context to specific error modes for each macrocognitive function contained in the SACADA method. A separate Bayesian network model is used for each macrocognitive function that produces the conditional probabilities of a human error based on a specific Contextual situation.

This paper also examines the various challenges in developing the Bayesian network models from SACADA data and provides recommended solutions to overcome these challenges. The results indicate that the approach provided in this paper is a reasonable and repeatable method for calculating HEPs from SACADA data. The paper also provides insights and recommendations for further improvements in SACADA relative to accuracy of data input and in the Bayesian modeling approach. The performance results shown using this paper’s methodology are considered reasonable for HRA purposes and are expected to further improve once additional operating experience with SACADA processes and data occurs. In the future, the addition of new data from other participating utilities would greatly enhance familiarity with SACADA and its processes and data structure. Also, in this paper, data uncertainty and other sources of inaccuracy are also discussed relative to their impacts on the HEP results.

The benefits of the methodology presented in this paper include the ability to calculate an HEP and to determine the most probable causes of the human error, given a particular Context. In terms of SACADA, a Context is defined by a set of SFs. These SFs are described in SACADA reference documents but include items such as workload, extent of communication, time criticality, etc. In most cases, the SFs can have several states to choose from. For example, in the case of time criticality, the states expansive, nominal or barely adequate are SFs that simulator instructors can select for operator performance to be measured against. Some of the SFs can have multiple states that can co-exist and so this must be considered in the model. This is handled by having separate nodes for the SF in the Bayesian model, with the yes/no states assigned to it. The model enables the Context (set of SFs) to be chosen for an HFE that the HRA analyst needs to quantify. The model then calculates the HEP based on the SACADA data initially provided by the simulator instructors developing the simulator exercise scenarios. The model can also be used to determine the most likely causes of human error given a specific Context. This is valuable information for the simulator instructors to improve training.

2. SACADA DATA PROCESSING

This section of the paper contains a discussion of the processing of SACADA data into a form suitable for further facilitating this paper’s technical approach. A spreadsheet was provided to the research team as the source for all data analysis. It contains two worksheets – CharSATSummaryReport with 2012 rows and 45 columns of data; DebriefSATSumaryReport with 456 rows and 34 columns.

CharSATSummaryReport ("CHAR") contains the training objective elements (TOEs) and related details, including test scenario, year, cycle, malfunction, and malfunction order, as well as the macrocognitive type, Context (SFs), total number of trials, and the number that were performed beyond satisfactorily (SAT+), satisfactorily (SAT), less than satisfactorily (SAT Δ), and unsatisfactorily (UNSAT).

DebriefSATSumaryReport ("DEBRIEF") has a row of data for each test crew that was rated as performing the TOE unsatisfactorily or less than satisfactorily (UNSAT or SAT Δ). It also contains the TOE and related details, allowing the Context information of CHAR to be associated with each row of DEBRIEF. In addition, the worksheet identifies the crew, the operational fundamental weakness, error mode (EM), error cause (EC) data (reflecting the Performance influencing factors (PIFs), and crew comments.

In both worksheets, color-coding of the TOE column provides additional task significance detail, which corresponds to the numerical column titled "Importance." The numerical values and their corresponding description and color are: 1 (Other, white), 2 (Significant, yellow), 3 (Safety Significant, orange), and 4 (Critical, red).

The first step of the analysis was summarizing the data that had been provided. Of the 2012 rows in CHAR, less than half (990) had been assigned a cognitive type (Table 1, Cognitive Type = Null). These entries without a cognitive type (i.e., Null) were not included in subsequent analyses, except where noted.

Table 1. Cognitive Types

Cognitive Type Number	Cognitive Type Description
0	Null
1	Monitoring / Detection
2	Diagnosis & Response Planning
3	Manipulation
4	External Communication

The number of each cognitive type, UNSAT, and SAT Δ is presented in Table 2. For a given cognitive type, the "ROWs with UNSAT" plus the "ROWs with SAT Δ" is less than the "Total UNSAT and SAT Δ". This is because each row of this summary worksheet reports the total for all crews (trials) that performed the TOE. For a given TOE, the value of UNSAT or SAT Δ may be greater than 1 if more than one crew performed unsatisfactorily or less than satisfactorily. That is, each row is a TOE that is performed by multiple crews. The crews performance on the TOE are different.

Table 2. Summary of data in spreadsheet "DATA Set 1 Simulator training data.xlsx" worksheet "CharSATSummaryReport"

Cognitive Type	Number of Rows	Rows with UNSAT	Rows with SAT Δ	Rows with UNSAT *and* SAT Δ	Rows with UNSAT *or* SAT Δ	Total UNSAT	Total SAT Δ	Total UNSAT and SAT Δ
0	1022	53	100	12	141	70	122	192
1	213	12	21	3	30	16	22	38
2	420	27	57	6	78	39	70	109
3	274	33	29	5	57	47	35	82
4	83	9	7	1	15	15	7	22
TOTAL	**2012**	**134**	**214**	**27**	**321**	**187**	**256**	**443**

A new worksheet was made called Debrief-Extended that was "expanded" to include a row for each success, in addition to the existing rows for each UNSAT and SAT Δ. The Context (SF columns) was copied from CHAR, but the error cause and error mode values were set to 0 for the successes. The resulting worksheet has 4358 rows, 187 corresponding to UNSAT, 256 to SAT Δ, and 3915 to successful trials.

This "extended" and "expanded" data, sorted by cognitive type, was used as the source data for Bayesian analysis using the software program "HUGIN". [2] A separate HUGIN model was made for each of the macrocognitive functions.

2.1 Associate Context with Human Error

The extended version of Debrief, Debrief-Extended mentioned above, was created by appending each row's corresponding context data from Char. This was done using the TOE description, scenario, and malfunction columns to identify unique matches. Starting with this Debrief-Extended worksheet, separate worksheets were created for each cognitive type, and analysis focused on these subsets of the data.

Prior to each simulation, the specific test case (combination of TOE, scenario, and malfunction) is assigned a cognitive type (Table 1), and the pertinent situational factors (SFs) are identified from the collection of SFs associated with the cognitive type.

For each cognitive type, the data was sorted by context. Because of the large number of test scenarios and limited number of SFs, there are many cases where distinct TOEs share the same context. Furthermore each scenario was performed by multiple crews, typically 12-15. Each case where one crew performed one scenario is considered a "trial."

The total number of trials for each context was calculated, along with the corresponding number of unsatisfactory (UNSAT) and less than satisfactory (SAT Δ) trials. These totals, calculated including the Overarching (OA) SFs and excluding them, are presented in Table 3.

Table 3: Count of Unique Contexts by Cognitive Type.

Cognitive Type	Total Number of Trials	Number of Unique Contexts (*including* Overarching SFs)	Maximum Number of Trials for One Context	Number of Unique Contexts (*excluding* Overarching SFs)	Maximum Number of Trials for One Context
1	2771	139	157	50	460
2	5485	222	315	78	1590
3	3463	144	108	34	764
4	1072	43	177	10	599

For cognitive type 1, two contexts had three UNSATs. Both of those contexts applied to only one simulation scenario. One had 13 trials (crews), and the other had 14. The case with 13 trials also had two SAT Δs. The context with 157 trials had 1 UNSAT and 1 SAT Δ. When the OA SFs were excluded, as expected, there were fewer unique contexts and a correspondingly larger number of trials for most contexts. The context with 460 trials had 2 UNSATs and 6 SAT Δs. The maximum number of UNSATs associated with one context was 4; the context had only 26 trials (2 simulation scenarios).

The results are similar for cognitive type 2. When OA SFs are included, one context, with 31 trials (2 simulation scenarios) had 8 UNSATs. The next-highest number of UNSATs associated with a single context was 3, for two different contexts. One had 84 trials, and the other had 315. When OA SFs are excluded, the context with 31 trials still has the maximum number of UNSATs, 8. A context with 1590 trials had 7 UNSATs and 12 SAT Δs.

For cognitive type 3, including OA SFs, two single-simulation scenarios had 5 UNSATs. One had 16 trials. The other, which had only 14 trials, also had 2 SAT Δs. Excluding OA SFs produced decidedly different results. A context with 764 trials had the maximum number of UNSATs, 9, along with 3 SAT Δs. Contexts with 195 and 221 trials had 6 UNSATs, and trials with 359 and 685 trials had 5 UNSATs.

For cognitive type 4, including OA SFs, a context with 26 trials (2 simulation scenarios) had 4 UNSATs. When OA SFs were excluded, over half of the UNSATs (9 of 15 for cognitive type 4) were associated with a single context. This context, which also had 4 SAT Δs, was associated with over half of the trials (599 of 1072). These results are presented in Table 4.

Table 4. Number of trials, UNSAT, and SATΔ, and UNSAT probability (%) by cognitive type.

Cognitive Type	Number of Trials	Total UNSAT	Total SAT Δ	Total UNSAT and SAT Δ	UNSAT Probability (%)
0	12993	70	122	192	0.54
1	2771	16	22	38	0.58
2	5485	39	70	109	0.71
3	3463	47	35	82	1.36
4	1072	15	7	22	1.40
TOTAL	**25784**	**187**	**256**	**443**	**0.73**

The results show that increasing trials results in a decrease in the rate of UNSATs. This may indicate or imply that increased training repetition on similar contexts with simulator crews produces reduced failure rates. In other words this may represent a quantitative measure of the effectiveness of training.

.

3. MACROCOGNITIVE FUNCTION MODELS

The development of Human Error Probabilities using the data produced through the SACADA program was done for each macrocognitive function. A Bayesian network approach is used that accounts for the various SFs and their influence on Control Room human errors and their associated contributing error modes.

Big Data and Artificial Intelligence alone cannot produce models that are suitable for calculating HEPs, this is due to the fact that the effect of SFs on human activity is a fundamentally causal question, and even the most sophisticated predictive models cannot correctly compute the causal effect of SFs. However, the fact that SACADA data is from a simulator environment, where causal effects on SFs can be directly observed and recorded, acts to reduce uncertainty in the resulting HEP calculations; however, uncertainty still exists since the simulator environment is not exactly the same as an actual control room environment. This adds uncertainty; however, it can be argued that uncertainty may be more or less in an actual control room environment as the impact to a real asset (the nuclear power plant) may result in either better or worse performance.

While models based on Bayesian networks do not automatically overcome this issue, they do allow us to directly encode causal assumptions from expert knowledge [3]. With that, we can correctly estimate the causal effect of SF effects from historical data, and we assert that over time, and with improved SACADA processes, this will provide a significantly improved human error probability basis, and ultimately produce adequate human error probabilities. Factor Analysis as well as expert knowledge can be used to discover latent factors that drive errors about operator behavior to improve the accuracy of the models. But even without this addition and with the limited data sets, the potential to develop the HEPs directly from the data provides an undeniable improvement to the HRA process.

As described above, the SACADA data was provided in Excel spreadsheets. Each row contains the information for each TOE for each crew, that is, all Situational Factors, including OA factors. OA factors are situational factors that are common to all the macrocognitive functions and appear to be driven by a latent factor that we will call ‘crew dynamics’ for now. The OA factors are workload, time criticality, communication within the control room and to other onsite operators and a grouping of miscellaneous factors, such as noisy background, coordination with onsite personnel, task demand on memory, and others. Thus, if a TOE is included in a scenario that is carried out by 12 crews, there will be 12 rows of information. If one crew had an UNSAT, that row will also contain information about error modes and error causes, that can be used for human performance improvement, as described in Section 3.3. The input is an Excel file that contains a column for each node in the model (not necessarily the same columns as the original data file because of the fact that some of the states can occur simultaneously. For these cases, the states should be divided out into separate nodes) and a row for each crew for each TOE, as explained above. As more data is included, it can be added to this sheet. With the data available now from data set 1, there are 2771 rows, for Cognitive Type 1 (COG 1), 5485 for COG 2, etc. Once the input file is ready, we can run the various models, compare to the results the data provides to results calculated from previous empirical studies as well as conduct queries.

As described by HSE [4] the basics of a qualitative assessment are required to demonstrate a good understanding of the tasks, and this is what the SACADA provides; that is, the process reflects both the instructor’s knowledge about the SFs that will affect the operators’ performance (in the coding of the TOEs) as well as a process to record dialogue among operators and instructor (in the debrief section) that provides invaluable information about the problems and issues encountered during scenarios, expressed as PIFs. Despite the fact that this is a simulation, it represents the best estimate of operator behavior during real accidents. As was the case for thermal hydraulic codes in the sixties the capabilities of the codes were strongly limited by lack of experimental data, details of modeling and the capacity of the computers, which is about where we find ourselves for human error analysis presently. SACADA provides a source of data to develop data driven methodologies and update the HRA process to incorporate real data into the HEPs.

3.1. Bayesian Network for Macrocognitive Functions

A preliminary model was developed for each of the macrocognitive functions. In the case of monitoring/detection, it consists of three error modes, Alarm issue, Indicator issue, and Other error mode, which are the error modes defined in SACADA for this macrocognitive function.

Figure 1 reflects the simple preliminary HUGIN model developed for the Monitor/Detection macrocognitive function, and includes the SFs (i.e., Context) impacting the error modes). For the objective of this paper, that of determining the human error probability, the model will consist of the Context (the set of situational factors) impact on the Error Modes, which in turn lead to the error probability of the human activity involved in the macrocognitive type being analyzed. Thus, the model consists of the SFs leading to the EMs and thus error probability, as shown below.

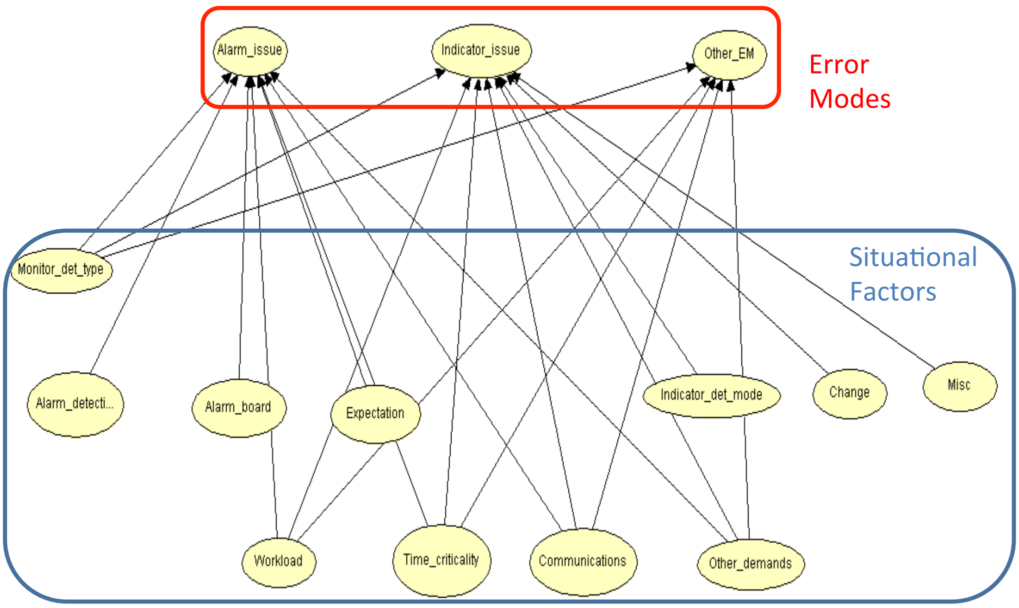

Figure 1. Model for Macrocognitive Function: Monitoring/Detection

The models for the other macrocognitive functions have been developed as well; the SFs included in the models are presented in Table 5. The way the models are used together is discussed in Section 5.

Table 5. Situational Factors for each Macrocognitive Function

	Detection/Monitoring						**OA**
SFs for Detection/ Monitoring	Alarm detection mode	Alarm board status	Expectation	Indicator detection mode	Change	Mimics	Workload Time-criticality
	Diagnosis						
SFs for Diagnosis	Diagnosis basis	Familiarity	Integration	Information specificity	Infor-mation quality	Diagnosis outcome	Commun-ication
	Response Planning						
SFs for Response Planning	Decision basis	Uncertainty	Familiarity	Outcome			Other

	Manipulation						
SFs for Manipulation	Type of action	Location	Guidance	Recover-ability	Miscel-laneous		

3.2 Models for Human Performance Improvement

The Context also impacts the error causes and the model shown in Figure 2 can provide the probabilities of error causes, given a specific Context. This model goes beyond the purpose of the data collection use in HRA for PRA purposes. That is, this model provides information that can be extremely valuable to the training center at a nuclear power plant to improve training and thus human performance. What can be observed in this figure is the use of the "debrief" section of the SACADA process. This data is the result of the operating crew and simulator instructor's debrief meeting after the scenarios have been run on the simulator. The SACADA system leads them through the questions about error modes and causes for any TOEs that they have identified as UNSAT or ΔSAT. The causes are defined by the performance influencing factors (PIFs) included in the SACADA system. The OA causes include scenario specific causes and person specific causes, such as knowledge gap, slow, lack of questioning attitude, failing to stop, think, act, and review, rushing, and distracted. Additionally, there are error causes specific to the macrocognitive function of detection / monitoring, which include identifying of the error was caused by multiple or unexpected alarms, label, mimic or display issues, among others. Due to the debriefing process, the operators should be encouraged to openly express the causes, and the instructor is present to support the evaluation. The instructor also reviews the final inputs and makes changes where considered necessary.

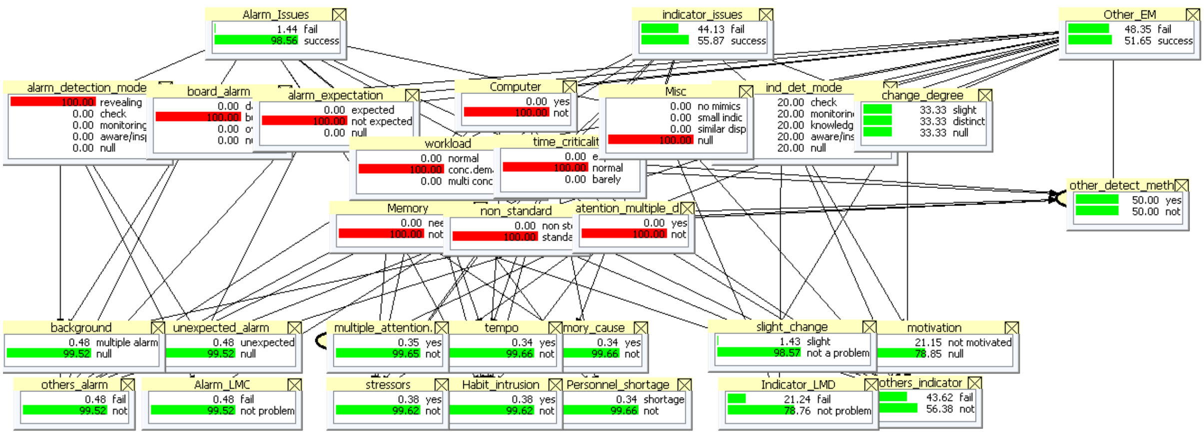

Figure 2. Detection / Monitoring Model for Queries

4. QUANTIFICATION APPROACH

The parameters necessary to know before any data is loaded are the SF state probabilities and the prior probability of each Context. The probabilities of the SF states are based on plant operating experience or expert judgment, these are known as the probability mass function (PMF) of each variable. For example, it is fairly easy to estimate that 60% of the time the alarm board is normal, 35% loaded and 5% overloaded from plant experience. On the other hand, the prior probabilities for each context input can be based on expert judgment, some other HRA method (e.g., SPAR-H as suggested by Groth [5]), or even another approach (e.g., weight factors developed from SACADA data [6]). While this may seem like a significant burdensome effort to calculate these priors, there are two reasons that this is not the biggest obstacle to overcome: 1) this prior only is a problem for the Contexts without trials (which can be solved by including these Contexts in the simulator scenarios); and 2) the fact that over time, these prior probabilities will come directly from SACADA data.

For the Contexts with trials, the updated probability for each Context is calculated by the HUGIN software using a counting-learning algorithm to update the prior from the SACADA input file, shown in Eq. 1.

((Prior probability * prior experience) + failures)/(prior experience + no. of trials)   Eq. 1

For example, if we have 12 trials with the context shown in Fig. 3 (6,3,3,3,0,0,0,3,3,3,6), and the prior was assigned 0.0057 (corresponding to the average UNSAT ratio for cognitive type 1from SACADA data), and prior experience of .001, the software adds the 12 trials to the experience, resulting in 12.001 for the updated experience and learning algorithm calculates the updated probability of the Context from Eq. 1, yielding 0.083.

When the HEP of the alarm detection aspect of an HFE is required and the human activity involves this specific context; the evidence is set by selecting those applicable states of the SFs and the HEP is shown in the resulting Alarm_issue node. This is discussed further in the case study presented in Section 5.

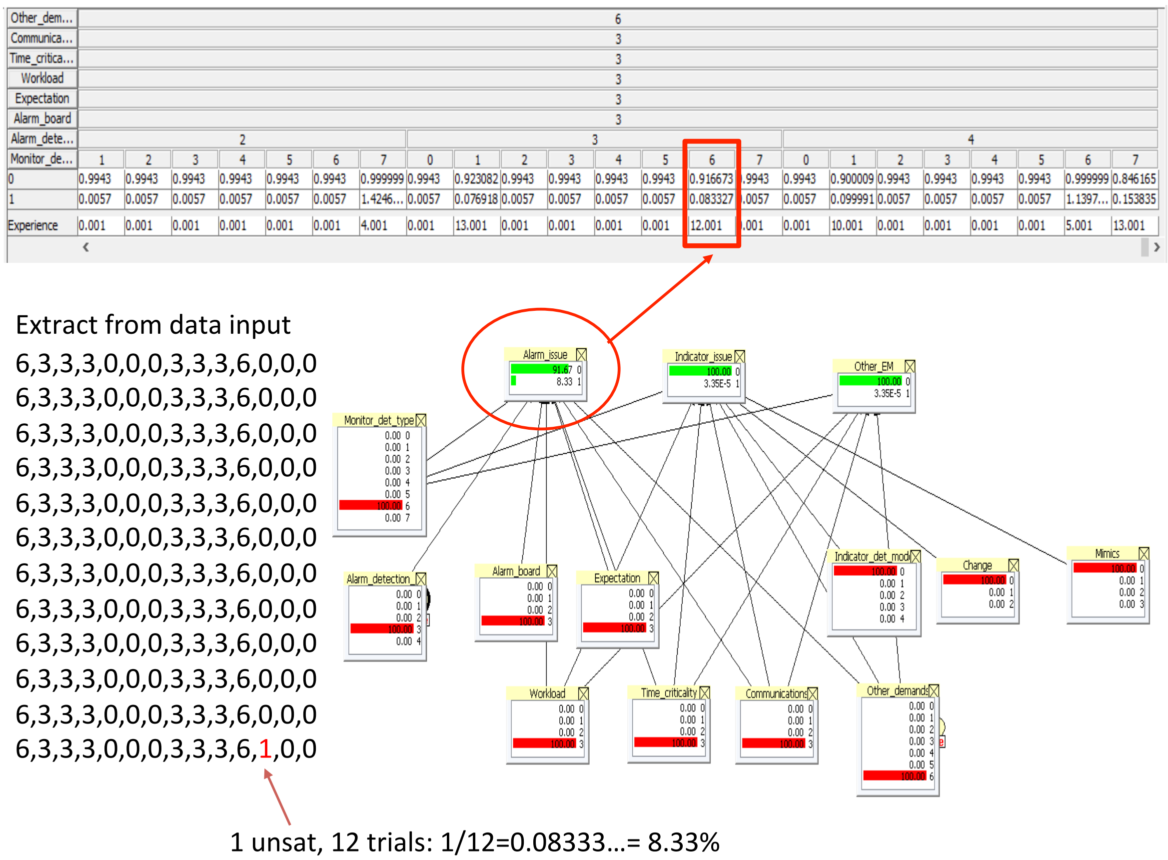

Figure 3. Data input for a Specific Context

5. CASE STUDY

At the beginning of this simulator scenario the plant was operating at 100% power. In a situation of total LOFW, the reactor core is cooled by the remaining water in the SGs. If feedwater cannot be re-established, the SGs will eventually become empty and unable to cool the core. It is important to establish another means of core cooling before the SGs are empty. This is done by initiating Bleed & Feed (i.e., starting SI and opening the Pressurizer PORVs). The criteria for starting Bleed & Feed (B&F) in the FR-H.1 procedure are that the WR level should be less than 12% in two out of three SGs, or that the reactor pressure should be high due to loss of secondary heat sink. To be able to start B&F in time, the crews need to monitor the SG levels. According to procedure FR-H.1, Bleed and Feed (B&F) shall be established when the WR level on any two SGs are less than 50%. The HFE is defined as: Failure to establish feed, given that the crews do not manually trip the reactor before an automatic reactor trip occurs.

The human actions considered are found in the following TOEs: COG1:Monitoring Critical Safety Functions Commences actual crew monitoring of Critical Safety Functions. COG1 Recognizes and informs Unit Supervisor of red path on Heat Sink., COG2 Transitions to 0POP05-EO-FRH1, Response to Loss Of Secondary Heat Sink when Addendum 5 is complete. COG3 Trip RCPs per FRH1 CIP or Step 2 due to inadequate WR S/G level. (<50% on 2 or more SG) and COG3 Initiate RCS bleed and feed so that the RCS depressurizes sufficiently for HHSI pump injection to occur. Embedded in these steps is the action to open PORVs, but should be separated and considered as another human action. Figure 4 shows the SFs identified for this last human action.

TOE COG3	COG type	Type of Action	Location	Guidance	Recoverability	Miscellaneous	Workload	Time Criticality	Comm.	Other
Initiate RCS bleed and feed so that the RCS depressurizes sufficiently for HHSI pump injection to occur	3	2	1	1	2	3	2	2	3	0

Figure 4. SF States for one COG3 Human Action in Feed and Bleed HFE

The SF states are determined from those identified in the data, as shown in Fig. 4 for one human action, for each of the macrocognitive functions and entered as evidence to the Bayesian models. In this way the HEP for the Context is retrieved and shown in the HEP monitor. This is shown in Fig. 5 for the calculation of the second manipulation HEP.

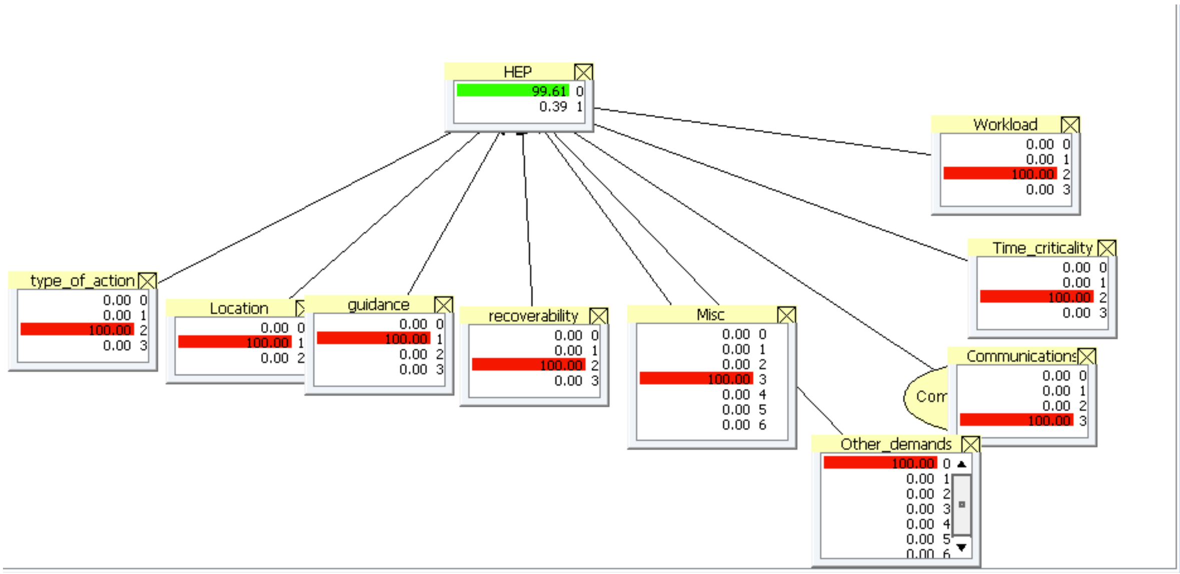

Figure 5. HEP for Manipulation Failure in Feed and Bleed

The total HEP for Feed and Bleed is the sum of each of the HEPs for the various macrocognitive functions of the HFE, as shown in Table 7. In this case, the result is 0.0613. This result falls in the uncertainty range calculated in the International HRA Empirical Study [7].

Table 7. Feed & Bleed Case Study Results

MCog 1	MCog2a	MCog2b	MCog3	HFE HEP Total
0.0033	0.0	0.0530	0.0011 + 0.0039	0.0613

5.1. SACADA Process Going Forward

It is seen from the above technical approach and case study that SACADA offers a transformative opportunity in the field of HRA. SACADA is still in its infancy and requires further refinement both in practice and in process. To further support SACADA development, a conceptualized SACADA-HRA process includes identifying the appropriate TOEs for the HFE of interest and the SFs are located for each of the macrocognitive functions, which can be more than one of each type. In this way the Contexts are defined, and the evidence is set accordingly. The HEP is obtained from the corresponding Bayesian model and can be registered in a library as generic data. The plant specific data is used to update the generic values. Important next steps would be to develop a library of human actions plus associated Context that would form the basis for HRA to be used in plant Specific PRAs, but that could also be used for improving training and human performance. As more plant simulators contribute to the SACADA database, the increased data will reduce overall uncertainties and also allow new methods to improve Operation's training and enhance Operator understanding of human reliability and performance.

Although the authors of this study consider uncertainties to be reduced using the SACADA method presented, there remain uncertainties that will need to be addressed. Some sources of uncertainty are associated with the SACADA data collection process and the assignment of SFs for simulator drill scenarios. Some uncertainties in these areas will be reduced with improved documentation and training on SACADA processes. This training should be targeted to simulator instructors. Items such as the following have been noted by the authors to be areas where uncertainties exist in the current process:

- TOE definition and degree of granularity in TOE definition,
- Improved clarity and more specific criteria for SF selection,
- Improved criteria for specific SF options and,
- Improved refinements in SACADA procedural processes (e.g., recording action times and specific procedure steps).

Also, the debriefing task where control room crews and simulator instructors meet to discuss simulator scenario performance could be improved so that there is improved criteria for SAT, UNSAT, and SAT Δs. It should be noted that some traditional sources of HRA uncertainty are better addressed through the SACADA method. Indeed, the actual actions of crews are captured with a robust basis (actual operational procedural actions).

6. CONCLUSIONS

This paper demonstrates that SACADA data and processes can be a proper foundational basis for developing HEPs/HFEs using simulator data. As more data is collected over time, this foundational basis will continue to strengthen. Improvements in data processing and in SACADA procedural processes will continue to facilitate reduced uncertainties in HEP estimates. In fact, it is plausible that SACADA methods could be extended to other crew types such as maintenance crews, fire brigades, emergency preparedness, and others. This paper has shown that quantified methods for HEPs supporting the development of HFEs in PRAs are well founded and should be continued. Although uncertainties in the estimates remain, the improvement in HRA through SACADA more than offsets these uncertainties and also allows some of the uncertainties to be better defined and therefore more addressable with future efforts.

Acknowledgements

Work Supported by the U.S. Nuclear Regulatory Commission (NRC), Office of Nuclear Regulatory Research, under Contract/order number: NRC-HQ-60-17-C-0004, Trinitek Services, Inc.

References

[1] Chang, Y.J., Bley, D., CriscioneL., Kirwan, B., Mosleh, A., Madary, T., Nowell, R., Richards, R., Roth, E. M., Sieben, S., & Zoulis, A. (2014). "*The SACADA Database for Human Reliability and Human Performance"*, Reliability Engineering & System Safety, 125, 117-133.
[2] HUGIN Researcher software,version7.6, 2011. URL /www.hugin.com.
[3] BayesiaLab*, Causal Inference in Marketing Science: Media Mix Modeling and Optimization with Bayesian Networks,* http://www.bayesia.com/2018-03-02-probabilistic-reasoning-under-uncertainty.
[4] Health and Safety Executive (HSE), Failure Rate and Event Data for use within Risk Assessments, http://www.hse.gov.uk/landuseplanning/failure-rates.pdf, 28/06/2012.
[5] Groth, K., Swiler, L., "*Bridging the gap between HRA research and HRA practice: A Bayesian network version of SPAR-H"*, Reliability Engineering and System Safety, 115, 33–42 (2013). .
[6] Azarm M.A., Kim I.S., Marks, C., and Azarm, F.,, "*Analyses Methods and Pilot Applications of SACADA Database*," PSAM14, Los Angeles, CA, Sept. 2018.
[7] NUREG-2127, The International HRA Empirical Study, Lessons Learned from Comparing HRA Methods Predictions to HAMMLAB Simulator Data, published August 2014.