Nuclear Engineering and Technology 54 (2022) 896–908

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

journal homepage: www.elsevier.com/locate/net

Original article

SACADA and HuREX part 2: The use of SACADA and HuREX data to estimate human error probabilities

Yochan Kim [a], Yung Hsien James Chang [b], Jinkyun Park [a,*], Lawrence Criscione [b]

[a] Korea Atomic Energy Research Institute, South Korea
[b] U.S. Nuclear Regulatory Commission, USA

ARTICLE INFO

Article history:
Received 12 April 2021
Received in revised form
9 July 2021
Accepted 2 September 2021
Available online 4 September 2021

Keywords:
PRA/PSA
HRA
Simulator data
HEP estimation
SACADA
HuREX

ABSTRACT

As a part of probabilistic risk (or safety) assessment (PRA or PSA) of nuclear power plants (NPPs), the primary role of human reliability analysis (HRA) is to provide credible estimations of the human error probabilities (HEPs) of safety-critical tasks. In this regard, it is vital to provide credible HEPs based on firm technical underpinnings including (but not limited to): (1) how to collect HRA data from available sources of information, and (2) how to inform HRA practitioners with the collected HRA data. Because of these necessities, the U.S. Nuclear Regulatory Commission and the Korea Atomic Energy Research Institute independently developed two dedicated HRA data collection systems, SACADA (Scenario Authoring, Characterization, And Debriefing Application) and HuREX (Human Reliability data EXtraction), respectively. These systems provide unique frameworks that can be used to secure HRA data from full-scope training simulators of NPPs (i.e., simulator data). In order to investigate the applicability of these two systems, two papers have been prepared with distinct purposes. The first paper, entitled "SACADA and HuREX: Part 1. The Use of SACADA and HuREX Systems to Collect Human Reliability Data", deals with technical issues pertaining to the collection of HRA data. This second paper explains how the two systems are able to inform HRA practitioners. To this end, the process of estimating HEPs is demonstrated based on feed-and-bleed operations using HRA data from the two systems.

1. Introduction

In complex systems such as nuclear power plants (NPPs), the reliability and performance of human operators interacting with the machine systems play an essential role in the risk of the system as a whole. Human reliability analysis (HRA) identifies significant human errors in the operations and maintenance of complex systems and estimates their probabilities. For enhancing the credibility of HRA activities and results, it is crucial to obtain human reliability data which provide a technical underpinning of the estimated human error probabilities (HEPs) [1]. The HRA community has collected various kinds of human reliability data to provide empirical evidence. Table 1 shows examples of previous works that collected human reliability data from simulator training.

Human reliability data collection often requires considerable expertise and resources to handle the following complicating factors. First, human errors in fields engaging highly trained operators are infrequently observed. To collect data for such kinds of human error, many tasks or events should be attempted. Second, the contexts that contribute to error occurrences, as represented by the performance influencing factors (PIFs) in HRA, are diverse. The identification of such contexts necessitates an in-depth understanding of human and machine interactions. The collection and analysis of data to estimate the PIFs' effects on HEPs could be thus resource-intensive. Lastly, the administrative process for gaining access to PIF information from the simulated or real-world incidents could be tedious.

As two of the examples listed in Table 1, the SACADA (Scenario Authoring, Characterization, And Debriefing Application) and HuREX (Human Reliability data EXtraction) systems have recently produced a relatively large amount of human reliability data from full-scope simulator training records. SACADA allows plant staff to select key tasks and supports them in evaluating their successes or failures during operator simulator training. HuREX invites HRA domain experts to review raw data, such as videos and chronological logs that archive the responses of systems and operators,

* Corresponding author.
E-mail addresses: James.Chang@nrc.gov (Y.H.J. Chang), kshpjk@kaeri.re.kr (J. Park).

https://doi.org/10.1016/j.net.2021.09.003

Table 1
Examples of collecting operator performance information for HRA.

Purpose	Data	Reference
HRA method development	ORE (Operator Reliability Experiments)	[2–4]
	Observations for MERMOS (Méthode d'Evaluation et de Réalisation des Missions Opérateurs pour la SÛreté)	[5,6]
	SACADA (Scenario Authoring, Characterization, and Debriefing Application)	[7]
	HuREX (Human Reliability data Extraction)	[8]
	CORE (Computerized Operator Reliability and Error)	[9]
Update of HEPs in PRA models	Simulator Data Collection in the Czech Republic	[10]
Evaluation of HRA methods	International HRA empirical studies	[11–14]
	US HRA empirical studies	[15]

and to evaluate the human error occurrences. The first paper described the similarities and differences between the SACADA and HuREX systems in detail. This (second) paper discusses the use of the two data collection systems for HEP quantification. This paper demonstrates the process of HEP quantification using the data from the two systems to exemplify the possible use of simulator data in HRA and to compare the characteristics of both data sources and related processes.

A large portion of studies generated human reliability data geared for specific HRA methods, and hence, the application process of the collected data was also dependent on specific HRA methods. For example, Moieni et al. [2–4] conducted the operator reliability experiments from the simulators of US plants and presented a quantification method for a failure probability to initiate a timely response based on the experiment data. UJV Rez researchers collected human reliability data from a simulator and updated the HEPs quantified by the existing HRA models based on the data insights [9]. On the other hand, the SACADA and HuREX systems collect data from full-scope simulators based on general human performance taxonomies instead of specific HRA methods. The current paper, therefore, tries to reveal how to use these kinds of data to practically estimate the HEPs of human failure events (HFEs). It is worth noting that while the processes presented in this paper describe use-cases for the two selected data sources, various approaches are available depending on the application purpose.

2. Use of SACADA data for quantifying the HEPs of HFEs

2.1. Data structure of the SACADA system

The scenario structure of the Westinghouse pressurized water reactors for operator simulator training was the blueprint for developing the SACADA data structure, which consists of three hieratical components: scenario, plant malfunction (malfunction, hereafter), and the training objective element (TOE). A scenario typically consists of a few malfunctions, and a malfunction consists of a few TOEs. The TOEs are the data units of SACADA's HEP quantification.

The TOEs of a malfunction represent the expected activities by operators in response to the malfunction and typically cover the macro-cognitive functions of detection, understanding, decision-making, and action. The number and composition of the macro-cognitive functions vary between malfunctions, and operator performance is evaluated against the performance expectations of the TOEs. The TOEs of a malfunction include the operator responses directly addressing the malfunction (e.g., open a valve) and the emergency plan requirements (e.g., declare the emergency action level and make a public address announcement to alert the onsite workers about the plant condition). Scenario developers specify the TOEs based on procedural instructions for effective and efficient discussion of operator performance in simulator training, but the scope of each TOE could correspond to a procedural instruction, a procedural step, or a few procedural steps. Depending on the training emphases, the time available for training, and other considerations, the scenario developers may adjust the TOEs per malfunction. As a result, the TOEs of the same plant malfunction could vary between scenarios. The following shows the 14 TOEs specified for a steam generator tube rupture (SGTR) event with a stuck-open SG power-operated relief valve (PORV):

1. Report the alarm(s)/indications of the SGTR to the unit supervisor using direct communication.
2. Manually initiate reactor trip and safety injection (SI) when directed or when the pressurizer (PZR) level decreases to < 8%.
3. Direct reactor trip and enter the emergency operating procedure (EOP-0) “Reactor Trip or Safety Injection."
4. Direct/ensure that the immediate actions of EOP-0 are completed.
5. Transition to EOP-3 ″Steam Generator Tube Rupture."
6. Identify the ruptured SG.
7. Control the PZR pressure to minimize break flow.
8. Identify the failed-open SG PORV.
9. Isolate the failed-open SG PORV.
10. Complete isolation of the ruptured SG:
    a. Isolate auxiliary feedwater (AFW) to the ruptured SG;
    b. Isolate blowdown from the ruptured SG;
    c. Adjust the PORV setpoint (1260–1265 psig) on the ruptured SG;
    d. Close the main steam isolation valve on the ruptured SG.
11. Properly select and maintain the target temperature for cooldown based on the chart provided in EOP-3.
12. Initiate the cooldown.
13. Depressurize the reactor coolant system (RCS) to meet the SI termination criteria before either of the following occur: (<9 min from target temp reached)
    a. SG PORV or safety valve opens;
    b. SG narrow range level goes off-scale high.
14. Terminate SI.

2.2. Assumption for quantification

SACADA uses the context similarity approach [7] in analyzing data to inform HEP estimates. The assumption is that different tasks with the same context have similar human reliability. Therefore, the grouping of data points to calculate human reliability is based on the context instead of the actual task (e.g., secure emergency diesel generator 1A). As stated previously, SACADA's data units are TOEs, and each TOE has its context, which is a combination of the cognitive type and the PIF statuses. Context similarity assessment between TOEs assesses the similarity of cognitive types and PIF statuses of the TOEs.

Each TOE has one and only one of the following four cognitive types to indicate the TOE's primary cognitive demands: detecting/monitoring information, understanding the situation, deciding how to respond, and taking action to change the scenario course. Each PIF has discrete statuses. For example, the *Cognitive basis* PIF distinguishes cognitive demands by three discrete statuses: skill-based, procedure-based, and knowledge-based. The *Workload* PIF also has three discrete statuses: normal, concurrent demands, and multiple concurrent demands. Human reliability experts and nuclear plant instructors jointly developed the PIFs to cover all the significant factors affecting operator reliability in the simulator exercises. The PIF statuses aim to be intuitive to the operators with distinguishable key differences because it is the nuclear plant operators themselves who enter the SACADA data.

When TOEs have the same cognitive type and PIF statuses, then the TOEs have a perfect context similarity. Conversely, the TOEs having different cognitive types have no context similarity. The TOEs having the same cognitive type but different PIF status combinations have a context similarity between no similarity and perfect similarity. Using the data points from the TOEs with perfect context similarity to calculate HEPs is desirable but not practical because the number of PIF status combinations ranges from 60,000 to 3 million between SACADA's four cognitive types [16]. The data points of a specific context are likely insufficient to calculate the HEP with a good statistical basis. SACADA data analysis only groups the TOEs with the same cognitive type to calculate HEPs. Step 3 of Section 2.3 discusses how to evaluate the context similarity.

2.3. Process of HEP calculation

This section discusses the process of using SACADA data directly to calculate the HEP of an HFE. The analysis process includes the following steps.

Step 1: Break down the HFE into TOEs: To calculate the HEP of an HFE, the analyst first represents the HFE by the TOEs critical to the success of the HFE. The SACADA database provides ample examples to specify the TOEs.
Step 2: Specify the context of the TOEs: Analysts use the SACADA taxonomy to specify the context of the TOEs identified in Step 1. The context is a combination of the cognitive type and the PIF statuses.
Step 3: Calculate the HEPs of the TOEs: This step has multiple approaches. Azarm et al. [17–19] use the statistical significance test method, while Chang [16] uses expert judgment to search the SACADA database to identify the TOEs with similar contexts. Nelson et al. [20–23] use a Bayesian belief network (BBN) based on the SACADA taxonomy to calculate the HEPs by the TOEs with the same contexts. Groth et al. [24] use SACADA to provide data supplements to a BBN model with a much larger scope than the main control activities. Groth's model [24–26] uses SACADA to provide data supplements to a BBN implementing a taxonomy different from SACADA, so this model is excluded from discussion in the following. Once the TOEs are identified, the HEP is simply the number of failures divided by the total number of TOEs.

Azarm et al. [17–19] and Chang [16] both implement the same process, starting with the PIF statuses (PIF combination) of the TOE of analysis then gradually relaxing the PIF restriction (i.e., increase PIF coverage) to identify the PIF combinations with similar context as the TOE of analysis. Azarm et al. [17–19] use HEP distributions to determine the similarity with mathematical rigor. The contexts of two PIF combinations are similar if their HEP distributions have a large overlap and their distribution means are close. This large overlap is satisfied if 90% of the 90% credible interval (i.e., interval between the 5th and 95th percentiles) of the original HEP distribution is covered by the 90% credible interval of the other HEP distribution. This implies a probability of 81% (0.9*0.9) that a sample taken from another PIF combination residing in its 90th percentile will be within the credible interval of the original PIF combination. This determination process is performed based on the statistical significance test technique.

Chang [16] relaxes PIF restrictions by first applying a screening rule to determine the PIFs whose scope cannot be relaxed quickly. Many of the PIFs in SACADA have three or four statuses representing different levels of adverse effects on human reliability. For example, among the three statuses of the *Workload* PIF (see Section 2.2), the 'multiple concurrent demands' is the worst status on human reliability. The screening rule is that two TOEs do not have a similar context if, for any PIF, the PIF status of a TOE is at the worst status and the PIF status of another TOE is not at the worst status. Therefore, TOEs with a *Workload* PIF status of 'multiple concurrent demands' do not have context similarity with TOEs with a *Workload* PIF status of 'normal' or 'concurrent demands'. However, TOEs with a *Workload* PIF status of 'normal' may or may not have context similarity with those with the 'concurrent demands'. Determination is made by expert judgment based on the HEP difference between the two PIF combinations calculated using the SACADA data.

Nelson et al. [20–23] use HUGIN software [27] to build a BBN using SACADA taxonomy and data. The SACADA data points having identical PIF combination with the TOE of analysis are identified, and their failure probability, represented by the number of failures divided by the total number of TOEs, is employed as the evidence in the Baysian update (see Equation (1)). HUGIN software applies the counting-learning algorithm [28,29] (see Equation (1) below) to calculate the posterior probability.

$$\textit{Posterior Probability} = \frac{(\textit{Prior Probability} \times \textit{Prior Experience}) + \textit{Number of Failures}}{\textit{Prior Experience} + \textit{Total Number of TOEs}} \tag{1}$$

Analysts provide the prior probabilities and prior experiences for all PIF combinations, ranging from 60,000 to 3 million combinations between the four macro-cognitive functions in SACADA. The prior probability indicates an HEP estimated from the prior knowledge such as literature or expert expectations, while the prior experience is the weight assigned to the prior probabilities. Assigning a very small prior experience (e.g., 0.001) in HUGIN lets the posterior probabilities strongly depend on the available data if data are available.

Step 4: Calculate the probabilistic sum of the HFE's HEP: Use the HEPs of the critical tasks from Step 3 to calculate the probabilistic sum of the HFE's HEP with Equation (2).

$$HEP\ (HFE) = 1 - \prod_{i=1}^{N}[1 - HEP(TOE_i)],\ for\ N\ is\ the\ number\ of\ TOEs \tag{2}$$

3. Use of HuREX data for quantifying the HEPs of HFEs

3.1. Data structure of the HuREX system

As mentioned in the first paper of this series, the HuREX system distinguishes primitive tasks (PTs), referring to the fundamental tasks to be performed under procedure instructions, and uses PTs as the main data points (Section 3.3.1 gives examples of PTs). Each data point contains information on the PT performance result and its context. The PTs are classified into information gathering (IG), situation interpreting (SI), response planning (RP), and execution (EX) activities. HuREX uses PIF variables to represent the contexts of a PT. The states of the PIF variables are determined by analyzing the characteristics of the scenario, procedure, procedure step, procedure sentences, and PTs.

The HuREX data are generated based on three information-gathering templates (IGTs): response, overview, and unsafe act (UA) templates, as shown in Fig. 1. The IGTs are connected using the simulation indices and unsafe act codes to generate an integrated data table to document a simulation run. The information of the other simulation runs is added to the same data table. All data collected to date from the Advanced Power Reactor − 1400 MWe (APR1400) simulator are placed in an integrated data table that consists of 45,000 rows and 60 columns. Table 2 shows the example contents of an integrated table in which each row is a data point. A data point contains the information of PT type (defined in HuREX), performance result (success or failure), error mode (i.e., errors of omission or errors of commission), and the context (PIF variables), as shown in the columns of Table 2.

Although the point estimation of an HEP with respect to a specific PT type can be directly determined from the HuREX table, its contents are also useful for deriving statistical insights based on diverse analysis techniques. The statistical analysis results from the data table are converted into decision tree forms or representative statistics for use in actual HRA applications.

3.2. Assumptions for quantification

Several assumptions considered for the use of the HuREX data in HRA applications are as follows. First, the HFE to be analyzed has to have the same operating environment, such as procedure and human–machine interface, as the data collection environment of the HuREX data. If the data collection environment and the actual environment in which the HEP is estimated are different, the human reliability gaps in those environments should be analyzed and reflected in the HEP. Second, the goals of any particular HFE are accomplished by following the procedures; therefore, the PTs that elucidate the HFE are identified based on the procedural instructions. If the event requires an action that is not described in the procedure, the action is not credited as feasible unless it is proved that the action is skillfully carried out through sufficient training. When a skill-of-the-craft action is analyzed, it is assumed that there is an imagined procedure for the action. The third assumption defines the HEP of an HFE as the sum of the probabilities that critical PTs fail. In general, the success of an HFE requires the successes of various PTs. To aggregate the failure probabilities of these PTs for quantifying the HEP of the HFE, this method applies the rare event approximation. Lastly, critical PTs are assumed to be those that are essential to satisfy the success criteria of a given HFE, such as a device manipulation or transferring to a procedural step that contains other critical PTs.

3.3. Process of HEP calculation

In this paper, the process of calculating an HEP using HuREX data is based on the EMBRACE (EMpirical data-Based crew Reliability Assessment and Cognitive Error analysis) method. The detailed process is illustrated in Ref. [30], which describes how to quantify the failure probability of cognitive errors during the performance of proceduralized tasks. Fig. 2 shows an overview of the HEP calculation process using the HuREX data. The uppermost sequence shows the HEP calculation process, and the two blocks in the middle show the statistical treatments of the HuREX data table to support the HEP calculation process.

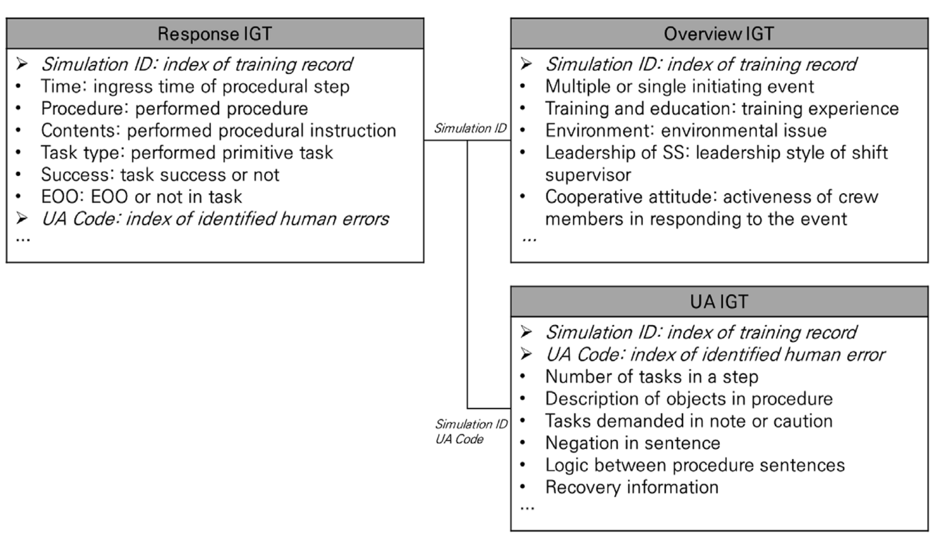

Fig. 1. Data association between different information-gathering templates.

Table 2
Cell examples of the HuREX data table.

Scenario ID	Procedure	Step	Instruction	Error of omission	PT type	Multiple initiating event	Existence of a failed indicator	Number of PTs per step	Note or caution
170314_#426	A	1	Goal: Check reactivity	False	Entering the succeeding step in the procedure	False	True	15	False
170314_#426	A	1	Is the reactivity criteria satisfied?	False	Directing information acquisition	False	True	15	False
170314_#426	A	1	Is the reactivity criteria satisfied?	True	Synthetically verifying information	False	True	15	False
170314_#426	B	2	Goal: Check electric supply	False	Entering the succeeding step in the procedure	False	True	8	False
170314_#426	B	2	Is the plant at low power?	False	Directing information acquisition	False	True	8	True

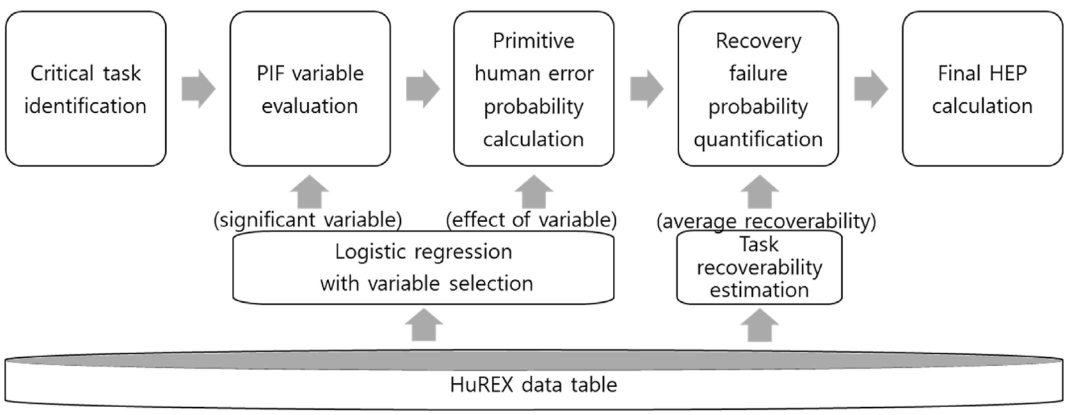

Fig. 2. Process of calculating HEP using HuREX data [30].

3.3.1. Critical task identification

One of the main features of HEP quantification based on EMBRACE [30] is that the HFE is decomposed into PTs involved in the instruction(s) of the procedures performed. In this sense, HEP is calculated by aggregating the failure probabilities of the PTs, which is termed primitive human error probability (PEP). Table 3 shows an example of identifying PTs from procedure instructions. For example, the instruction to verify whether a particular valve has been opened is generally performed with (1) a PT in recognizing the need for the valve verification action according to the procedure and requesting another operator to perform the action, and (2) a PT in verifying the condition of the valve according to the request and reporting the valve state. This distinction is plausible when operators in the main control room (MCR) of a nuclear power plant distinguish the required cognitive activities in accordance with a command-and-control protocol. For example, the shift supervisor of the APR1400 can direct other operators to observe the plant situation or operate machines according to relevant procedures. At this time, reactor or turbine operators report plant parameter values or manipulate components based on the instructions of the shift supervisor.

Because not all PTs in the procedure are related to a given HFE, it is necessary to identify a set of critical PTs. The critical PTs are those that satisfy any of the following criteria.

(1) The PTs must be carried out to satisfy the HFE's success criteria. Example PTs are equipment operation and communication.
(2) The PTs are the essential transitions to the procedural steps documenting the PTs that were already selected.
(3) The PTs are related to the checking of the plant conditions as essential to perform the PTs of (1) or (2) above.
(4) The PTs are entries to the procedural steps where other critical PTs such as (1) to (3) above are located.

In order to select the critical PTs, the critical procedural steps and the important instructions in the steps are identified in advance. Guidelines to consistently select the critical steps and important instructions were developed to enhance the transparency of task analysis; Fig. 3 schematically explains the basic principle to select a critical step and instruction regarding procedure transitions. In each case, the red node indicates a critical step that has been selected previously, and the blue node means a newly defined critical step. Indices in each node represent the step sequence number, and it is assumed that the steps in the procedure are sequentially followed in the order of $x-1$, x, and $x+1$. Steps y and z belong to Y and Z procedures, respectively. For example, cases 1–4 each show that when a procedural step marked with a red circle is already picked out as a critical step, the steps before that critical step are not declared as critical steps. This is because their success or failure does not affect the performance of the red step. The dotted arrow to step z in case 4 implies that step x has an instruction connecting to two different steps, y and z, but it is evident that the crew would not proceed to step z. In cases 5 and 6, it is possible that the crew will skip step y when they inappropriately carry out the tasks in step x; hence, step x is newly added as a critical step. The process of step identification is discussed in detail in Ref. [30].

3.3.2. PIF variable evaluation

To calculate the PEP of each critical PT, the PIF variables that affect the reliability of the critical PT are evaluated. The PIF variables evaluated here are the variables revealed as significant from the regression analysis of the HuREX data. For example, it was found

Table 3
Examples of PTs included in procedural instructions.

Procedure Instruction	PT
Verify valve A has opened.	RP – directing information acquisition IG – verifying the state of an indicator
Verify steam generator pressure is stable.	RP – directing information acquisition IG – evaluating a trend
Verify the level of tank B with reference to the attached graph.	RP – directing information acquisition IG – comparing in graph constraint
Check if residual heat removal pumps are operable.	RP – directing information acquisition IG – synthetically verifying information
Is the flow of line C lower than 44 l/s and the level of tank D lowering?	RP – directing information acquisition IG – comparing parameters RP – directing information acquisition IG – evaluating a trend
Start Pump D.	RP – directing manipulation (discrete) EX – simple (discrete) control
Maintain pressurizer level.	RP – directing manipulation (dynamic) EX – dynamic manipulation
If pump E has stopped, start pump F.	RP – directing information acquisition IG – verifying the state of an indicator RP – directing manipulation (discrete) EX – simple (discrete) control
(Entering a subsequent step)	RP – entering the succeeding step in the procedure
Go to PROC-01 procedure.	RP – directing procedure transfer

*EX: execution; IG: information gathering; RP: response planning.

through regression analysis that the presence of 'Urgent additional tasks' significantly affects the PEP of the information gathering PTs. Therefore, it should be evaluated whether the operators are urgently required to perform additional tasks during the process of performing a critical PT in the given HFE situation. Table 4 lists the significant PIF variables with respect to representative PT groups.

Table 4
Significant PIF variables.

PT group	Significant PIF variable
Information gathering PTs	Urgent additional task
Situation interpreting PTs	(No variable found)
Response planning PTs	PT type Change of procedure Negative plant state check Urgent additional task
Execution PTs	PT type Multiple initiating events Caution for manipulation Task attention Complexity of interface in local controller Training level of local operation

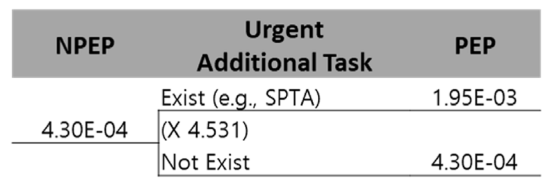

Fig. 4. Decision tree to calculate the PEP of information gathering PTs.

3.3.3. Primitive human error probability calculation

After evaluating all the variables related to the critical PTs, the PEP of each PT is calculated using a decision tree created based on the regression result. For example, the PEP of the information gathering PTs is determined based on the decision tree shown in Fig. 4 that considers whether or not there is an urgent additional task. Without the presence of an urgent additional task, the PEP is 4.30E-04. However, in the initial stages of an emergency where urgent additional tasks exist (e.g., identifying the overall plant condition and crucial components statuses), the PEP increases by about 4.5 times.

3.3.4. Recovery failure probability quantification

EMBRACE considers seven recovery sources: (1) self/peer review, (2) diagnosis re-evaluation, (3) apparent cue, (4) shift change, (5) recheck in stable status, (6) shift technical advisor (STA) monitoring, and (7) procedure. The recovery failure probability, which implies the probability of failure in the performance of a recovery action (i.e., 1 – recovery probability), is calculated for each recovery source. The HuREX data contains estimates regarding the failure probability of self/peer review recovery; since there are no empirical data for the other recovery sources, those failure probabilities of error recoveries were estimated based on the HRA

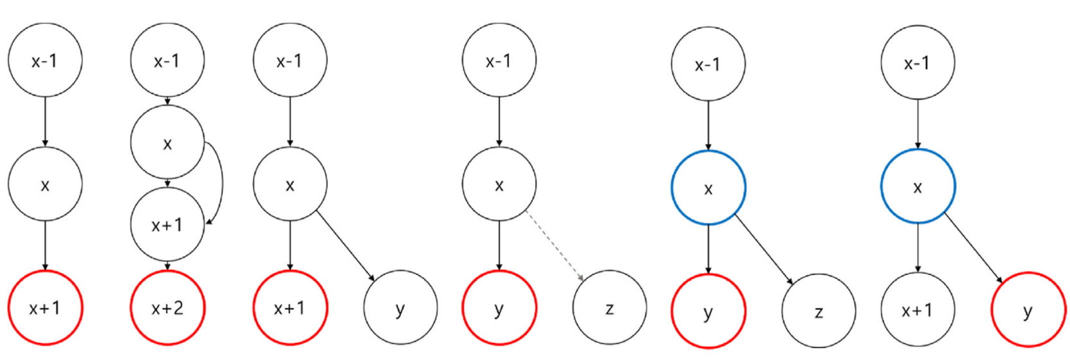

Fig. 3. Critical step selection cases.

dependency rule [31]. The meanings and evaluation methods of each source are as follows.

- Self/peer review: This means recovery activities based on the MCR instrumentation information or personal knowledge. According to HuREX data analysis, the recovery failure probability of each PT ranges from 0.5 to 0.9.
- Diagnosis re-evaluation: In the case of emergency diagnosis procedures, it is assumed that recoveries of the PTs in a diagnosis procedure are possible with a probability of 0.5 if there exists a step in the procedure that reviews the overall situation diagnosis results.
- Apparent cue: There are human events in which all operators can immediately confirm a task performance result without accessing the instrumentation or procedure. An example is that the operators know that the power supply has been lost simply by the obvious symptoms of the loss of regular lighting and the actuation of emergency lighting in the MCR. This kind of event is judged to be recoverable independently from the HFE. In other words, the recovery failure probability is estimated to be the total PEP of the event.
- Shift change recovery: If the HFE is a long-term action of more than 8 h, crew shift change is credited for the error recovery of the event. At this time, the errors by the existing crew can be recovered by the next crew with a recovery failure probability of 0.05.
- Recheck in stable status: If the HFE grants enough time to recheck the entire process after performing all the procedures responding to the accident, the potential error of the event is recovered with a recovery failure probability of 0.14.
- STA monitoring: Events in which the shift technical advisor can periodically monitor the crew's results have a recovery failure probability between 0.14 and 0.5.
- Procedure: If a separate procedural step instructs to verify the success criteria of the human event, it is assumed that recovery is possible with a recovery failure probability of 0.5.

Since the recovery failure probabilities regarding the self/peer reviews are estimated for each PT type from HuREX data, the recovery failure probability for this recovery source is calculated and applied to each PEP. In addition, because the diagnosis re-evaluation can only be considered in the emergency diagnosis procedures, this recovery failure probability is applicable to the PEPs regarding the diagnosis procedure. On the other hand, the recovery failure probabilities for the other kinds of recovery sources are integrated into a decision tree for application convenience. By multiplying the total PEP by the integrated recovery failure probability, the final HEP is quantified.

3.3.5. Final HEP calculation

The total HEP for a given HFE is calculated by the formula shown in Fig. 5, where $NPEP_i$ is the nominal primitive human error probability of the i-th critical PT, $PIFv_{ij}$ is the level of the j-th PIF variable for the i-th PT, SPR_i is the recovery failure probability of self/peer review, DRR_i is the recovery failure probability of diagnosis re-evaluation, and RFP is the recovery failure probability considering the other kinds of recovery sources. The detailed calculation process is discussed in Ref. [30].

4. Case study

This section demonstrates the use of SACADA and HuREX data to calculate the HEP of the feed-and-bleed (F&B) HFE in a loss of all feedwater scenario. The main purpose is to use the data available at the time of writing this paper to show the different HEP calculation processes discussed in Sections 2.3 and 3.3. The HEPs calculated in this section may change when more data become available. On account of the differences between NPPs, procedures, and MCRs, etc., the HEPs in this work should not be directly applied to other purposes without adequate analysis.

4.1. Case study: SACADA

The first part of this case study uses SACADA data collected from a Westinghouse four-loop pressurized water reactor with a conventional MCR simulator. The database used for this demonstration contains about 18,000 data points (i.e., 18,000 TOEs per crew) [16].

4.1.1. Event description

The initiating event is a loss of all feedwater scenario. The auxiliary feedwater system (AFW) is not available, leading the operators to conduct F&B to remove decay heat while trying to restore feedwater to the SGs. The F&B action is in the procedure FR-H1 ″Response to Loss of Secondary Heat Sink.” The heat sink critical safety function status tree (CSFST), which is constantly monitored by the shift technical advisor during emergency events, provides the FR-H1 entry conditions that include (1) no SG narrow range water level is greater than 14% when no adverse containment condition exists, and (2) the total AFW flow to the SGs is less than 576 gallons per minute. The heat sink CSFST is shown on a computer in the MCR. When a branch's entry condition is reached, the branch's color changes to alert the operators.

FR-H1 directs operators to check and restore feedwater to the SGs using the AFW system (e.g., turbine-driven and motor-drive pumps), main feedwater system, or condensate system. If all these methods fail, FR-H1 instructs operators to implement F&B to remove the heat from the RCS. In this scenario, operators need to trip all reactor coolant pumps (via FR-H1).

The available SACADA data have three F&B scenario variations. All scenarios start with an SG tube leak event and then progress to an SGTR event. In two scenarios, all the feedwater systems are not available that leads to F&B. The other scenario has additional complications before performing F&B, including a main steamline leakage, 250 DC ground fault, and turbine control failure. All three

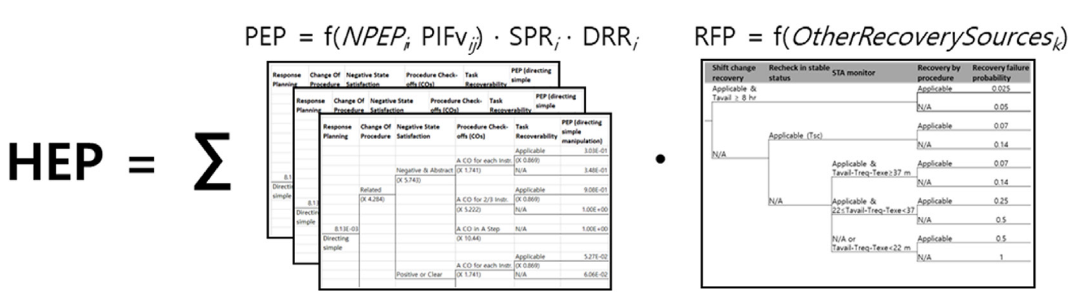

Fig. 5. Equation calculating HEP using estimates from the HuREX data.

scenarios involve multiple component and system failures. The following four TOEs directly relate to the F&B HFE of this discussion, including TOE descriptions and cognitive types as specified by the scenario developers.

- Monitoring Critical Safety Functions: The STA monitors the CSFSTs shown on a computer monitor in the control room to detect and notify the unit supervisor upon reaching the heat sink CSFST red path to enter the FR-H1 procedure. The cognitive type is detecting or monitoring information.
- Transitions to FR-H1: The shift supervisor announces the plant condition and enters the FR-H1 procedure. The cognitive type is deciding.
- Trip RCPs: The reactor operator trips the RCPs per the shift supervisor's command. The cognitive type is action.
- Initiate RCS F&B: The reactor operator performs the following items per the shift supervisor's commands: (1) actuate safety injection, (2) verify that the feed path alignment is correct, (3) establish the bleed path, (4) verify that the bleed opening is sufficient to depressurize the RCS enough for safety injection flow to enter the RCS, and (5) maintain the F&B until the safety injection reset criteria is reached. The cognitive type of this TOE is action.

The above four TOEs are specified from a training perspective to evaluate the performance of an individual operator or a sub-team. The first TOE evaluates whether the STA monitors the CSFSTs closely. The second evaluates the unit supervisor in implementing the FR-H1 procedure. The third and fourth TOEs evaluate the sub-team that consists of the reactor operator and unit supervisor in implementing the actions.

4.1.2. Task analysis and HEP quantification

This section applies the four HEP quantification steps discussed in Section 2.3 with the use of SACADA data to calculate the HEP of the F&B HFE. Three different analysis methods are discussed, namely, those by Azarm et al. [18,19], Nelson et al. [20,23], and Chang [16], which generate three different F&B HEPs that can act as an uncertainty range.

Step 1: Break down the HFE into TOEs

To calculate the F&B HEP, Azarm et al. [19] identified 11 TOEs while Nelson et al. [23] and Chang [16] identified 3 TOEs. The difference in the number of TOEs mainly stems from the differences in defining the scope of the F&B HFE. Azarm analyzed the scenario timing and concluded that the operator's primary procedure path is started with the reactor trip and safety injection procedure (EOP-0) because of reactor trip then followed by the reactor trip without safety injection procedure (ES-01) because SI is not needed. In ES-01, the operator starts to monitor the heat sink CSFST that leads to transferring to the loss of the secondary cooling procedure (FR-H1) because of the red path on the CSFST. The 11 TOEs could be identified from the employed procedures whose types of primary cognitive demands and description are the following:

1. Monitoring: Monitor indications to determine Reactor Trip
2. Monitoring: Monitor indications to determine Turbine Trip
3. Monitoring: Monitor indications to determine AC Power Available
4. Monitoring: Monitor indications to determine SI Status
5. Monitoring: Monitor indications to determine whether SI is required and to initiate Heat Sink CSFST
6. Monitoring: Monitor SG level and flow to determine the status of Heat Sink CSFST
7. Deciding: Transition to FR-H1(Loss of Secondary Heat Sink)
8. Manipulation: Actuate SI
9. Monitoring: Verify feed path
10. Manipulation: Establish bleed path
11. Monitoring: Verify Bleed path

Chang's and Nelson's 3 TOEs start from the monitoring of the CSFST and end at F&B completion, as described below.

1. Monitoring TOE: The STA monitors the CSFSTs and notifies the unit supervisor upon reaching the heat sink CSFST red path. When reaching a branch condition in the CSFST, the branch's color changes from green to red. Therefore, the signal is salient. The main challenge to human reliability is that the STA may divert their attention to other activities that could cause them to not detect that the FR-H1's entry condition has been reached.
2. Deciding TOE: The unit supervisor implements the FR-H1 procedure. All the collective decisions to provide an adequate heat sink are grouped together and represented by a TOE to calculate their joint HEP. These decisions include entering FR-H1, determining that none of the feedwater systems are available, dispatching plant staff to recover the feedwater systems, directing RCP trips and F&B actions, etc. This practice is because of the strong cognitive dependencies and many error recovery opportunities between the decisions; stated alternatively, from a human reliability assessment perspective, modeling each decision separately does not necessarily provide a more confident HEP estimate.
3. Action TOE: The RO trips all RCPs and implements the F&B actions. All the actions to achieve a successful F&B are grouped together and represented by a TOE.

Step 2: Specify TOE Contexts

Azarm's analysis [19] used the SACADA taxonomy to characterize the context of each of the identified 11 TOEs. Table 5 shows the contexts of 3 of the 11 TOEs.

Nelson et al. [23] and Chang [16] used three identical TOEs to calculate the F&B HFE's HEP. Nelson's analysis used the SACADA data points with the same context as the TOEs of analysis to calculate HEPs. The contexts are shown in the second column (Basic Context) in Table 6. Chang used the SACADA's data points with the same or similar contexts as the TOEs of analysis to calculate the HEP. As a result, Chang used more data points to calculate the HEP of each TOE than in Nelson's analysis. The contexts used in Chang's analysis are the combination of the second and third columns (i.e., basic context and expanded context, respectively) in Table 6.

Steps 3 and 4: Calculate the HEPs of the TOEs and F&B HFE

Azarm et al. [19] employ individual statistical significance tests to identify the TOEs in the SACADA database with similar contexts to the 11 TOEs identified for this analysis. Nelson et al. [20] apply the prior probability of 0.001 and the prior experience of 0.001 globally to all PIF combinations, then use HUGIN software [27] to import the SACADA data to calculate the posterior probabilities. Nelson's approach only uses the data points with contexts exactly identical to the contexts of the TOEs of analysis. The HEPs of the monitoring, deciding, and action TOEs are 3.3E-3, 5.3E-2, and 5.0E-3, respectively.

Chang's analysis [16] searched the SACADA data to identify the TOEs with the same or similar context as shown in Table 6, then calculated the HEPs by directly dividing the number of UNSAT and the number of TOEs via Microsoft Excel. The calculated HEPs of monitoring, deciding, and action TOEs are 1.3E-2 (2, 155), 3.4E-3 (4, 1169), and 2.4E-3 (1, 420), respectively, where the values inside the

Table 5
Example TOEs' context used in Azarm's analysis.

TOE	Context
Monitor SG level and flow to determine the status of Heat Sink CSFST	**Cognitive type:** Monitoring/Detection **Information source:** Computer **Detection mode:** Procedure directed monitoring **Signal:** Distinctive change **Workload:** Concurrent demands **Time criticality:** Extensive time available **Extend of communication required:** Extensive onsite communication **Other overarching issue**: Coordination
Transition to FR-H1	**Cognitive type:** Deciding **Decision basis:** Procedure **Familiarity:** Standard **Uncertainty:** Clear **Workload:** Concurrent demands **Time criticality:** Normal time available **Extend of communication required:** Normal communication
Establish bleed path	**Cognitive type:** Action **Action type:** Simple and distinct **Location:** Main or auxiliary control panel **Guidance:** Procedure **Recoverability:** Immediately recoverable **Workload:** Normal **Time criticality:** Extensive time available **Extend of communication required:** Normal communication

parentheses give the UNSAT number and the number of data points.

The F&B's HEPs calculated by the three different data analysis approaches are 3.9E-2 [19], 6.1E-2 [23], and 1.9E-2 [16]. The factors contributing to the HEP differences include the use of prior HEPs in the Bayesian update approach, different assumptions about the context resulting in differences in the data points used for analysis, and different definitions in the scope of the F&B.

4.2. Case study: HuREX

The HuREX data collected so far include reliability data extracted from both analog and digital instrumentation-and-control (I&C)-based MCRs. To concretely explain the proposed method, this case study restricts its focus on the data collected from an APR1400 simulator with a fully digitalized MCR to estimate the F&B HEP for an APR1400 PRA.

4.2.1. Event description

The event analyzed was a F&B operation during a loss of all feedwater accident. When a reactor trip occurred, the operators quickly recognize that the water levels of two steam generators (SGs) are less than a setpoint and lowering. These symptoms lead the operators to transfer to a loss of all feedwater (LOAF) procedure owing to an event-based diagnosis (ED) procedure that consists of a series of flowcharts to identify the nature of an accident in progress. When the operators enter the LOAF procedure, one of the procedural steps instructs them to check the operation of auxiliary feedwater (AFW) systems. At this point, since all of the AFW systems are not available, the operators have to confirm the integrity of the RCS heat removal function. If the SG level is low and the pilot valves are opened, operators supply electric power to the pressurizer safety valves using a power switch in a cabinet outside the MCR and manually open the safety valves. From timeline analysis, it was revealed that the temporal success criterion is 40 min, and the level of the SG lowers 20 min after the reactor trip.

The APR1400 procedures provide detailed instructions on handling this accident; Fig. 6 depicts the procedure flow to mitigate the event. The reactor trips soon after the loss of all feedwater, after which operators then enter the ED procedure immediately to

Table 6
Contexts used in Nelsons' analysis (basic context) and Chang's analysis (basic context plus expanded context).

TOE	Basic Context	Expanded Context (Basic Context plus the Following Specified Context)
Monitoring	**Cognitive type:** Monitoring/Detection	–
	Information source: Computer	All indicator-related sources, including meters, indication lights, flags, and computers
	Detection mode: Procedure directed monitoring	"knowledge-driven monitoring"
	Signal: Distinctive change	–
	Workload: Concurrent demands	"normal"
	Time criticality: Normal time available	"extensive time available"
	Extend of communication required: Extensive communication within the control room	"normal communication"
	Other factors: None	All “other factors"
Deciding	**Cognitive type:** Deciding	–
	Decision basis: Procedure	"skill"
	Familiarity: Standard	–
	Uncertainty: Clear	–
	Workload: Concurrent demands	"normal"
	Time criticality: Normal time available	"extensive time available"
	Extend of communication required: Extensive communication within the control room	"normal communication"
	Other factors: None	All “other factors"
Action	**Cognitive type:** Action	–
	Action type: Order (i.e., perform actions in sequence)	–
	Location: Main or auxiliary control panel	–
	Guidance: Procedure	–
	Recoverability: Recoverable with significant efforts	–
	Workload: Concurrent demands	"normal"
	Time criticality: Normal time available	"expensive time available"
	Extend of communication required: Extensive communication within the control room	"normal communication"
	Other factors: None	All “other factors"

identify the accident type and to transfer to the corresponding response procedure for that event, e.g., the LOAF procedure in this case. The 10th step of the LOAF procedure directs the operators to check the SG water levels to determine whether to transfer to the F&B procedure. In the F&B procedure, the 7th step instructs the supply of electric power to the safety valves, the 9th step directs the manual opening of the safety valves, and the 10th step provides an action list to review the results of the F&B operation.

4.2.2. Task analysis and HEP quantification

The HEP of the F&B is calculated based on the critical PTs identified from the procedural steps that contain the actions and procedure transitions critical to implement the F&B. The blue circles in Fig. 6 indicate the critical procedural steps. There are four critical steps. The 7th step and 9th step in the F&B procedure contain the component manipulation PTs relevant to the HFE's success criteria. The 5th step of the ED procedure and the 10th step of the LOAF procedure have the PTs required to transfer to the F&B procedure. The instructions in the critical procedural steps are described in Table 7. Among the instructions, some instructions were not thought to be significant to the procedure flow for the F&B operation (e.g., the 4th and 5th instructions in the 10th step of the LOAF procedure and the 2nd, 3rd, and 4th instructions in the 7th step of the F&B procedure); hence, those instructions were not analyzed. For example, evacuating workers in a containment building, which is described in the 7th step of the F&B procedure, is a very important action from a worker safety perspective. However, this action is not associated with the success criteria of the given HFE. Once the important instructions are selected, they are decomposed into the PTs defined by the HuREX framework. The last column of Table 7 shows the critical PTs.

For quantifying the PEPs of the PTs identified in Table 7, the states of the PIF variables are determined based on Table 4. As an example, Table 8 lists the critical PTs extracted from the 7th step of the F&B procedure with related PIF variables and states. Based on decision trees such as in Fig. 4, the HEP of the critical PTs (i.e., PEP), which is the product of a critical PT's nominal PEP and the PIF multipliers on the critical PT, is derived. The PEPs calculated for this event are shown in the last column of Table 8. In this case study, all recoveries from self or peer review activities are assumed to be possible. The total PEP of the critical PTs of the F&B HFE is 7.79E-02, which is the HEP without consideration of the other kinds of error recovery.

To calculate the recovery failure probability for the F&B HEP, the time available and required, the STA availability, and the procedural steps for recovery were analyzed. The time available is about 20 min; hence, it is infeasible for error recoveries by the mechanisms of shift change, recheck plant status, and STA monitoring. The apparent cue was also not available in this situation. However, there is a separate step in the related procedure to check the results of the F&B operation, giving a probability of error recovery of 0.5. Therefore, the final HEP of F&B was quantified as 3.89E-02 (= 7.79E-2 × 0.5).

5. Discussion and conclusion

5.1. HEP estimation using simulator data

In this paper, we presented different approaches to demonstrate the use of SACADA and HuREX data to estimate the HEPs of the F&B operation. The two systems acquire their data from simulator training sessions using different data taxonomies. Therefore, there are distinguishable differences in how to use the data to calculate HEPs.

SACADA relies on the scenario developers to define the TOEs' scopes; thus, the scopes are subject to the variability between scenario developers. This practice consequently affects HRA analysts in specifying the TOEs to calculate HEPs. One mechanism to reduce variability is by providing TOE examples from the SACADA database to scenario developers and HRA analysts. HuREX takes a different approach by identifying PTs explicitly according to the procedure instructions and providing comprehensive and detailed guidelines to identify the critical PTs from the procedures. In comparison, SACADA's TOE selection is more flexible but prone to analyst-to-analyst variability, while HuREX's critical PT selection is rigorous and, therefore, less subject to analyst-to-analyst variability.

Second, the demonstrated SACADA approach to calculate the F&B HEP directly analyzes the SACADA database by applying different analysis techniques. HuREX's way of calculating HEPs, on the other hand, does not analyze the HuREX data directly. Instead, HuREX data are pre-analyzed to generate decision trees for HEP calculation. The pre-generated decision trees include the nominal HEPs of the PTs and the effects of the PIFs on human reliability. To calculate the HEP of an HFE, the HRA analysts follow HuREX's rules to identify the critical PTs and the PIFs, and then apply the pre-generated decision trees to calculate the relevant HEPs. The two different approaches are driven by the ways that the SACADA and HuREX data are collected. SACADA data are entered by nuclear plant staff in simulator training, thereby the size of the database grows when simulator training is conducted. HuREX data collection relies on analysts examining the simulator sessions in the formats of videos and simulator logs, etc. The data collection approach maximizes the generation of data points for statistical analysis. With sufficient data, SACADA's data analysis approach can be more fit to calculate HFEs in specific contexts. The HuREX data analysis approach can generate HEPs with simulator data from a relatively small set of scenarios.

Finally, SACADA demonstrates three HEP calculation methods, while HuREX shows the use of PT error probabilities, PIF effects, and error recovery to calculate HEPs. Both systems have intentions to explore other data analysis techniques, perform additional

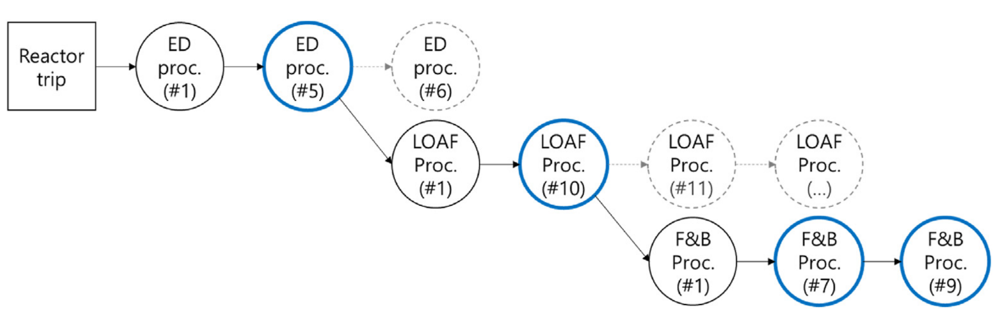

Fig. 6. Procedural path from accident initiation to F&B operation in the APR1400.

Table 7
Instructions of the critical steps in the F&B operation and their primitive tasks.

Procedure, step	Instruction index	Instruction	Primitive Task(s)
ED, 5	1	Goal: Check SG level	RP – entering the succeeding step in the procedure
	2	Is the SG level lower than 30%? (if yes, go to the LOAF procedure, otherwise, go to step #6 of the ED procedure)	RP – directing information acquisition IG – comparing parameters RP – transferring steps
LOAF, 10	1	Goal: Check heat removal from SG	RP – entering the succeeding step in the procedure
	2	Check if the SG level is lower than 1%; if unsatisfied, go to the contingency action part	RP – directing information acquisition IG – comparing parameters RP – transferring steps
	3	Go to the 1st step of F&B procedure	RP – transferring the procedure
	4	Check if the RCS cooling tube temperature is lowering	(Not important instruction)
	5	Go to the 1st step of F&B procedure if the temperature is raising quickly	(Not important instruction)
F&B,7	1	Goal: Ready the feed-and-bleed	RP – entering the succeeding step in the procedure
	2	Ready the feed and bleed by conduct the followings	(Not important instruction)
	3	Evacuate the workers in the containment	(Not important instruction)
	4	Verify the safety valves are operable	(Not important instruction)
	5	Supply power to the safety valves	RP – directing contact with outside the MCR EX – communicating with outside the MCR EX – local simple control
F&B,9	1	Goal: Initiate feed-and-bleed	RP – entering the succeeding step in the procedure
	2	If a leak alarm of the safety valves has activated and the discharge temperature is higher than 160°;	RP – directing information acquisition IG – verifying alarms RP – directing information acquisition IG – comparing parameters
	3	Then manually open the safety valves	RP – directing manipulation (discrete) EX – simple (discrete) control

Table 8
PIF variable evaluation and PEP estimates for the PTs in the 7th step of the F&B procedure.

Procedure, step	Instruction index	PTs	PIF variables and their states	PEP
F&B,7	0	RP – entering the succeeding step in the procedure	• PT type: Entering the succeeding step in the procedure • Urgent additional task: False	6.76E-04
F&B,7	1(3)	RP – directing contact with outside the MCR	• PT type: Directing contact with outside the MCR • Change of procedure: False • Negative plant state check: False • Urgent additional task: False	6.76E-04
		EX – communicating with outside the MCR	• PT type: Communicating with outside the MCR • Multiple IE: False • Caution for manipulation: Provided or Not required • - Task attention: Attentive	1.21E-02
		EX – local simple control	• PT type: Simple manipulation • Complexity of interface in local controller: Low • Training level of local operators: High	4.47E-03

analyses, and acquire additional data to advance the use of simulator data for HEP estimates and enhance the data basis of HRA.

5.2. Insights on the differences in the HEP estimates of the F&B HFE

The case study exemplified the data use processes for estimating the HEPs of the F&B HFEs. The HEPs quantified from HuREX and SACADA data had differences of up to a factor of 3.2. This section discusses the main contributors to the difference.

First, the plant types, the I&Cs, and the procedures in the MCRs where the case study was conducted were different between the HuREX and SACADA approaches. In this paper, the SACADA data were from a Westinghouse type reactor, which has an analog I&C MCR. Therefore, human events were analyzed assuming a situation that occurred in a similar type of power plant. For example, SACADA analyses assumed that the operators recognize the need for F&B from the monitoring of the CSFST. The HuREX data on the other hand were from an APR1400 plant, which has a fully digitalized I&C

MCR, and the data was used for estimating the F&B event in the APR1400. In this analysis, operators recognize the need for F&B from step 10 of the LOAF procedure.

Second, the contextual factors applied to the HEP estimations, such as time and man–machine interfaces, were not the same. For example, one of the SACADA analyses assumed that the factor of time available is normal or extensive. On the other hand, the HuREX analysis evaluated the scenario in which the time margin for the HFE was relatively shorter than other kinds of scenarios.

Third, the human actions that crews must perform were different. For example, the APR1400 required a local operator to supply power for the valve operation to implement F&B, while the Westinghouse plant did not need local action.

Because of the above contributors, it is difficult to explicitly compare the effectiveness of the two quantification approaches or the safety levels of the operating systems from the difference in the HEP values. However, this case study revealed some considerations in applying simulator data to estimate HEPs. First, it is important to define the scope of an HFE. In Chang and Nelson's analyses of SACADA data, they assumed the F&B event was detected by monitoring the CSFST. In the HuREX analysis, the initiation of the F&B event was by concluding the loss of all feedwater accidents in the ED procedure. The differences in cues affect the F&B scopes between the analyses, including the critical tasks for analysis; hence, it is necessary to establish guidelines for determining the scope of the human actions to be assessed.

Second, it is important to define the PT/TOE units or human errors in the HFE. The PTs in the HuREX systems are defined with more detailed tasks than the TOEs of SACADA. Detailed task definitions are useful to have clear specifications on the error mechanisms of the HEPs. However, there is also a risk that the quantification results may be too conservative unless the recovery possibilities of detailed tasks are properly considered.

Finally, after critical PTs or TOEs are identified for an HFE, the ways that the numeric information extracted from the simulator data can affect the HEP estimates should be considered. As mentioned earlier, the SACADA case study showed three different ways to calculate the HEP of the same TOE. The HuREX case utilized the results of regression analysis to calculate the HEP of each PT. It should be reasonable and trackable to select data related to a given TOE and TOE context and to calculate the HEP based on the data. Thus, it is beneficial to develop guidance on how to use the data in the future.

5.3. Additional issue

Several remaining research issues in the two quantification approaches presented in this study need to be discussed. The SACADA and HuREX data collected so far consist of data extracted from simulators of specific plants with limited data quantity. Therefore, it is necessary to verify whether the statistics generated from the collected data are sufficiently general. Additional data collection in various environments and plants is required, and for this, multinational and multi-disciplinary cooperation studies are necessary.

Although the two data collection systems include a significant number of PIFs, a strong data basis can only be provided to the extent that the data are available. For the contexts not covered by typical simulator training (e.g., extreme conditions such as large earthquakes and MCR fires or conditions demanding local actions), other data sources such as expert judgment and literature that provide relevant information would need to be researched to fill the gap. For example, as pointed out in Ref. [30], for considering the effect of some exceptional circumstances, the HEP quantified in this study can be additionally multiplied by the PIF effects proposed by existing HRA methods.

It is also important to understand and quantify the uncertainties hidden in the HEP estimates. The success/failure information in the SACADA and HuREX data can be used to predict the parameter uncertainties through Bayesian inference. For example, the HuREX data were analyzed to estimate the HEPs or PIF effects through Bayesian regression. In this situation, the predictive intervals of the regression model can be the evidence for estimating the parameter uncertainty of the HEP. However, it is necessary to understand the sources of uncertainty and to characterize the detailed uncertainty considering those sources. Basically, simulator data entails uncertainty issues such as (1) variability of the tasks within the category of PTs or macro-cognitive functions, (2) variability according to the operator's performance characteristics relevant to personal ability and operation/coordination style, and (3) the differences in operator performances from different scenarios. Greco et al. provide a good guide on the treatment of uncertainties according to these factors [32]. In addition, there is another kind of uncertainty factor due to limitations in the HFE modeling, such as a lack of coherence in the process of using the simulator data. As presented in this study, various HEP quantification methods are available, and it is beneficial to characterize the model uncertainty through comparative studies of these quantification methods. The establishment of a realistic and transparent quantification procedure is also expected to help improve the effectiveness of simulator data use.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by a grant from the Nuclear Research & Development Program of the National Research Foundation of Korea, funded by the Korean government, Ministry of Science, ICT & Future Planning (grant number 2017M2A8A4015291), and by the Nuclear Safety Research Program through the Korea Foundation Of Nuclear Safety (KoFONS), funded by the Nuclear Safety and Security Commission of the Republic of Korea (No. grant number 2101054).

References

[1] R. Patriarca, M. Ramos, N. Paltrinieri, S. Massaiu, F. Costantino, G. Di Gravio, R.L. Boring, Human reliability analysis: exploring the intellectual structure of a research field, Reliab. Eng. Syst. Saf. 203 (2020) 107102. https://www.sciencedirect.com/science/article/pii/S0951832020306037.

[2] Spurgin, A.J., P. Moieni, C.D. Gaddy, G.W. Parry, D.D. Orvis, J.P. Spurgin, V. Joksimovich, D.P. Gaver, and G.W. Hannaman, *Operator Reliability Experiments Using Power Plant Simulators Volume 1: Executive Summary*, 1990, EPRI NP-6937, Vol 1, Electric Power Research Institute, https://www.epri.com/research/products/NP-6937-V1.

[3] A.J. Spurgin, P. Moieni, C.D. Gaddy, G.W. Parry, D.D. Orvis, J.P. Spurgin, V. Joksimovich, D.P. Gaver, G.W. Hannaman, Appende\ixes, Operator Reliability Experiments Using Power Plant Simulators Volume 3, 3, Electric Power Research Institute, 1990. EPRI NP-6937, https://www.epri.com/research/products/NP-6937-V1.

[4] A.J. Spurgin, P. Moieni, C.D. Gaddy, G.W. Parry, D.D. Orvis, J.P. Spurgin, V. Joksimovich, D.P. Gaver, G.W. Hannaman, Technical Report, Operator Reliability Experiments Using Power Plant Simulators Volume 2, 2, Electric Power Research Institute, 1990. EPRI NP-6937, https://www.epri.com/research/products/NP-6937-V1.

[5] C. Bieder, P. Le-Bot, F. Cara, MERMOS, a second generation HRA method: what it does and doesn't, in: International Topical Meeting on Probabilistic Safety Assessment (PSA'99), 1999. Washington D.C.: American Nuclear Society.Aug 22 -26, 1999.

[6] C. Bieder, P. Le-Bot, E. Desmares, J.-L. Bonnet, F. Cara, Mermos : EDF's new advanced HRA method, in: A. Mosleh, R.A. Bari (Eds.), IAPSAM 4, Springer, New York City, 1998, pp. 129–134.

[7] Y.J. Chang, D. Bley, L. Criscione, B. Kirwan, A. Mosleh, T. Madary, R. Nowell, R. Richards, E.M. Roth, S. Sieben, A. Zoulis, The SACADA database for human

reliability and human performance, Reliab. Eng. Syst. Saf. 125 (2014) 117–133, https://doi.org/10.1016/j.ress.2013.07.014.
[8] W. Jung, J. Park, Y. Kim, S.Y. Choi, S. Kim, HuREX – a framework of HRA data collection from simulators in nuclear power plants, Reliab. Eng. Syst. Saf. 194 (2020) 106235. http://www.sciencedirect.com/science/article/pii/S0951832017309857.
[9] B. Kirwan, G. Basra, S.E. Taylor-Adams, CORE-DATA: a computerised human error database for human reliability support, in: Global Perpectives of Human Factors in Power Generation. Proceeding of the 1997 IEEE Sixth Conference on Human Factors and Power Plants. 1997, IEEE, 1997. http://ieeexplore.ieee.org/document/624858/.
[10] K. Jan, R. Dolezal, HRA data collection projects in Czech republic, in: HRA Data Workshop, 2020. Rockville Maryland, U.S. Nuclear Regulatory Commission.ADAMS Accession No. ML20083J308.
[11] A. Bye, E. Lois, V.N. Dang, G. Parry, J. Forester, S. Massaiu, R. Boring, P.Ø. Braarud, H. Broberg, J. Julius, I. Männistö, P. Nelson, NUREG/IA-0216, International HRA Empirical Study - Phase 2 Report: Results from Comparing HRA Method Predictions to Simulator Data from SGTR Scenarios, 2, 2011. ADAMS Accession No. ML11250A010), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1125/ML11250A010.pdf.
[12] V.N. Dang, J. Forester, R. Boring, H. Broberg, S. Massaiu, J. Julius, I. Männistö, H. Liao, P. Nelson, E. Lois, A. Bye, NUREG/IA-0216, International HRA Empirical Study - Phase 3 Report: Results from Comparing HRA Methods Predictions to HAMMLAB Simulator Data on LOFW Scenarios, 3, 2014. ADAMS Accession No. ML14358A254), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1435/ML14358A254.pdf.
[13] J. Forester, V.N. Dang, A. Bye, E. Lois, S. Massaiu, H. Broberg, P.Ø. Braarud, R. Boring, I. Männistö, H. Liao, J. Julius, G. Parry, P. Nelson, The International HRA Empirical Study: Lessons Learned from Comparing HRA Methods Predictions to HAMMLAB Simulator Data, 2014. NUREG-2127 (ADAMS Accession No. ML14227A197), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1422/ML14227A197.pdf.
[14] E. Lois, V.N. Dang, J. Forester, H. Broberg, S. Massaiu, M. Hildebrandt, P.Ø. Braarud, G. Parry, J. Julius, R. Boring, I. Männistö, A. Bye, NUREG/IA-0216, International HRA Empirical Study - Phase 1 Report: Description of Overall Approach and Pilot Phase Results from Comparing HRA Methods to Simulator Performance Data, 1, 2009. ADAMS Accession No. ML093380283), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML0933/ML093380283.pdf.
[15] J. Forester, H. Liao, V.N. Dang, A. Bye, E. Lois, M. Presley, J. Marble, R. Nowell, H. Broberg, M. Hildebrandt, B. Halbert, T. Morgan, The U.S. HRA empirical study – assessment of HRA method predictions against operating crew performance on a U.S. Nuclear power plant simulator, NUREG-2156 (ADAMS Accession No. ML16179A124), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1617/ML16179A124.pdf, 2016.
[16] Y.J. Chang, Preliminary analysis results of SACADA data for HEPs, in: HRA Data Workshop, ADAMS Accession No. ML20083J313, 2020. Rockville Maryland, March 12-13, 2020.
[17] M.A. Azarm, I.S. Kim, C. Marks, F. Azarm, Proposed methods for using SACADA DATA for empirical HRA estimation, in: HRA Data Workshop, the NRC Headquarters, March 15-16, 2018, 2018. Rockville, Maryland: U.S. Nuclear Regulatory Commission.ADAMS Accession No. ML18079A161.
[18] M.A. Azarm, I.S. Kim, C. Marks, F. Azarm, Analyses methods and pilot applications of SACADA data, in: Probabilistic Safety Assessment and Management (PSAM 14), UCLA Meyer & Renee Conference Center, Los Angeles, CA, 2018, in: http://www.iapsam.org/psam14/proceedings/paper/paper_60_1.pdf, in: http://www.iapsam.org/psam14/proceedings/Presentations/78.pdf.
[19] M.A. Azarm, I.S. Kim, C. Marks, F. Azarm, Data Analysis and a Pilot Application Using SACADA [Scenario Authoring, Characterization, and Debriefing Application], 2018. Innovative Engineering and Safety Solutions, LLC, ADAMS Access No. ML21012A293: U.S. Nuclear Regulatory Commission, http://www.iapsam.org/psam14/proceedings/paper/paper_60_1.pdf, http://www.iapsam.org/psam14/proceedings/Presentations/78.pdf.
[20] P. Nelson, C.R. Grantom, D. Quintanar-Gago, Methodology for supporting the determination of human error probabilities from simulator sourced data, in: Probabilistic Safety Assessment and Management (PSAM 14), UCLA Meyer & Renee Conference Center, Los Angeles, CA, 2018, in: http://www.iapsam.org/psam14/proceedings/paper/paper_78_1.pdf, in: http://www.iapsam.org/psam14/proceedings/Presentations/78.pdf.
[21] S. Fitch, P. Nelson, C.R. Grantom, A method to use SACADA data for estimating human error probabilities of human failure events, in: Human Reliability Analysis Data Workshop, March 15-16, 2018, ADAMS Accession No. ML18079A160, Rockville, Maryland, 2018. U.S. Nuclear Regulatory Commission.
[22] S. Fitch, P.F. Nelson, C.R. Grantom, A method to use SACADA data for estimating human error probabilities of human failure events, in: HRA Data Workshop, the NRC Headquarters, March 15-16, 2018, 2018. Rockville, Maryland.ADAMS Accession No. ML18079A160.
[23] S.A. Fitch, P. Nelson, C.R. Grantom, A Methodology to Use SACADA Data for Estimating Human Error Probabilities of Human Failure Events, Trinitek Services, Inc., 2018. ADAMS Accession No. ML21012A294: U.S. Nuclear Regulatory Commission, http://www.iapsam.org/psam14/proceedings/paper/paper_78_1.pdf, http://www.iapsam.org/psam14/proceedings/Presentations/78.pdf.
[24] K.M. Groth, J.D. Tedeschi, R. Moradi, Transforming HRA using SACADA data, Bayesian methods, and DBNs, in: HRA Data Worksop, the NRC Headquarters, March 15-16, 2018, 2018. Rockville, Maryland.ADAMS Accession No. ML18079A162.
[25] K.M. Groth, R. Smith, R. Moradi, Exploration of Methods for Using SACADA Data to Estimate HEPs: Final Report, Center for Risk and Reliability, University of Maryland, 2018. ADAMS Accession No. ML21012A295: U.S. Nuclear Regulatory Commission.
[26] K.M. Groth, R. Smith, R. Moradi, A hybrid algorithm for developing third generation HRA methods using simulator data, causal models, and cognitive science, Reliab. Eng. Syst. Saf. 191 (2019) 106507. http://www.sciencedirect.com/science/article/pii/S0951832018312456.
[27] M. Neil, N. Fenton, L. Nielsen, Building large-scale Bayesian networks, Knowl. Eng. Rev. 15 (3) (2000) 257–284.
[28] NORSYS Software Corp, Counting-Learning Algorithm, 2020. https://www.norsys.com/WebHelp/NETICA/X_Counting_Learning_Algorithm.htm.
[29] T.M. Sousa, T. Pinto, I. Praca, Z. Vale, H. Morais, Reinforcement learning based on the bayesian theorem for electricity markets decision support, in: S. Ōmatu, et al. (Eds.), Distributed Computing and Artificial Intelligence, 11th International Conference, Springer, New York, 2014.
[30] Y. Kim, J. Kim, J. Park, S. Choi, H. Kim, An HRA Method for Digital Main Control Rooms -Part II: Estimating the Failure Probability Due to Cognitive Error, 2020.
[31] J. Kim, S.J. Lee, S. Jang, HuRECA: the Human Reliability Analysis Method for Computer-Based Advanced Control Rooms, Korea Atomic Energy Research Institute, 2011. KAERI/TR-4385.
[32] S.F. Greco, L. Podofillini, V.N. Dang, A Bayesian model to treat within-category and crew-to-crew variability in simulator data for Human Reliability Analysis, Reliab. Eng. Syst. Saf. 206 (2021) 107309.