Reliability Engineering and System Safety 148 (2016) 44–56

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Human error probabilities from operational experience of German nuclear power plants, Part II

Wolfgang Preischl [a], Mario Hellmich [b,*]

[a] Gesellschaft für Anlagen- und Reaktorsicherheit (GRS) mbh, Forschungszentrum, Boltzmannstraße 14, 85748 Garching, Germany
[b] Bundesamt für Strahlenschutz (Federal Office for Radiation Protection), Willy-Brandt-Straße 5, 38226 Salzgitter, Germany

ARTICLE INFO

Article history:
Received 13 August 2015
Received in revised form
3 November 2015
Accepted 7 November 2015
Available online 27 November 2015

Keywords:
Human reliability assessment
Human reliability data
Probabilistic risk assessment
Operational experience

ABSTRACT

This paper is a continuation of an earlier publication (Preischl et al., Reliab Eng Syst Saf 2013;109:150–9) and presents the second part of a project aimed to collect human reliability data from the operational experience of German nuclear power plants. We employ a method which utilizes the German licensee event reporting system to gather the data. In this way, in addition to the data already presented in the previous paper, another 30 estimates for human error probabilities (HEP) are obtained. Moreover, a new method to access parts of the operational experience below the notification threshold of the German event reporting system is described. This method is demonstrated in cooperation with a reference nuclear power plant, resulting in 18 additional HEP estimates. As a result of both projects altogether 74 usable HEP estimates for a wide variety of tasks were derived. Notably, a number of them concern memory related or cognitive errors. A comparison with the THERP database shows that for 48 of these HEP estimates THERP provides no data, whereas in the 26 cases where THERP proposes a HEP it agrees with our data in all but eight cases.

1. Introduction

This paper is a continuation of an earlier publication [25]. It presents the results of the second part of a project aimed to infer human reliability data from operational experience of German nuclear power plants. In this part of the project [24], as well as in its predecessor [23], the German licensee event report system is utilized to gather human reliability data, in order to both validate and extend existing databases. Both projects were funded by the German federal nuclear regulator, the Ministry for the Environment, Nature Conservation, Building and Nuclear Safety, and conducted by Gesellschaft für Anlagen- und Reaktorsicherheit (GRS) under the technical attendance of Bundesamt für Strahlenschutz (BfS). As in [25], we apply a specifically designed method to infer human error probabilities from licensee event reports corresponding to events which have certain properties. In the present paper, besides supplying another 24 human error estimates for a wide variety of tasks, we describe a new method to access parts of the operational experience below the notification threshold, i.e. for tasks in which an error would lead to a reportable event, but which have not yet produced an event. This method is demonstrated in cooperation with a reference nuclear power plant, and 18 human error estimates obtained in this way by zero failure estimation are presented.

The scarcity of validated and traceable human reliability data is frequently quoted as a major problem in human reliability analysis (HRA), for both the development of new HRA methodologies and for applications in the context of probabilistic risk assessment (PRA). See [22] for a thorough recent literature review concerning the HRA data problem, and [31,2] for a historical perspective. Due to the shortage of relevant human reliability data, several data collection efforts have taken place in the past, where in some cases “HRA data” is understood as any relevant operational experience and is not confined to just human error probabilities for specific tasks; moreover, some databases collect data from various sources and not only from operational experience of nuclear power plants. Note that in most cases the data is not publicly available. General references concerning the problem of HRA data collection and database requirements are [31,29,16,17,14]. Some recent examples of data collection efforts in the nuclear sector are the SACADA database by the US NRC [4], which is intended as a long term data collection program, the Computerized Operator Reliability and Error Database (CORE-DATA) [11,12], supported by the UK Health and Safety Executive, which is used in conjunction with the NARA HRA methodology [19] in the UK, the older Nuclear Computerized Library for Assessing Reactor Reliability (NUCLARR) by the US NRC [10], which is reported to contain more than 2000 task based

* Corresponding author.
E-mail addresses: wolfgang.preischl@grs.de (W. Preischl), mhellmich@bfs.de (M. Hellmich).

http://dx.doi.org/10.1016/j.ress.2015.11.011
0951-8320/

human error estimates [3], and the Operator Performance and Reliability Analysis Database (OPERA) developed by the Korean Atomic Energy Research Institute (KAERI) [20], which includes data from real operational experience as well as simulator data. See [22] for further examples and discussion. Moreover, many HRA methods propose their own data, such as THERP [30], with its database consisting of over 100 human error probabilities (HEP). In spite of these data collection efforts and the data already available the importance of inferring data from actual plant experience remains, as has been repeatedly pointed out in the literature [8,29,18,31].

Since the German guideline for PRA in the context of periodic safety reviews of nuclear power plants recommends THERP as the primary method to be employed in the HRA part [5], the main goal of the data collection projects reported in the present paper and in [25] is to check whether the THERP database is in accordance with the operational practice of German nuclear power plants, and moreover, to extend it by HEP estimates for tasks for which THERP provides no data. Nevertheless, the data we obtain is not specific to any HRA method.

Summarizing the results of the two research projects [23,24], we contribute 67 HEP estimates from samples that were generated using actual operational experience (however, in ten of them the sample size is considered to be too small for precise data generation, and in one case the sample is probabilistically trivial). Moreover, another 18 samples from operational experience below the notification threshold are reported, resulting in a total of 74 usable HEP estimates. The samples cover a wide variety of tasks; notably, a number of them are memory related or involve cognitive errors, for which THERP does not provide data. For altogether 48 samples THERP provides no data, and in those 26 cases in which THERP proposes a HEP estimate, it agrees with our data within the uncertainty bounds in all but eight cases (in most of them the disagreement is slight, and THERP deviates in the conservative direction, see also the remarks on the comparison in Section 4). Hence we conclude that the THERP database is generally in good agreement with the German operational practice, at least for those tasks for which samples are available.

This paper is organized as follows. Section 2 briefly recalls the statistical method that is used to estimate human error probabilities from the sample data (the number of errors and the number of opportunities for errors). Moreover, uncertainty analysis of the data is addressed, which is not only of theoretical interest but also relevant for application of the data in PRA studies. In Section 3 our data source, the German licensee event report system, is introduced: Section 3.1 briefly recalls the method that is employed to gather HRA data from reportable events and, in particular, how it is possible to infer the number of opportunities for errors (the so-called "denominator problem" [22]) for certain tasks, whereas Section 3.2 details the method that was developed to access parts of the operational experience below the notification threshold of the event report system. Section 4 presents the altogether 48 samples obtained in the present project in the form of data tables. A detailed discussion and interpretation of the results of the present project, taking into account also our earlier results [25], is provided in Section 5. The paper closes with concluding remarks in Section 6.

2. Statistical inference of human error probabilities and uncertainty analysis

In this section we briefly describe the method of statistical inference used to analyze the samples and to infer human error probability (HEP) estimates from the sample data. A more detailed description was given in [25].

Recall that the HEP is the probability that "when a given task is performed, an error will occur" [30]. Thus, according to the relative frequency interpretation of probability, for a particular task labeled by i the corresponding HEP can be estimated by

$$\mathrm{HEP}_i \approx \frac{m_i}{n_i}, \tag{1}$$

where n_i is the number of times the task i was performed, and m_i is the number of errors that occurred. Clearly, the parameter HEP_i always lies in the interval $[0,1]$. A more powerful inference method than simply taking (1) as an estimator for HEP_i is Bayesian analysis [7], which, for the purpose of the present application, is now briefly recalled.

As with all input data for PRA, uncertainty analysis should also be done for human reliability data and the corresponding uncertainty should be appropriately propagated through the PRA model. To perform uncertainty analysis for the HEP estimates, two conceptually different uncertainty contributions can be distinguished. First, according to the relative frequency interpretation it is an underlying assumption that every individual has a certain error probability at a particular time, given a particular task to be performed under given circumstances [6]. Thus, there is a variability of HEP_i both with the individual selected from the population of shift personnel as well as due to the individual's temporal variability of its fitness for duty (e.g. during night shift), or due to the variability of the boundary conditions under which the task is performed. This variability can be modeled by considering HEP_i as a random variable with a distribution concentrated on the interval $[0,1]$. In the context of PRA, this variability becomes manifest as an uncertainty about HEP_i since it cannot be known in advance which individual of the population is in charge of the task to be performed at the (random) time the PRA initiating event occurs. It was argued in [25] that a beta distribution (with suitably chosen parameters) should be used to describe this variability, since this distribution is unimodal (with only a single maximum, reflecting the fact that among the population of operators one usually cannot find two or more subgroups with grossly different performance levels) and in general unsymmetrical, as well as being concentrated on the interval $[0,1]$. Recall that the beta distribution depends on two parameters $\alpha, \beta > 0$; it is absolutely continuous and is given by the density

$$f_{\alpha,\beta}(x) = \begin{cases} \dfrac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & x \in]0,1[\\ 0 & \text{else} \end{cases}, \tag{2}$$

where $B(\alpha,\beta)$ is the beta function. The graph of (2) has a bell-shaped appearance for parameter ranges of interest in the present application.

There is another contribution to the uncertainty of HEP_i, attributable to our limited knowledge about the "system" under consideration (the operator in the socio-technical context of the power plant), due to the limited amount of data on which the estimation of HEP_i is based: the so-called epistemic uncertainty. According to the Bayesian interpretation of probability as a "degree of belief" this uncertainty can also be modeled by a probability distribution, which we choose from the beta family as well. Consequently, samples of a smaller size n_i with a larger epistemic uncertainty tend to have a beta distribution with a greater variance (i.e. a wider bell-shaped density curve), corresponding to larger uncertainty intervals for the single point HEP estimates.

We remark that, in the present paper and in [25], due to the way the samples are taken they reflect the performance of a group of operators in charge of a specific activity and summarize the performance differences of the individuals in the group. However, the HEP variability due to individual differences or randomly