Reliability Engineering and System Safety 109 (2013) 150–159

Contents lists available at SciVerse ScienceDirect

Reliability Engineering and System Safety

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

Human error probabilities from operational experience of German nuclear power plants

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 9 March 2012
Received in revised form
13 August 2012
Accepted 16 August 2012
Available online 25 August 2012

Keywords:
Human reliability analysis
Human reliability data
Probabilistic safety assessment

ABSTRACT

We report about a project aimed to collect human reliability data from the operational experience of German nuclear power plants. Its objective is the validation and extension of existing human reliability databases (in particular, the THERP database). A method utilizing the German licensee event report system to gather the data is described. For certain tasks with specific attributes this method allows to determine the number of times the task was performed in the past, as well as the number of errors that occurred. A statistical method to estimate the corresponding human error probability (HEP) based on these numbers is provided. We have applied this method to the reportable events stored in the database collecting the reportable events in German nuclear installations. In this way up to now 37 HEPs for a wide variety of tasks were obtained, together with information about relevant performance shaping factors. We discuss these HEP estimates and compare them to the THERP database if it provides a HEP for the task in question. In all except three cases we find an agreement within the uncertainty bounds. Moreover, we contribute 21 HEP estimates for which the THERP handbook provides no data, so they serve to extend the THERP database, among them a number of memory related errors. Therefore, this data may serve as an input for the discussion of second generation HRA methods.

1. Introduction

Probabilistic risk analyses (PRA) are an important tool in assessing the safety of nuclear power plants, and they play a prominent role in the regulatory process. Beyond the nuclear sector, safety assessments for high risk ventures in other industries, for example chemical or offshore, rely increasingly on PRAs. A human reliability analysis (HRA) constitutes a vital part of any comprehensive PRA of a sufficiently complex system like a nuclear power or chemical plant. For this purpose a number of HRA methods are available. Since the early 1990's the development of new HRA methods has become an active field of research; in particular, the so-called second generation methods [11] (see also [24,25] for an overview) propose a new paradigm in HRA. Moreover, various aspects of first generation methods have become the subject of recent research, such as the dependence of human failure events and performance shaping factors [8,9]. However, despite the recent research efforts, the widespread use of HRA and its importance in the PRA field [19,7], a major problem remains the lack of plant specific, or at least industry specific, human reliability data, which is at the basis of any attempt to quantify human reliability. Hence it is important to gather human reliability data from operational experience, to both validate and extend the existing databases. For this reason the German federal nuclear regulator, the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, sponsored a research project to gather human reliability data from the operational experience of German nuclear power plants. This project was conducted by Gesellschaft für Anlagen- und Reaktorsicherheit (GRS) under the technical attendance of Bundesamt für Strahlenschutz (BfS). In the present paper our purpose is to report about this project and present the human reliability data that were established, together with the methods that were employed. The objective of this project is, on the one hand, to validate existing HRA data (in particular, the data contained in the THERP handbook [28]) for the use in PRAs of German nuclear power plants, as well as to extend the available data. It constitutes the first systematic attempt to generate human reliability data from the operational experience of German nuclear power plants.

The human reliability data currently available for PRA use stems from three major resources: (1) Databases associated with a particular HRA method; here the method THERP [28] with its database consisting of 27 tables with more than 100 human error probabilities is the most voluminous and best known, but other HRA methods propose their own human error probabilities.

* Corresponding author. Tel.: +49 30 18333 1033.
E-mail addresses: wolfgang.preischl@grs.de (W. Preischl), mhellmich@bfs.de (M. Hellmich).

0951-8320/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ress.2012.08.004

(2) Stand-alone databases independent of a particular HRA method, such as [12,15], which aim to record operational experience for data generation purposes. (3) Simulator databases, which cumulate data collected in simulator studies, refer to [26,14] for an overview. For an excellent account of the historical perspective on the collection of human reliability data refer to [29] and the references therein, as well as to [26,14].

The importance of collecting human reliability data from actual plant experience has been pointed out by many authors [27,10,20,29], because the validity of the presently existing data is often difficult to assess due to the following issues [26]:

- It is often unclear how the data was obtained.
- It is often unclear if the data is relevant to the current application.
- The accuracy of the data is unknown and it is impossible to track it to its sources.
- The data is not peer reviewed or publicly available.

Moreover, the amount of data available is still insufficient to cover all needs.

The German guideline for probabilistic risk assessments in the context of periodic safety reviews [2] (which is currently under revision) recommends to use THERP [28] for the HRA part of a PRA. Only in a few cases an equivalent HRA method can be used. Consequently our primary purpose 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 a number of human error estimates for which it contributes no data.

However, the data we obtain is not specific to any particular HRA method. It is our goal to collect and analyze the data in a transparent, comprehensible and traceable way, and to put our results and methods up for discussion by the HRA community. In this way we hope to meet the above issues head on, and thereby contribute to the HRA data problem solution.

The paper is organized as follows. In Section 2 we start with a detailed description of the statistical method which we use to obtain human error estimates together with the corresponding uncertainty bounds on the basis of counting data (i.e. the number of errors and the number of opportunities for errors). We also provide some discussion about uncertainty of human reliability data. The next Section 3 introduces our data source, the German licensee event report system. Since this system is not intended to gather human reliability data, some problems have to be solved in order to use it for this purpose. We report on our method to address these problems and how it was applied in the present project. Section 4 presents the 37 human error estimates in the form of data tables. Section 5 discusses these results and compares them, if appropriate, to the THERP data. The final Section 6 provides conclusions and an outlook on further studies planned.

2. Human error probabilities and statistical inference

2.1. Human error probabilities

The human error probability (HEP) is the basic parameter describing human performance. According to [28], "the HEP is the probability that when a given task is performed, an error will occur". Thus, according to the relative frequency interpretation the HEP for a particular task labeled by i can be approximated 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 the number errors that occurred. Of course the parameter HEP_i always lies in the interval [0,1]. Below in Section 2.2 we will provide a rigorous statistical method to estimate the left hand side of (1) based on observed numbers n_i and m_i.

It is an underlying assumption of the relative frequency characterization of HEPs that every individual has a certain error probability at a particular time, given a particular task to be performed under particular circumstances [4]. Thus, if an individual is randomly selected from a population, the probability for making an error in performing a certain task under given conditions at a given point of time depends on the individual's error probability, and thus becomes uncertain. This uncertainty about HEP_i due to the individual being randomly chosen from a population at a random time can be modeled by considering HEP_i as a random variable with a distribution concentrated on the interval [0,1]. In other words, this uncertainty reflects the fact that it cannot be known in advance which individual of the population is in charge of the task to be evaluated by the human reliability analysis when the PRA initiating event occurs.

This type of uncertainty should, at least conceptually, be distinguished from the *epistemic uncertainty* about HEP_i, which is due to our incomplete knowledge about the system and the limited amount of data on which our inference of the true value of HEP_i is based [17,18,21]. According to the Bayesian interpretation of probability as a measure of the "degree of belief" this uncertainty can be modeled by a probability distribution as well.

Each of our samples taken from operational experience reflects the performance of a population of operators in charge of a specific activity. It summarizes performance differences of the individuals of the population (e.g. due to fitness for duty differences), which all have an impact on sample parameter m_i. With an increasing sample size n_i it can be expected that every individual in the population is contributing to m_i, thus with our statistical method to be described below we obtain a HEP estimate which is an average of the performance of the population. The epistemic uncertainty about this group performance can be expected to tend to zero for the sample size tending to infinity.

For PRA purposes HEP estimates for groups of operators (e.g. shift supervisors) acting in the context of given boundary conditions (e.g. experience, stress level, ergonomic layout) are needed. As explained above our HEP estimates accomplish that requirement. However, in order to perform uncertainty analysis it is necessary to take into account the HEP variability due to the impact of individual differences (deviation of individuals from group average). The error probability of a randomly selected task performer acting in the context of randomly selected boundary conditions (e.g. fitness for duty during night shift) is not known and not described by the samples. This uncertainty source has still to be assessed by expert judgment and requires a modification of the uncertainty bounds of the HEPs derived by the statistical method described in Section 2.2.

Let us write θ_i for the random variable describing the HEP for performing a certain task i. To account for the uncertainty about θ_i, Swain and Guttmann [28] suggest a log-normal distribution (i.e. $\log \theta_i$ is assumed to be normally distributed). Since the log-normal distribution is not concentrated on the interval [0,1] it has to be truncated (and appropriately renormalized). Swain and Guttmann choose this distribution on the grounds that since "the performance of skilled persons tends to bunch towards the low HEPs [...] it is appropriate, for PRA purposes, to select a nonsymmetric distribution". Moreover, an unimodal distribution which