ZEDB DATA BASE – USING RAW RELIABILITY DATA FOR MAINTENANCE PLANNING

Günter Becker[1], Alexios Camarinopoulos[1], Yousef Abusharkh[2], Mariana Jockenhövel-Barttfeld[2], Heiko Kollasko[2]

[1] *RISA Sicherheitsanalysen GmbH, Krumme Str. 55, D-10627 Berlin, risa@risa.de*
[2] *AREVA NP GmbH, Henry-Dunant-Str. 50, D-91058 Erlangen, deinfo.areva.@areva.com*

The Centralized Reliability and Events Data Base (ZEDB) is a long-term data collection project for reliability measures of nuclear power plants in Germany and other European countries, which started more than twenty years ago with the goal of obtaining trustworthy reliability data for use in probabilistic safety analyses (PSA).

The final evaluation of reliability data considers the operation experience of power plants until December 2013 and has been created and published in August 2015. This paper summarizes the ZEDB in terms of data base organization, data contents, properties and capabilities of the data base software, data collection, results generated and the quality assurance approach applied. Finally possible future use of ZEDB is discussed.

I. DATA STRUCTURE AND OPERATING EXPERIENCE OF THE ZEDB

The Centralized Reliability and Events Data Base (ZEDB) gathers and analyses operating experience of safety relevant components gained at 19 nuclear power plants in Germany and 2 plants in the Netherlands and Switzerland [1].

The next chapter presents the data structure of the ZEDB database. Afterwards, data collection, evaluation, contents so far and possibilities of subsequent usage are discussed.

I.A. Data Structure

The smallest unit examined in the ZEDB database is the component [2], which is described by means of a plant-specific ID code (place of installation). The components are characterized by the fact that they perform an independent function within a mechanical or electrical system. Each component type is defined by a component boundary.

For each component contained in the database, component master data, operating reports and events reports are stored. These reports are essential for the organization of ZEDB data content and are defined as follows:

I.A.1. ***Component Master Data***

The technical attributes of the components, like design and operating parameters, are recorded in the master data. These technical attributes, which are called “aspects” in the ZEDB database, represent the master key to create homogeneous populations. The format of the master data reports is identical for all component prototypes. Most of the attributes are mandatory aspects for the database and are provided by the utilities.

The inspection intervals and observation periods for the component are also specified in the master data, which identify the component and the plants more precisely.

Inspections are considered as component tests to check if a stand-by component will function when required (this concerns failure on demand and stand-by failure rate).

I.A.2. ***Operation Reports***

In the operating reports stating the start and end times of the reporting period are stored for each component. The operating time, any disregarded periods within the reporting period, and the frequency of demands are also stated.

Start time and end times determine the length of the observation period for the component in the report. The component standby time is then obtained by subtracting the stated disregarded periods and the stated operating time from this time interval.

I.A.3. ***Event Reports***

All events leading to unavailability of a component as a consequence of a failure or malfunction of the component itself are reported in the ZEDB database as a component failure event. The failure events to be considered in the evaluation of the reliability data must occur within the component boundary of a component prototype, and must be capable of being unequivocally allocated to a specific component prototype. The format of the event reports is virtually identical for all component prototypes. The contents of the dropdown lists for the aspects “failure mode” and “subassembly/component part” differ.

I.A.6. ***Queries and analysis options***

As the data contained in the ZEDB database are assigned to their respective plants and components by means of the plant abbreviation and component ID, it is possible to perform database queries to select components and events according to certain criteria and to create component populations for which the reliability parameters can be determined. For the evaluation of reliability data, two steps are required:

1. The selection of the component populations according to certain technical and/or operational attributes using the component master data.
2. The relevant failure events for these component populations are selected from the event reports by choosing certain aspects, e.g. failure mode “fails to open”.

Using the associated operating reports of the selected components and failure events, the following reliability data can be determined:

- Failure rates per time [1/h] for periods of operation,
- Failure rates per time [1/h] for periods on standby, and
- Failure probabilities per demand.

These parameters are determined by the hierarchical super-population approach, which enables the evaluation of both generic and plant-specific values [3 – 9]. The generic reliability values are obtained using evidence from all participating plants. To obtain the plant-specific values, the data from all plants, with the exception of the particular plant concerned, are firstly used to determine a prior distribution, which is then updated in a second step to obtain a posterior distribution by including the data from the plant in question.

I.B. Example of an evaluation

Figure 1 below presents an example of a ZEDB result sheet. The component considered is a generator transformer with the failure mode “no transmission of power”. Nine failures of these components have been observed in a total observation period of $4.8317 \cdot 10^{+06}$ hours. Assuming a homogenous stratum, this would lead to the uncertainty distribution in TABLE **1** below, which is compared with the hierarchical (2 stage Bayesian) approach used in ZEDB [10].

TABLE 1: Comparison of results

Type	5 % Quantile	Median	95 % Quantile	Exp. value	K
Homogenous	$1.05 \cdot 10^{-06}$	$1.90 \cdot 10^{-06}$	$3.12 \cdot 10^{-06}$	$1.97 \cdot 10^{-06}$	1.73
Hierarchical	$3.36 \cdot 10^{-07}$	$1.45 \cdot 10^{-06}$	$5.74 \cdot 10^{-06}$	$2.17 \cdot 10^{-06}$	4.13

Concerning the expected value results appear not to differ too much, however, the uncertainty ranges differ considerably.

However, two out of 15 plants in **Figure 1** below are outside the uncertainty region of the simple homogeneous calculation, where 0 to 1 such outliers at most are expected. The hierarchical approach used in ZEDB provides a more realistic result, as the uncertainty distribution is broader.

I.C. Data Contents

I.C.1. ***Queries and analysis options***

The total operating experience collected in the ZEDB until December 2013 corresponds to approximately 336 thousand years.

There are approx. 18200 components in the ZEDB. TABLE 2 summarizes the number of components for the different prototypes. The distribution over the prototypes is presented in Figure 2

The continuous growth of the operating experience for all thirteen prototypes included in the database for the period 1998-2013 is illustrated graphically in Figure 2. The operating experience's growth rate for the presented prototypes reflects the number of components in the respective prototype, which is given in TABLE 1. Furthermore, the diagram shows the stability of the growth rate of each prototype over time, which confirms the continuous gapless data delivered by the participating power plants. Due to the shutdown of approx. 40% of the German Nuclear Power Plants, a decrease of the operating experience growth rate can be observed after 2010.

TABLE 2: Number of components contained in the ZEDB database distributed over different prototypes

Prototype	Number of components
Valves	9763
Circuit breakers	2264
Pumps	1219
Control rods	929
Bus bars	875
Vessels/Tanks	736
Static inverters	575
Fans	496
Batteries	457
Heat exchangers	352
Transformers	326
Emergency diesel generators	144
Rotating inverters	83
Total Number of Components	18219

Transformers VGB ZEDB

Components: Generator transformer
Type of transformer: Two-winding
Primary voltage: 220 – 425 kV

Failure mode: **Failure to transmit power**

Analysis: Plant-specific, failure rate (operating time) [1/h]

K.1-2: Generator Transformer, Plant-Specific Failure Rate

Plant	5 % Quantile	50 % Quantile	95 % Quantile	k	Mean Value
1	-	-	-	-	-
2	-	-	-	-	-
3	2.97E-07	1.14E-06	3.23E-06	3.30	1.37E-06
4	1.39E-06	4.10E-06	1.20E-05	2.94	5.07E-06
5	3.06E-07	1.22E-06	3.58E-06	3.42	1.49E-06
6	3.11E-07	1.25E-06	3.79E-06	3.49	1.55E-06
7	-	-	-	-	-
8	8.24E-07	2.56E-06	8.57E-06	3.23	3.34E-06
9	4.87E-07	1.77E-06	5.52E-06	3.37	2.22E-06
10	3.22E-07	1.34E-06	4.39E-06	3.69	1.73E-06
11	-	-	-	-	-
12	-	-	-	-	-
13	3.06E-07	1.22E-06	3.58E-06	3.42	1.49E-06
14	-	-	-	-	-
15	3.06E-07	1.21E-06	3.57E-06	3.42	1.49E-06
16	3.84E-07	1.34E-06	3.38E-06	2.97	1.54E-06
17	3.23E-07	1.35E-06	4.45E-06	3.71	1.75E-06
18	2.95E-07	1.12E-06	3.14E-06	3.27	1.34E-06
19	2.94E-07	1.12E-06	3.13E-06	3.26	1.34E-06
20	3.10E-07	1.24E-06	3.73E-06	3.47	1.53E-06
21	4.48E-07	1.60E-06	4.54E-06	3.18	1.93E-06
generic, empirical	3.36E-07	1.45E-06	5.74E-06	4.13	2.17E-06
generic, lognormal	3.61E-07	1.49E-06	6.15E-06	4.13	2.17E-06

Figure 1: Example: Results for Generator Transformer

I.C.2. ***Failure event data***

The total number of critical/non-critical failure events contained in the ZEDB amounts to 3932. Critical failures are those affecting the component such, that they do not perform the function modeled in PSA. Note that only critical failure events are considered for the evaluation of reliability parameters. Non-critical failure events are included for the sake of completeness. In some cases, regulators demand event reports for failures, which are not relevant for PSA. Such failure is marked as non-critical. The number of failure events assigned to each different prototype is summarized in **Error! Reference source not found.**.

During the repair time of a component, no further failures are considered until the component has been repaired and normalised. For this reason, replicated failures are extremely rare in ZEDB.

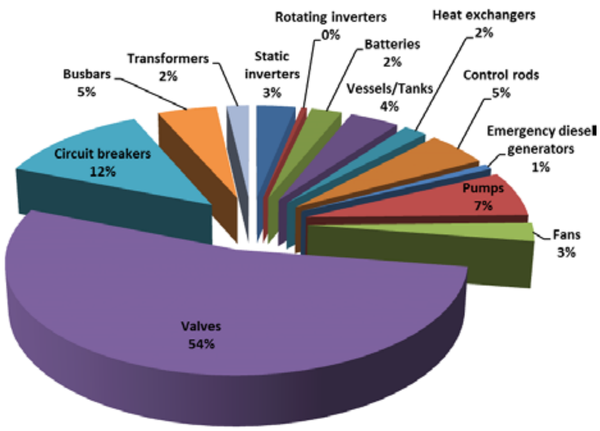

Figure 2: Distribution of the components over the prototypes.

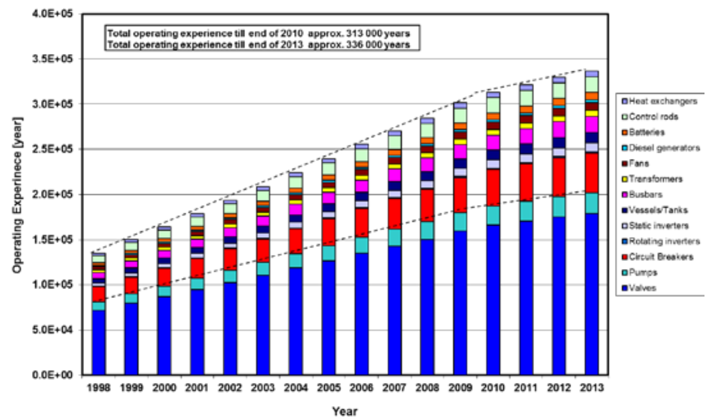

Figure 3: Operating Experience of the Different Prototype over time from 1998 until End of 2013 [1].

II. PROPERTIES AND CAPABILITIES OF THE DATABASE SOFTWARE

The ZEDB gathers and analyzes operating experience gained at nuclear power plants. The analysis is performed using a two-stage and one stage Bayesian model to calculate plant specific and generic reliability data. ZEDB brings together professionals of diverse research backgrounds, including algorithms, data management, privacy and security, user interfaces, and visualization.

II.A. Introduction

The use of a generic database application is required when an ever-evolving technical task has to be solved, which involves uncertainties regarding development of the structure, scope or level of detail of the data model. The technological innovation in particular for ZEDB is the possibility to manage a central data model which is common for all plants simultaneously with specific model extensions for each individual plant with one software application.

Moreover, the ZEDB facilitates all the functionality required for both data collection at local plants and the demands of the central operator. Keeping this in the same software enables easy maintenance, and reduces problems of compatibility.

II.B. Views in tables and forms

The design of the user interface is table-oriented, allowing the user to view and modify data in table views, in addition to a form-based view, which is provided to allow the user to modify data in a more structured manner. Examples of both can be seen in the two figures below.

In Figure 4 a horizontal split of three table views is shown.

1. The component (pump),
2. Operational information and
3. Failure events.

Obviously, this data is interrelated so that after selecting one specific component all related objects of this component are selected, as shown in Figure 5 which illustrates a form-based view containing the specific data for this component.

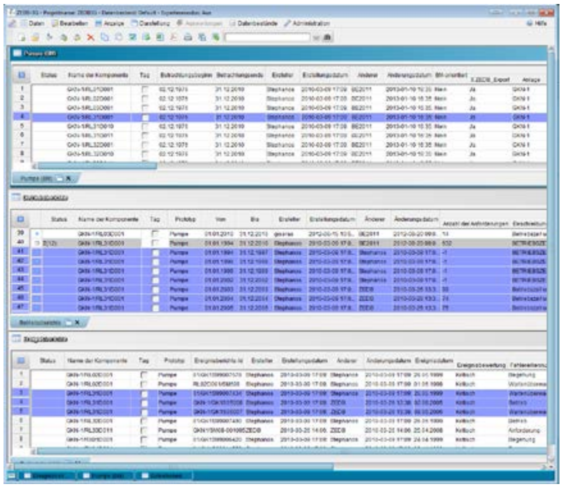

Figure 4: Horizontal split table view in the main application

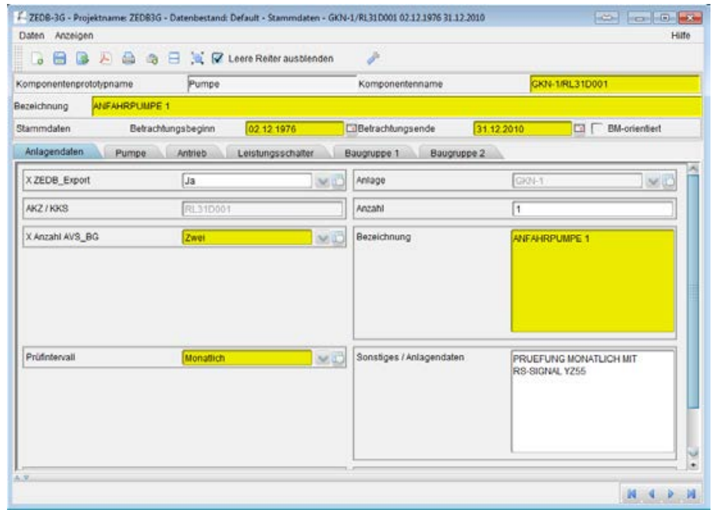

Figure 5: Classical form based view of data

II.C. Automatic statistical analysis

One stage (simple) and two stage (hierarchical) Bayesian approaches were chosen as mathematical tools to analyse the ZEDB data using Bayesian methods. In both cases, the probability distributions of failure rates and failure probabilities per demand are estimated.

Within the simple approach, a lognormal distribution for the failure rates/failure probabilities on demand is assumed. The underlying Likelihood function is the Poisson distribution for failure rate estimations and the Binomial distribution for the estimation of failure probabilities per demand.

The two-stage Bayesian models are standard practice nowadays (see [3] to [7]), although they may differ in their mathematical models and software implementation. The similarity of these applications is the assimilation of data from different sources, as illustrated in Figure 6. This is very attractive in cases where the data from the given plant are sparse.

Using these models, the so-called super-population models, the reliability parameter for any component in a particular plant can be assessed with reference to the operating experience from all others. More specific, the general form of the posterior distribution for failure rate λ_0 of a component at plant of interest 0, given failures X_i and observation times T_i of similar components at plants 0, 1, … n is derived. The exposure T_i is not considered stochastic, as it can usually be observed with certainty.

The number of failures X_i for a given exposure follows a fixed distribution type, in this case Poisson. The parameter(s) of this fixed distribution type are uncertain, and are drawn from a prior distribution. The prior distribution is also of a fixed type, but with uncertain parameters. In other words, the prior distribution itself is uncertain. This uncertainty is characterized by a hyper-prior distribution over the parameter(s) of the prior. In Figure 3, the hyper-prior is a distribution P(Q) over the parameters Q of the prior distribution from which the Poisson intensities $\lambda_1,\ldots\lambda_n$ are drawn. In sum, the model is characterized by a joint distribution:

$$P(X_1,\ldots X_n, \lambda_1,\ldots\lambda_n, Q) \qquad (1)$$

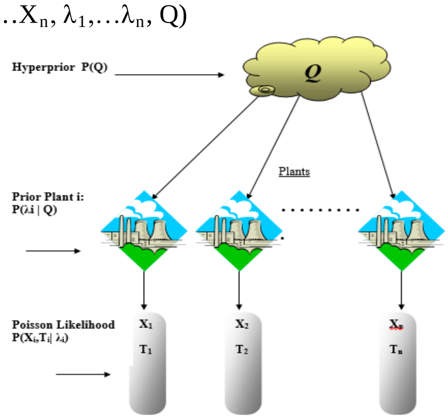

Figure 6: Bayesian Two Stage Hierarchical model

To yield tractable models, two types of assumptions are needed. First, conditional independence assumptions are made to factor (1) [8, 9]. Second, assumptions must be made regarding the fixed distribution types and the hype-prior distribution P(Q).

Although ZEDB recommends the lognormal model for the prior distribution of λ, both the lognormal and gamma models are supported. The choice of the hyper-prior is based partly on the non-informative rule of Jeffrey and partly on engineering justification [10].

II.D. Interoperability

Due to the heterogeneity of PSA software solutions that are available, ZEDB makes a strong commitment to interoperability. There are flexible interfaces for exchange data with external systems. In addition, ZEDB already includes interfaces to most known PSA software suites (Riskspectrum©, Riskman©, etc.).

III. QUALITY ASSURANCE APPROACH

Lots of effort and energy has been invested in the data collection process. Beside the measures mentioned in this chapter, yearly workshops have been held to train the staff from participating plants on how to collect and import the relevant data into the application. At these workshops, the participants gave valuable feedback on practical use of the software to help further develop the whole data collection process.

Whenever demanded, experts from the ZEDB team provided on- and offsite support during the data collection or even conducted the data collection independently on behalf of several utilities. A multistage quality assurance procedure beginning from the collecting of component data on the power plants until the evaluation and documentation of the reliability data is applied.

A ZEDB manual 2 has been prepared to provide instructions on how to compile the master data as well as the event and operating reports. This has the aim of ensuring that the data collection is consistent among all participating plants and all submitted data are of the same quality. The manual specifies the component boundaries, describes the component prototypes and explains the dropdown lists assigned to the component prototypes along with the terms contained in them.

The raw data is collected by qualified personnel from various sources at the plant and then input into the local ZEDB3G program. The software only permits the data to be incorporated into the database in a ZEDB-compatible format. To ensure that the data transferred to the database operator are free of errors and inconsistencies, an integrated testing module checks the data.

When the database operator transfers the data supplied by the plant into the central ZEDB database, data is checked via the integrated testing module. If inconsistencies or errors occur, the incorporation of the data into the database is denied and the database operator is informed by means of a test report about the errors that were found.

The data is then subjected to a plausibility check by qualified staff employed by the database operator. If the database operator should detect incorrect or at least questionable data, the data concerned is sent back to the plant with an explanation and a request for correction or clarification. If the plausibility check is successful and the data content is correct, the data is used for the evaluation of reliability parameters.

The documentation of evaluated reliability data is also subjected to a comparative plausibility check by qualified staff employed by the database operator and the release for publishing is performed by the ZEDB Steering Committee.

IV. TREND OF THE RELIABILITY PARAMETER ESTIMATED WITH THE ZEDB DATA

The increase, decrease or constancy of the failure rates estimated in the ZEDB depends on the ratio of the growth of the number of failures events to the growth of the operating time, which mainly depends on the number of components within the considered collective

Based on the observation that the relative increase of the operating experience gradually decreases over time, the relative change of the failure rate is mainly dependent on the relative change of the number of the failure events. Because of the fact that the greater the total number of failure event is, the smaller the relative change of the failure events will be, three representative component populations (with different number of events in 1998) were defined.

For these component populations, the trend of the failure rate has been evaluated for different points in time, starting from 1998 until the end of 2013 with a constant linear step width of one year. The trend analyses shows that the component populations with a total number of failure events larger than 10, can be considered as rather stable and insensitive to the occurrence of additional failure events. The consideration of additional operational experience less than 10% of the total number of the existing failure events will lead to a reduction of the failure rate. Thus, a further evaluation of these collectives is not strictly necessary.

For component populations with a total number of failure events $1 \leq N < 10$, the most sensitive collective is the one with the total number of failures equal one. This is because of the approximately doubling of the mean value of the failure rate by considering only one further additional failure event. In contrast, the component populations without previous failure events are the most sensitive ones. Therefore, a revaluation of these component populations is recommended.

In order to obtain a general assessment of failure rate tendency for all collectives evaluated in the ZEDB, the percentage fractions of the prototype specific collectives depending upon the total number of failure, which defines the trend of the reliability data, is presented in TABLE 3. About 25% of all formed component populations are sensitive against the consideration of further failure of events. About 47% of all component populations with a total number of failure events 1≤ N< 10 can be considered as less sensitive. Due to the sufficient operating experience and in particular due to the large number of failure events, 27% of the total number of formed component populations can be considered as insensitive against further consideration of operating experience.

TABLE 3: Fraction of the prototype-specific collectives depending on the total number of failure events

Prototype	Number of failures		
	N = 0	**1 <N<10**	**N ≥10**
Static inverters	0%	0%	100%
Rotating inverters	0%	0%	100%
Batteries	45%	0%	55%
Vessels/Tanks	1%	54%	45%
Heat exchangers	38%	23%	39%
Control rods	0%	67%	33%
Emergency diesel generators	0%	67%	33%
Pumps	41%	33%	26%
Fans	75%	0%	25%
Valves	27%	52%	21%
Circuit breakers	69%	18%	13%
Bus bars	50%	50%	0%
Transformers	85%	15%	0%
Related to all component population	**25%**	**47%**	**27%**

Since the operating experience contained in the ZEDB is not going to be further updated, a revaluation of these components population considering the generic prior from other reliability data sources is recommended. There is a free text describing the details of the failure in the database, which can be used to conclude on failure mechanisms and to locate the failure in plant documentation, which may provide additional insight.

V. POTENTIAL USES OF THE ZEDB DATA BASE

ZEDB comprises a large amount of operating experience, which enables the creation and the evaluation of collectives with respect to the component function.

V.A. Use as Prior Information in Probabilistic Safety Analysis

The trend analyses have shown that the evaluated populations with a total number of failure events ≥ 10 are insensitive and that the consideration of further operational experience with additional failure events < 10% of the total number of existing events will lead to a lower failure rate. Furthermore, the ZEDB provides the reliability data for more than 75% of the mechanical and electrical components included in Probabilistic Safety Analyses (PSA) for pressurized water reactors (PWR) and boiling water reactors (BWR). Thus, the ZEDB is the reliable basis for the use in PSA.

As ZEDB considers variability across different plants, the generic results can be used as prior information for other plants, or as a generic basis for a new plant, which has no plant specific information yet.

V.B. Use for Aging Studies

Specific analysis concerning aging has been performed for Diesel generators [11], though some additional data (times of renewal) had to be collected to perform this task. Due to the flexibility of the query tool and its capabilities, it has been easy to collect the information, and it could be shown, that there is nearly no aging in emergency Diesel generators (in German NPP).

V.C. Use for Maintenance Management

PSA studies and maintenance management usual reside in different parts of the organization. This issue should be questioned, as component failures considered for failure rate generation in PSA influence cost of maintenance. As far as safety related components are concerned, a lot of data is already available in ZEDB. Maintenance events would have to be added, and also cost terms for maintenance, like cost per failure event and cost per unit time under repair. Whereas reliability centered maintenance (much like risk based regulation) have turned out to be too restricted approaches, a concept of reliability informed maintenance may turn out to work. It will reveal components, where maintenance could be reduced without prohibitive consequences in reliability.

V.C. Use for CCF Data Management

Due to the generic concept of the data base, structures to hold CCF group data could be configured easily. A code implementing calculation of alpha factors is available and has to be integrated in this case as an additional analysis option.

VI. SUMMARY AND CONCLUSIONS

ZEDB is a data base and data collection and evaluation project with an experience background of some 20 years. Due to rigorous quality management procedures, the data in ZEDB has extremely high quality. Evaluation includes both plant specific data and generic data reflecting variation of the failure rates or failure probabilities across different plants. The latter can be used for new designs, where no specific data is available yet.

ZEDB is available to add raw data of additional plants. In this case, the existing information can be used to generate the prior data. As a lot of technical data for the components is available, there is a base to define collectives to generate prior distributions very precisely.

ZEDB is extensible towards optimization of maintenance costs. There is some additional data required, but much less in comparison to a project starting from zero.

Especially the last mentioned possibility appears to have big potential on return on investment. As maintenance management and probabilistic safety assessment are normally tasks of different parts of the modulation, and they are normally not seen simultaneously.

ACKNOWLEDGMENTS

ZEDB has been substantially supported by international NPP utilities organised in the VGB PowerTech Service GmbH.

REFERENCES

1. ZEDB Evaluation 2014, VGB PowerTech, 2014, Essen, Germany
2. Handbuch zur ZEDB-Datenerfassung, Version 4.2, erstellt durch AREVA GmbH im Auftrag der VGB, Erlangen 2008
3. Kaplan, S. (1983) "On a two-stage Bayesian procedure for determining failure rates from experimental data",*IEEE Trans. On Power App. and Sys.* Vol. PAS-102, 195-202.
4. Kaplan, S. (1985) "Two-stage Poisson-type problem in probabilistic risk analysis" *Risk Analysis*, 5, no. 3 227-230.
5. T-Book Reliability Data of Components in Nordic Nuclear Power Plants, ATV Office, Vattenfall, AB, S-162, 87, Vallingby Sweden.
6. Pörn, K. (1990) *On Empirical Bayesian Inference Applied to Poisson Probability Models*, Linköping Studies in Science and Technology, Dissertation No. 234, Linköping.
7. Becker, G. and Hofer, E. (2001) Schätzung von Zuverlässigkeits-Kennwerten aus der Evidenz verschiedener Anlagen GRS-A-2918.
8. Iman R. & Hora S. 1990. Bayesian modeling of initiating event frequencies at nuclear power plants, Risk Analysis, vol. 10, no.1, pg. 103–109.
9. Pörn K. 1990. On Empirical Bayesian Inference Applied to Poisson Probability Models, Linköping Studies in Science and Technology, Dissertation No. 234, Linköping.
10. Becker, G. 2009. Theorie des zweistufigen Bayes-Verfahrens und deren Implementierung im Programm Bayes 2.0.
11. Blombach, J., Brahmstaedt, K.U., Camarinopoulos, L., Does ageing of NPPs require the incorporation of time dependent failure rates in PSA models - Verification using operating experience of emergency Diesel generators collected in ZEDB, 31st ESREDA seminary, 2006.