Evaluation of Radiographic Testing Performance with an Advanced POD Approach

Daniel KANZLER*, Christina MÜLLER*, Jorma PITKÄNEN**
* BAM Federal Institute for Materials Research and Testing; Berlin, Germany;
daniel.kanzler@bam.de, christina.mueller@bam.de,
** Posiva Oy; Eurajoki, Finland; jorma.pitkanen@posiva.fi

Abstract. POD is the most often used approach for the evaluation of the reliability of non-destructive testing methods in critical areas. But especially for those cases the requirements for a sufficient amount of data, adequate type of defects and the statistical requirements are high. The claim for checking these requirements is not new, but never the less often ignored. Especially the large amount of data for critical real defects seems too often to be not really solvable. In this presentation we show an approach to use real defect data with knowledge from former testing results with the same equipment, in a useful adequate way. Furthermore we introduce a method to evaluate the real defect description parameters for the radiographic testing. This approach was used for a joint project with the company POSIVA, which is building a final depository for spent nuclear fuel in Finland. Radiographic testing is one of the NDT-methods they use to test the electron beam weld of the copper canister. The copper canister will be used in the deposit as a corrosion barrier within the deposit concept. Through the high sophisticated manufacturing methods of the canister there are almost no material defects, which could be used for reliability evaluation. In this context the idea of Bayesian Updating of the POD with real defect data with former results from artificial defect data was born to get a useful evaluation.

Introduction

The evaluation of a non-destructive testing system (NDT) is essential for the structural integrity of the tested objects. The question "how reliable is a NDT system" can be answered in different ways. One answer is given by the probability of detection (POD), which allows evaluating the biggest defect which might be missed by the testing process. Over the last four decades the POD became the most often used method to evaluate the reliability of NDT systems. But the main challenge still remains - the evaluation is often based only on artificial defects or simulations. Based on this, assumptions about the "reality" are made. In contrast, to estimate the POD using real defect data the approach which is described here tries to combine the artificially made defects with the few real defect data to receive a statistically-based method on real defects.

Advantages of combine data instead of pooling

The easiest way to put data together is the so-called "pooling". In pooling, data get mixed without weighting or under consideration of different statistical models. The amount of data is rising with every experimental data point. There is no complex statistical model needed and the results are objective without subjective considerations. Often the pooling of data is not adequate and may give misleading results. But despite that, a sense-making combination gives the opportunity to include human operator, which can have influence on the POD results. In the case of a POD with real and artificial defects, the real defects count more for the evaluation of the NDT system. This should be taken into account in the combination. A pooling of similar data cannot be done, when a combination in a logical way is possible. The combination is necessary, because both data sources should influence the result of the POD.

Bayesian approach

One popular approach to combine data is the Bayesian statistics. In Bayesian statistics, the experimental data is used in a so-called likelihood function [P(B|A)] and is combined with a priori knowledge [P(A)] from the former experiment. In that case, the likelihood function includes a small amount of real defects, while the a priori knowledge expresses the information from other sources of data, typically former experiments, simulations, expert opinions, etc. As a result we get an updated posterior knowledge [P(A|B)], as it can be seen in the following equation, which contains both kind of data [6]:

$$P(A \mid B) = \frac{P(B \mid A)}{P(B)} P(A)$$

In the case of the calculation of the POD, with an approach published by Berens, the task is to estimate the parameters for the normal distribution, which will be the POD curve. The combination takes part in the estimation of these parameters. We assume that the artificial data and the real data can be expressed by a normally distributed model, which simplifies the calculation.

Use of the Bayesian approach for pod

The Bayesian approach, presented here, was used for the evaluation of radiographic testing of the electron beam weld of the copper canister for the final disposal of high-level radioactive waste. The aim is to evaluate the reliability of RT in the production process using a POD based on real defects, which may appear in the production process. This is particularly vital for highly sophisticated systems and in case of processes, during which a failure may lead to serious consequences. However, producing real defects is related to high costs and is sometimes hardly possible. On the other hand a big amount of data on various kinds of defects is needed for the statistical analysis, which is the mathematical basis the POD operates upon.

We consider supporting the data of the experiments on real defects with the prior knowledge which we possess before these experiments. The prior knowledge is achieved from experiments with artificial defects. The so-called likelihood knowledge is built up of the data from experiments with real defects. Both prior knowledge and likelihood knowledge can be combined in the posterior information. The Bayesian approach allows us to express the posterior information and the prior and likelihood knowledge through different distribution functions. The expected benefit is the smaller confidence band and a better estimation of the distribution parameters.

The mathematical process is similar to the often-used Berens [5] approach: The maximum amplitude is plotted versus the significant defect parameter causing the signal (Figure 2a) –in this case, the depth of the crack (size) [5]. The plotted graph is called â vs. a graph. In the common POD model for radiographic testing the highest grey values or contrast to noise ratio (CNR) (â) and the adequate penetrated length (a) are evaluated [3]. First this linearity of defects and signals should be shown.

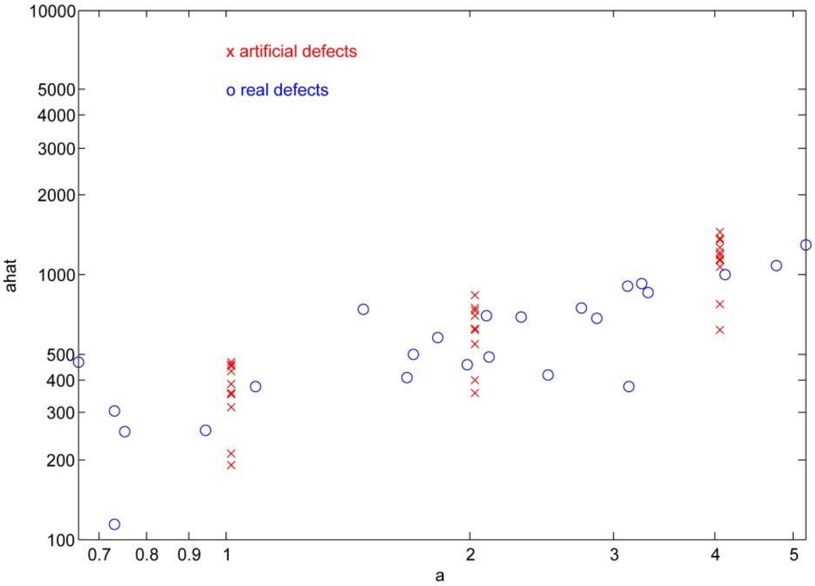

Figure 1: Data of artificial defects and real defects for RT

We assume that both real defects and artificial defects have a normally distributed spreading around the linear models. So in both cases, following the Berens approach, we estimated a normal distribution with a mean value and a deviation for the probability of detection for real and for artificial defects. From this step on, there is a profit of using the Bayesian statistics - through the update of the likelihood distribution function for real defects with a prior distribution for artificial defects the estimation will be better than only based on one kind of data.

We take care about this additional information in the confidence bounds in an adequate way [17]. The POD curve with the lower 95% confidence band is a typical way to present the capability of the NDT system to detect a flaw [5]. To guarantee the integrity of the canister and to evaluate the NDT method, the size of the defect that is detected with 90% probability and 95% confidence, has to be determined. This measure is called $a_{90/95}$.

The maximum likelihood estimation (MLE) of the â versus a relation creates a distribution of the data points, that relate to â in dependence of a. We used the following two parameters of the MLE to define the distribution for the Bayesian approach: One parameter is the mean value, which is defined as the linear function, and the other parameter is the variance, which describes the scattering of the experimental values around the linear function.

Furthermore, the calculated covariance matrix of the MLE provides an estimation of the goodness of the mean value and the variance value [9]. The confidence band is defined by the covariance matrix, by the amount of data and by the level of confidence, which was defined as 95% [9].

The estimated parameters and their confidence band create a bivariate normal distribution for the prior function and another one for the likelihood function. Each of the distribution functions describes the behavior of the NDT system.

We update the likelihood bivariate distribution of the real defects with the prior bivariate distribution of the artificial defects to the bivariate posterior normal distribution function, according to the Bayesian theorem. Therefore, the number of data for the posterior function, which has a large influence on the width of the confidence band, equals to the sum of a scaled amount of prior data plus the amount of likelihood data. For the calculation of the scaled amount of the prior data the equation of reference [11] was used.

The difference of the data sources

We decide to use the Bayesian approach for combining artificial data and real data, because of different behavior of data sources: the cost to create the artificial defects is, in respect to real defects or realistic defects, much higher. However, the value of information of real defects is higher, because the information we get is much nearer to the real situation and the defects which might occur in the later production. Both facts are leading to a concept that real defect data should have a higher weight than the artificial data. One possibility to do a combination based on that fact is using the Bayesian approach [6].

The uncertainty we have through the different data sources is coming from different facts. A main influencing factor on experiments with artificial defects is the drilling process. In realistic defects the uncertainty of the process happens with metallographic process. Furthermore, the NDT signal in radiography is changing much faster in case of a realistic defect than in an artificial defect due to the form of the defect, which can have an influence on the signal distribution for both defects. This additional fact can be considered within the Bayesian approach.

Results: Bayesian approach for the pod

We first calculate the POD based on a smaller amount of real data (24 in black). Based on that result we used the above mentioned Bayesian combination of data to receive a better POD (green).

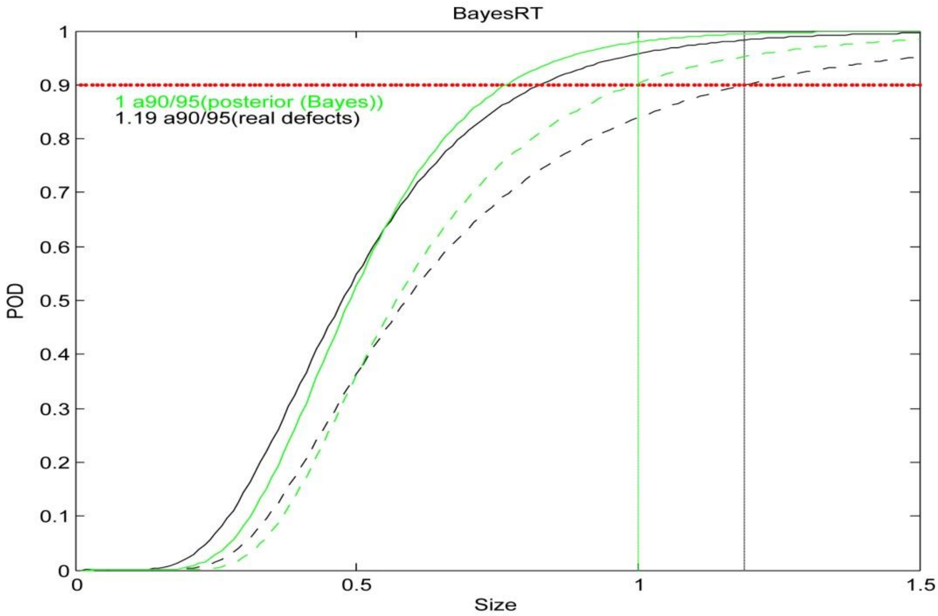

Figure 2: POD of real defects with and without the Bayesian approach

As a comparable key parameter we chose the size of the defect that is detected with 90% probability and 95% confidence, the so-called a90/95. The number of experiments increased from 24 on the real defects to 29 on the combined amount of defects. Through this increased number of data and the joint values of the mean and the variance value, the a90/95 decreased from 1.2 mm to 1.0 mm.

The comparison with the POD calculated with a three times greater pool of real defects which was in this investigation available (see the POD curve in figure 7), the POD curve of the Bayesian approach provided a better result than the POD curve of only a few real defects. The a90/95 from the Bayesian approach of 1.03 mm was verified through the result from bigger data pool of real defects of 1.05 mm. For an overview about the results see table 1.

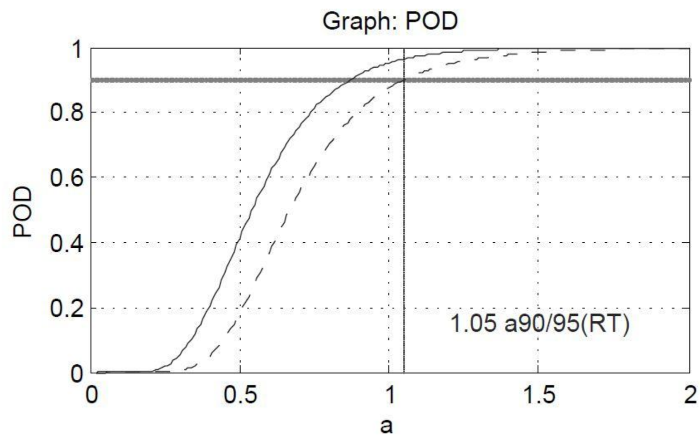

Figure 3: POD with a bigger amount of data

Table 1: Comparison of the results

Characteristics	Small amount of real defects (24 defects)	Bayesian approach	Big amount of real defects (72 defects)
a90/95	**1.29**	**1.03**	**1.05**
μ of norm. Distr.	-0.74	-0.71	-0.61
σ of norm. Distr.	0.43	0.35	0.37

Conclusion

Our approach reveals the characteristic a90/95 based on a small amount of real data and the Bayesian approach that is comparable with the evaluation of a bigger amount of data. The approach improves the estimation of the statistical parameters and provides the confidence band with a scaled amount of data from artificial defects and real defects. The approach combines the data in a statistically correct way. Instead of having a small amount of data, we have to pay close attention to deal with a more sophisticated procedure to calculate a meaningful POD. The demanding numerical calculation needs a cautious procedure with information and data.

The Bayesian approach enables a combination of different kinds of data, i.e. data from real and artificial defects. Furthermore, there is a possibility to weight the data according to their value for the evaluation. Additionally, it is possible to develop a POD approach based not only on one amplitude value but on data fields (data field POD), more sophisticated thresholds (observer POD), etc. [15][16].

In future work, we aim to focus on the testing of statistical requirements, on using other distributions, beside the ones previously used, on other NDT methods (e.g. eddy current and ultrasonic testing [13][14]).

References

[1] Pitkänen J, Salonen T, Sandlin S, Ronneteg U, "Defect detectability in EB-welded copper disposal canister with 9 MeV accelerator", *6th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurized Components*, Budapest, 2007

[2] Pitkänen J. Posiva Report 2010-04, Inspection of bottom and lid welds for disposal canisters, September 2010, 98p.

[3] Kanzler D., Milsch S., Pavlovic M., Müller C., Pitkänen J., "Concept of total reliability of NDT methods for inspection of the EB weld of the copper canister used for a long-term storage of spent nuclear fuel", Proceedings of the Eighth International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurized Components, Berlin, 2010

[4] Rummel W. D., "Recommended Practice for a Demonstration of Nondestructive Evaluation (NDE) Reliability on Aircraft Production Parts", Materials Evaluation, Vol. 40 August 1982

[5] Berens A P, NDE Reliability Data Analysis – Metals Handbook, Volume 17, 9th Edition: Nondestructive Evaluation and Quality Control, ASM International, OH, 1989

[6] Bury, Karl V. John Wiley & Sons, J. W. (Ed.) Statistical Models in Applied Science *Wiley Series in Probability and Mathematical Statistics*, 1975

[7] Pitkänen, J.; Paussu, R.; Pohjanne, P.; Virkkunen, I.; Kemppainen, M.; Lipponen, A.; Sarkimo, M.; Simola, K. & Reddy, K.-M. Metallographic study of Detected Indications in EB-Copper welds for

Verifying the NDT Reliability of Inspection *8th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurized Components* September 29 - October 1 2010
[8] Pavlovic, M., Müller C., Takahashi K., Pitkänen J., Ronneteg U., 2009, Multiparameter influence on the response of the flaw to the phased array ultrasonic NDT system., 4th European-American Workshop on Reliability of NDE, 2009
[9] Cheng R.C.H., Iles T. C., "Confidence Bands for Cumulative Distribution Function of Continuous Random Variables", Technometrics, Vol. 25, No. 1. 1983
[10] Cheng R.C.H., Iles T. C., "One-Sided Confidence Bands for Cumulative Distribution Functions", Technometrics, Vol. 30, No. 2., 1988
[11] Stange K., "Bayes-Verfahren, Schätz- und Testverfahren bei Berücksichtigung von Vorinformationen", Springer-Verlag, Berlin, 1977
[12] MIL-HDBK-1823A, "Nondestructive evaluation system reliability assessment", Department of Defense Handbook, 2009
[13] Pitkänen, J., Salonen, T., Sandlin, S. & Ronneteg. U., "Defect detectability in EB-welded copper disposal canister with 9 MeV accelerator ", 2008
[14] Pitkänen J. "Inspection of Bottom and Lid Welds for Disposal Canisters", Posiva 2010-04, September 2010
[15] Pavlovic, M.. Müller, ,C., Takahashi:, K.., Pitkänen, J., Ronneteg, U., 2009, Multiparameter influence on the response of the flaw to the phased array ultrasonic NDT system., 4th European-American Workshop on Reliability of NDE, 2009.
[16] Takahashi, K., Pavlovic, M., Ronneteg, U., Pitkänen, J., Müller, C., 2009, POD from C-Scan Data for Reliability Analysis of Ultrasonic NDT, 4th European-American Workshop on Reliability of NDE, 2009.
[17] Kanzler, D., Müller, C., Ewert, U., Pitkänen, J., 2012, Einsatz der Bayesschen Statistik in der Zuverlässigkeitsbestimmung von zerstörungsfreien Prüfsystemen., Materials Testing 4/2013 Volume 55 Hanser