



Evaluating RT systems with a new POD approach

Daniel KANZLER¹, Christina MÜLLER²

¹ KNC-Reliability, Berlin, Germany

² BAM Bundesanstalt für Materialforschung und -prüfung, Berlin, Germany

Contact e-mail: KanzlerD@knc-reliability.com

Abstract. The usefulness and purpose of evaluating nondestructive testing (NDT) systems and their capabilities has changed in the last decade. The conventional method of simply applying a familiar statistical algorithm to say whether the system is usable for the tasks is history.

Nowadays, multiple parameter methods which describe the probabilities of detection (POD) of different systems or real defects need new characteristics and a broader variety of statistical models to describe the true system behaviour.

The appraisal of the NDT system involves diverse departments within a company (engineering, NDT-operators, and statisticians), but is, at the same time, more needed and requested than in the past.

In this article, an approach is discussed in which professionals from different fields worked well together, accomplishing cost-intensive metallographic studies in correlation with well-understood physical behaviour of NDT-methods as well as deep-discussed mathematical methods to create a holistic evaluation of the technical reliability for a specific radiographic testing (RT) equipment.

The first part of the publication will show the comparison between metallographic grinding and the RT indications. An essential innovation over past evaluation methods was the use of a multi-scale smoothing algorithm, which describes physical parameters, which were not used in evaluation like the POD in this way in the past.

In the second part the statistical requirements for the POD take the focus. It can often be hard to make significant statements; especially in the case where only a small amount of data is available. The combination of data and the use of knowledge from simulations are essential. One possible solution will be shown for the RT evaluation.

The methodology is used for evaluating the digital RT system for the inspection of electron-beam welds, which was method considered to seal the Finnish copper canisters for the final deposit of spent nuclear fuel.

1. Probability of Detection as versatile tool

Producing 100% defect-free components is not possible. To be certain that the product can fulfil its designated task, nondestructive testing (NDT) is essential. The evaluation of an NDT system is therefore an important task in order to determine which kind and what size of defects might still be in the component after testing. The evaluation itself can be used to make a comparison between different testing systems and is able to assess if the system is adequate for the situation, in which it will be used. Another, and perhaps even more important, role of evaluation is to communicate between the different departments to build a holistic lean quality management over the whole life time of the product, e.g. designers,



structural engineers, and NDT-operators, which enables high quality and safety for the use of the product. The evaluation, which is often used in this case, is the probability of detection evaluation (POD); an objective and understandable summary of the evaluation which fulfils these tasks.

Simultaneously while the range of use for the POD is expanding, the requirements on the POD are becoming more demanding. The one-parametric approach (according to MIL-HDBK-1823 A) cannot be used in more critical or ambitious situations, e.g. for real defect situations.

In this paper a comparison of metallographic grindings and radiographic testing (RT) indications is made, to show that the handling of real defects provide a further challenge for POD evaluation. Additionally a new approach using a mathematical method based on smoothing will be used to describe the correlation between the defect and its indication more rigorously. The new method to calculate the POD leads to a more realistic evaluation of the NDT method.

1.1 Basic mathematical model for calculating the POD

For creating new approaches based on the well-known POD methods, there will be a short introduction of the requirements and the general process of calculating the POD. A larger defect has, in general, a higher probability of detection in comparison to smaller defects. This is essential for reducing the necessary amount of data to get a significant result. This relationship is based on the functional physical behavior of the NDT system and makes the signal-response POD very useful in comparison to hit/miss approaches. In the case of the RT the physical law, which describes the relationship between an RT signal and a main defect parameter is the attenuation law: The relative contrast C_r of small wall thickness variation Δw (small defects relatively to the wall thickness) is directly proportional to the attenuation coefficient μ_{eff} and the defect size in the beam direction Δw , which is called penetrated length:

$$C_r \propto \mu_{\text{eff}} \Delta w \quad (1)$$

In the POD for RT this relationship is used by the regression analysis, while (\hat{a}) is the contrast and (a) is the penetrated length of the defect (Δw). It can be described as approximately linear for small defect:

$$\hat{a} = \beta_0 + \beta_1 \cdot a + \varepsilon \quad (2)$$

β_0 and β_1 are the parameters taken into the regression model and ε is the residual error (variance of the data which are not explained by the regression model).

The second essential mathematical element in the POD evaluation is the description of the error of the regression model ε . Often the distribution model is assumed to be normally distributed. It is defined by two parameters: μ the mean value and σ the residual deviation of the data. In the case of the error of the regression model, the mean value μ equals 0 (equation (3)).

$$\varepsilon = N(0, \sigma) \quad (3)$$

The residual deviation σ of the experimental data around the mean value – in our case the main physical law - is assumed to remain constant along the regression line (homoscedasticity) (1). An additional assumption is the independence of the experiments (2). These assumptions and models are necessary to meet to calculate a conventional POD, which is often used in the practice to evaluate NDT systems (1).

1.2 Decision thresholds

A decision threshold indicates whether an indication will be defined as a defect signal or noise signal. Which means that if there is a defect, that it is detected or not. The decision threshold which will be used here is a detection based threshold, for which the visual detectability of a defect plays a role (1). The usual procedure to define the detection threshold is based on the noise and the expected false alarm rate (typically used in receiver operating characteristic (ROC)) (3). The most frequently used threshold is a constant that does not change with signal height and it is easy to handle in the POD calculation (see Figure 1).

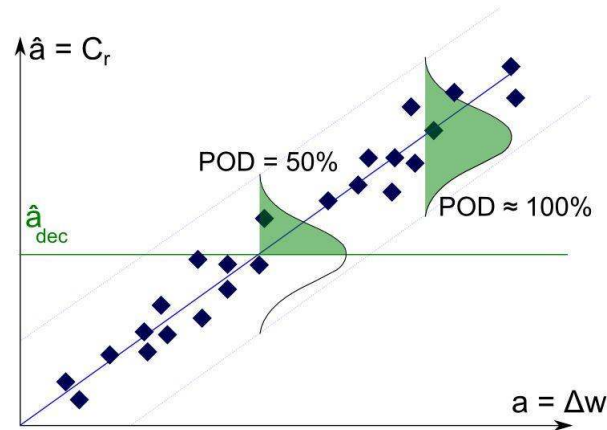


Fig. 1. The signal-response (\hat{a} vs. a) graph for the calculation of the POD according to the radiographic attenuation with a constant decision threshold for the main defect parameter Δw and the relative contrast C_r .

It has been shown, that in radiographic testing beside the signal information the size and the form of the indication are important for the detection process (4) (5). In the original POD additional conditions, like these, are not considered. This might lead to an underestimation in the range of small values of the defect parameters and to an overestimation for higher values, which demands a more adequate threshold for a realistic POD evaluation.

State of the art solutions are multi-threshold approaches (6). Some approaches define the threshold for specific areas or different forms (cylinder hole-shapes). But also the indication size has an important influence on the detectability (7), (8). Based on this idea the operator threshold was introduced into the POD for RT systems (9).

In this article another approach will be introduced in which the smoothing of data is the key element to introduce the dependence of the indication area. Therefore metallographic data will be used for the penetrated length of the defect and the corresponding RT image as the perceived contrast signal of the defect. This work was part of a project that evaluated NDT systems for assessing the weld of a canister for the final deposit of spent nuclear fuel in cooperation with the Swedish and Finnish companies responsible for that task (SKB and Posiva Oy).

2. An indication size dependant threshold for the POD

Based on the ideas of the detection theory for RT indications (10) a POD approach with a size dependant threshold was introduced (9). With simulated RT images and different test subjects the size dependence was validated. An example for the observer threshold is

shown in **Fehler! Verweisquelle konnte nicht gefunden werden.** Therefore, it was assumed, that for the evaluation the CNR_{min} is objective and operator independent.

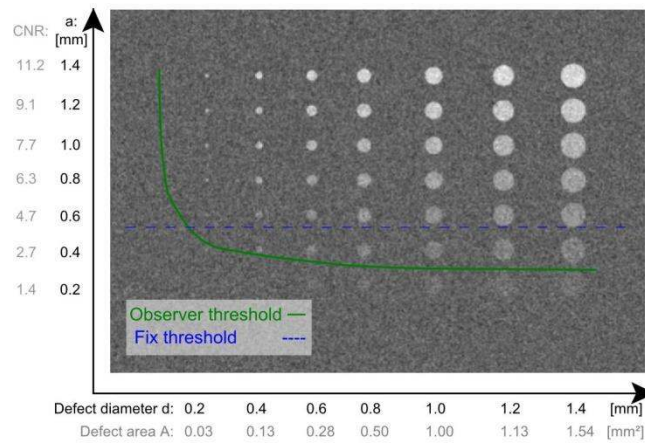


Fig. 2. Observer threshold in comparison with a fix threshold for RT indications of hole like defects

Due to the immense amount of data and work necessary to implement an observer threshold another approach was introduced. This article will give an overview about the process.

2.1 POD requirements for real data

With the evaluation of real defects the above mentioned approach will be hard to use. The amount of real defects are less in their number and even for realistic defects the costs are too high to get a sufficient amount of data. Another fact is the individual form of real defects, which are significantly different from round holes (see Figure 3).

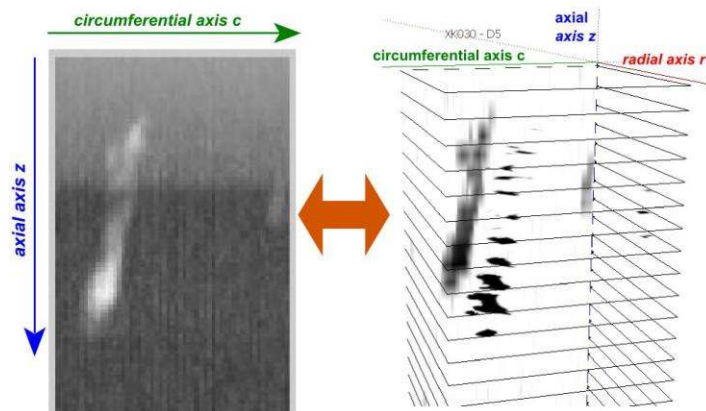


Fig. 3. Exemplary analysis of one realistic welding defect of an electron beam weld for copper canisters for the final disposal: Indication in the RT image and metallographic slices in few millimetre intervals.

Therefore an approach based on a smoothing algorithm is introduced. The first approach was based on the indication dimensions in circumferential direction. At the same time the amount of data was sufficient to evaluate a POD based on the amount of metallographic data.

2.2 Smoothing approach for calculating a POD for real defects

The idea of the smoothing approach is to use a kernel G to convolve the metallographic data and the RT image for the slice with this kernel. For the metallographic data the physical attributes in the test, like uncertainty of the RT system, can be established in the model. The smoothing of the RT indications can, on the one hand, decrease the noise, which is a typical detection behaviour, and on the other hand small indications will decline in their indication height. This is the same behaviour which was seen in the observer threshold.

In this case an Epanechnikov kernel was used (11):

$$G(t) = \begin{cases} \frac{3}{4}(1-t^2), & t \leq 1 \\ 0, & t > 1 \end{cases} \quad (4)$$

For the use of different indication sizes different bandwidth were introduced:

$$G_h(t-x) = \frac{1}{h} G\left(\frac{t-x}{h}\right) \quad (5)$$

In the next step the original data $y(t)$ were convolved with the abovementioned kernel:

$$T_h(x, t) = \sum_t y(t) \cdot G_h(t-x) \quad (6)$$

As an example one slice from the above mentioned defect is shown here, convolved with a kernel with different bandwidth h to see the influence of the kernel (**Fehler! Verweisquelle konnte nicht gefunden werden.**).

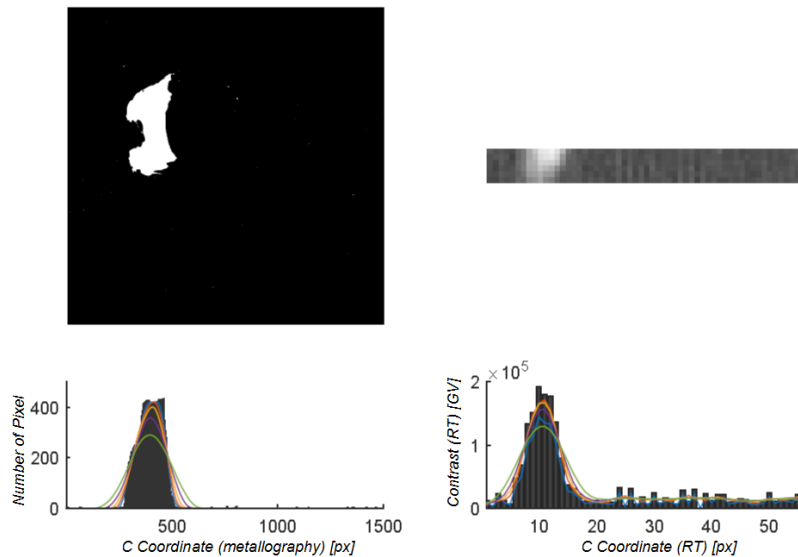


Fig. 4. Evaluation of a defect slice and its RT indication area with a kernel with different bandwidth (RT: $h(\text{blue}) = 2\text{px}$, $h(\text{violet}) = 4.2\text{px}$, $h(\text{red}) = 6\text{px}$, $h(\text{yellow}) = 7\text{px}$ and $h(\text{green}) = 8.5\text{px}$) to include the circumferential indication size in the POD

The bandwidth parameter, in this case, was adapted to the width of the defect. Therefore, different comparison with the above mentioned observer POD were performed and a model for using the bandwidth according to the observer threshold was included. An example of choosing the bandwidth is shown in Figure 5. The process of how to choose bandwidth and connect h with the observer threshold and the size of the defect is described in (11). In this article our focus is on the procedure for using the general approach.

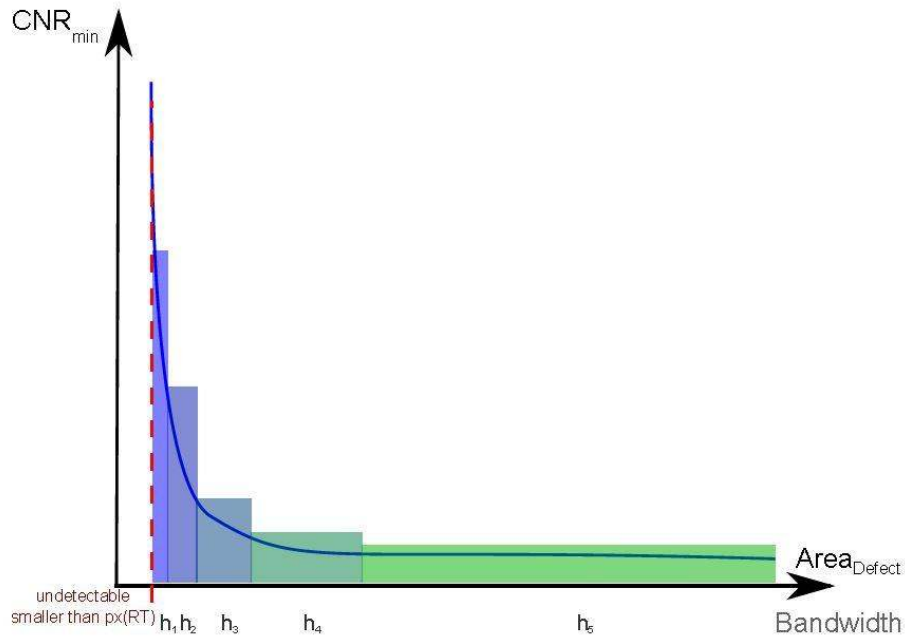


Fig. 5. The connection of the observer threshold and different bandwidth for the smoothing approach can be seen in the schematic graph

3. Detectability for different real defects

The approach of using a smoothing kernel to establish the indication size in the POD was first used on modelled data and after that on artificial data to verify the functionality of the approach. Both experiments can be found under (11). For the real defect evaluation every slice was connected with a part of the RT image with a height of 6 px to take into account uncertainties, in rotation and on the exact sliced position. The resolution for the RT image was 0.4 mm, while the resolution of the metallographic picture was 0.01 mm. At this point another advantage of the smoothing is visible. The different resolution of the data is after the smoothing present no further problem. 43 different defects were evaluated in this approach. For this amount of data a multilevel threshold evaluation is not possible to receive a statistically meaningful result. For the decision threshold a contrast to noise ratio of 2.5 was chosen. For the bandwidth of $h = 2\text{px}$ and $h = 8.5\text{px}$ the decision threshold was between 65 000 GW and 43 000 GW. Already here the assumption of the decreasing noise level is seen.

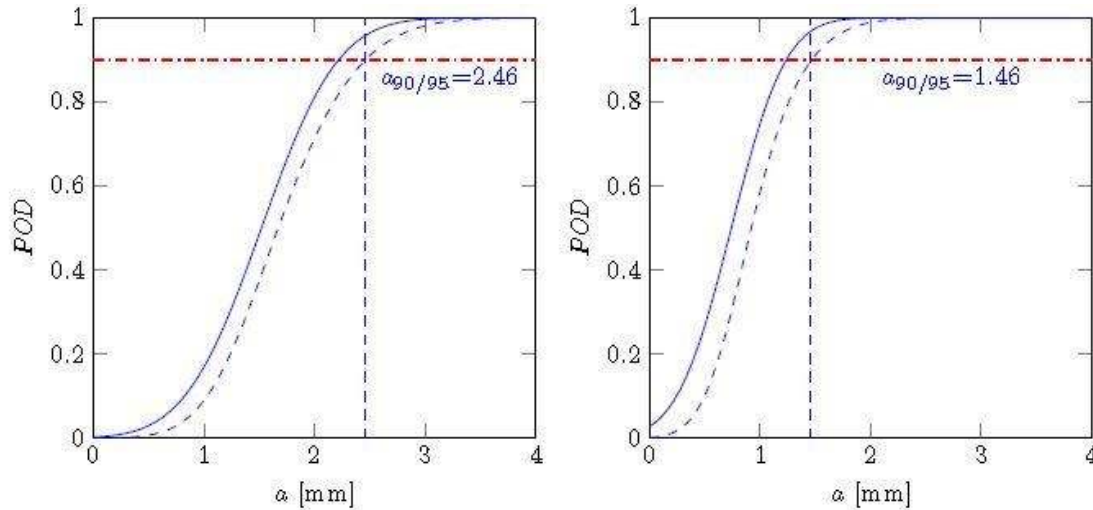


Fig. 6. Results for POD with smoothing kernel with the bandwidth of 2(left) and 8.5(right)

In the further analysis the bandwidth of 2 stands for a defect area of 0.5 mm² and $h = 8.5\text{px}$ identifies a defect of 2.5 mm² perpendicular to the penetrated length.

4. Discussion of the results

The results of the two different bandwidths yielded two different results. As expected, with a larger bandwidth the reliable detectable defect size (in penetrated length) is smaller. This is the result of the declining noise and the scattering which narrows with higher bandwidth. In this area, the reduction of the signal of small defects has a secondary influence.

The bandwidth has an influence on the results and on the decision threshold as it was expected. It was also possible with the help of the observer threshold to establish a connection with the area of the indication, which makes it useful for holistic damage tolerance considerations

4.1 Conclusion and future steps

The described method uses the dependency of the detectability and of the indication size to improve the concept of POD towards the real operator based detectability of defects. It was shown that the area of the projected indication is relevant for the POD and needs to be considered. The implementation of the smoothing worked for real defects, even for a small number (43), which is not possible for multilevel or observer thresholds. With the discussed approaches the detection threshold is near to a real one for pores, and the POD evaluation is close to the real detection behaviour of human operators.

In the next steps additional data will be evaluated to verify the process and provide further insights. The connection with the observer threshold will be verified. A set of guidelines on how to use the different bandwidth is also planned for the future. Finally, the expansion for the indication area in width, height and form is still under development

5. Acknowledgments

The authors would like to thank Ulf Ronneteg from SKB Sweden for his valuable comments and feedback and also Prof. Dr. Spokoiny for his expertise about the multi-scale smoothing method.

References

- [1] **Berens, Alan P.** NDE reliability data analysis. [book auth.] ASM International. Metals Handbook Vol. 17 Ed. 9. 1989.
- [2] **U.S. Department of Defense.** Non-destructive evaluation system reliability assessment - MIL-STD-1823. 2009.
- [3] **Feistkorn, Sascha and Taffe, Alexander.** Die POD - eine Vorgehensweise zum qualitativen Gütenachweis zerstörungsfreier Prüfverfahren im Bauwesen am Beispiel des Impulsradars. DGZfP - Jahrestagung 2011. 2011.
- [4] **Sikora, Ryszard, et al., et al.** Comparison of selected weld defects extraction methods. Prezeglad elektrotechniczny. 83, 2007, Vol. 11, pp. 181-183.
- [5] **Souza, M. P.** Detection of Lack of Fusion weld defects by radiography. Review of Quantitative Nondestructive~Evaluation. 28, 2009, Vol. CP1096, pp. 1174--1181.
- [6] **Povagin, V. A. and Lavrentev, A. G.** Determination of Classification Characteristics of Magnetic Images for Sorting of Useful und Noise Signals during Flaw Detection in Gas Pipelines. Russian Journal of Nondestructive Testing. 2008, Vol. 11, pp. 787-791.
- [7] **Rose, Albert.** Television Pickup Tubes and the Problem of Vision. Advances in Electronics and electron Physics. 1, 1948, pp. 131-166.
- [8] **Burgess, Arthur E.** The Rose model, revisited: Brigham and Women's Hospital and Harvard Medical School, Boston Massachusetts 02115. Optical Society of America. 16, 1999, Vol. 3, pp. 633-646.
- [9] **Kanzler, Daniel and Müller, Christina.** How much information do we need? A reflection on the correct use of real defects in POD evaluations. AIP Conference Proceedings . 2016, 42th Annual Review of Progress in Quantitative Nondestructive Evaluation.
- [10] **Ewert, Uwe, et al., et al.** Image Quality in Digital Industrial Radiography. Material Evaluation. 70, 2012, Vol. 8, pp. 961-970.
- [11] **Kanzler, Daniel.** Zuverlässigkeitsanalyse von digitalen Radiographie-Systemen bei der Prüfung von realen Materialdefekten. Dissertation (Doctoral Thesis) . Berlin : s.n., 2015 (is submitted).
- [12] **Deutsch, Volker, et al., et al.** Die Ultraschallprüfung (UT). Wuppertal : Castell-Verlag GmbH, 2010.
- [13] **Ewert, Uwe, Zscherpel, Uwe and Jechow, Mirko.** Essential parameters and conditions for optimum image quality in digital radiology. 2012, Vol. 18th World Conference on Nondestructive Testing.