

Reliability Considerations of NDT by Probability of Detection (POD) Determination Using Ultrasound Phased Array – Results from a Project in Frame of the German Nuclear Safety Research Program

Jochen H. KURZ *, Sandra DUGAN **, Anne JÜNGERT **

* Fraunhofer Institute for Nondestructive Testing IZFP, Germany
Campus E3 1, 66123 Saarbrücken, Phone +49 681 9302 3880, Fax +49 681 9302 11 3880,
Email jochen.kurz@izfp.fraunhofer.de

** MPA University of Stuttgart, Stuttgart, Germany
Pfaffenwaldring 32, 70569 Stuttgart, Phone +49 711 685 66787, Fax +49 711 685 62761,
Email anne.juengert@mpa.uni-stuttgart.de

Abstract. Reliable assessment procedures are an important aspect of maintenance concepts. Non-destructive testing (NDT) methods are an essential part of a variety of maintenance plans. Fracture mechanical assessments require knowledge of flaw dimensions, loads and material parameters. NDT methods are able to acquire information on all of these areas. However, it has to be considered that the level of detail information depends on the case investigated and therefore on the applicable methods. Reliability aspects of NDT methods are of importance if quantitative information is required. Different design concepts e.g. the damage tolerance approach in aerospace already include reliability criteria of NDT methods applied in maintenance plans. NDT is also an essential part during construction and maintenance of nuclear power plants. In Germany, type and extent of inspection are specified in Safety Standards of the Nuclear Safety Standards Commission (KTA). Only certified inspections are allowed in the nuclear industry. The qualification of NDT is carried out in form of performance demonstrations of the inspection teams and the equipment, witnessed by an authorized inspector. The results of these tests are mainly statements regarding the detection capabilities of certain artificial flaws. In other countries, e.g. the U.S., additional blind tests on test blocks with hidden and unknown flaws may be required, in which a certain percentage of these flaws has to be detected. The knowledge of the probability of detection (POD) curves of specific flaws in specific testing conditions is often not present. This paper shows the results of a research project designed for POD determination of ultrasound phased array inspections of real and artificial cracks. The continuative objective of this project was to generate quantitative POD results. The distribution of the crack sizes of the specimens and the inspection planning is discussed, and results of the ultrasound inspections are presented. In the context of the results, the remaining uncertainty of the inspections has to be taken into consideration for failure analysis.

1. Introduction

A reliable prognosis of the condition and behavior of a structure or component is an important basis for an effective service life management. Non-destructive testing (NDT) methods are able to retrieve information on the geometry of flaws [1-3], on states of material degradation [4], on material parameters [5] and on stress related material states [6]. This information is required for damage assessment concepts, especially those related to the applied design concepts. NDT is also an essential part during construction and maintenance of nuclear power plants. In Germany, type and extent of inspection are specified by the Nuclear Safety Standards Commission (KTA). Only certified inspections are allowed in the nuclear industry. The qualification of NDT is carried out in form of performance demonstrations of the inspection teams and the equipment, witnessed by an authorized inspector. The results of these tests are mainly statements regarding the detection capabilities of certain artificial flaws. In other countries, e.g. the U.S., additional blind tests on test blocks with hidden and unknown flaws may be required, in which a certain percentage of these flaws has to be detected. The knowledge of the probability of detection (POD) curves of specific flaws in specific testing conditions is often not present.

A quantitative consideration of NDT methods in assessment procedures requires quantitative statements about the reliability of the applied NDT method. The determination of the probability of detection (POD) is one possibility to quantify the probability to detect a specific flaw with an NDT method. Each NDT measurement is influenced by a variety of factors: the stage of development and the quality of NDT equipment, the quality of written procedures, the skills and attitude of the operators, the geometry and material of the component, the environment in which the inspection takes place, the location, orientation and size of flaws, as well as the material state. Therefore, reliability information about an NDT method requires defining a probability of detection which can be described as: the proportion of cracks that will be detected in the total number of existing cracks in a component by an NDE technique when applied by qualified operators to a population of components in a defined working environment.

In general, publications dealing with POD applications can be found in different fields (medical sciences, biology, materials science, physics, engineering sciences, telecommunication etc.) [7-12]. From which of these fields the POD concept originated cannot be determined with certainty. However, some of the earliest studies of a variety of aspects concerning POD can be found in the development of radar technology during World War II.

NDT in aerospace is the field where the modern concepts of POD determination were developed and where POD determination is most widespread for the qualification of NDT procedures. The basic POD concepts are documented in detail in [8] and [13]. Brown [14] gives also an overview about the POD principle and the underlying philosophy. Furthermore, the compound of NDT and damage assessment by considering the POD information shows an increasing amount of concept developments and applications in fields where safety and reliability issues are concerned. Since the POD determination has to be performed for every entity of flaws and each NDT method which is applied, there are many different studies, also recent ones, especially for materials of practical relevance [15-19].

The work presented in this paper deals with the experimental POD determination from inspections of test blocks with real cracks, done within the framework of a German nuclear safety research project.

In this paper, the detection of real and artificial flaws of different sizes and different types using ultrasound is considered. The POD provides the probability for the detection of a certain flaw size. A high probability of detection does not exclude the occurrence of an isolated event that a larger crack is not detected. The POD delivers the realistic, statistical

assessment of the reliability for an NDT method. The knowledge of a POD does not increase or decrease the occurrence of such an exception. However, it allows a realistic, safety-oriented design and assessment since risks can be evaluated. With the objective of POD determination, ultrasonic phased array [20] and sampling phased array [21] inspections were carried out. Due to the increasing number of applications of the phased array technique in ultrasonic testing, the focus was put on the inspection using phased array transducers.

2. Specimens, procedure and inspections

Within the scope of the investigations, the Material Testing Institute (MPA) at the University of Stuttgart was responsible for manufacturing and supplying suitable test specimens for the ultrasonic investigations which are the basis for the calculation of the POD curves. The MPA holds a wide variety of test specimens and has broad experience with the fabrication of test specimen with realistic flaws. The selection of suitable test specimens is essential, because the flaws should be as realistic as possible and not limited to artificial flaws. The test specimens used in frame of the investigations presented here are listed in table 1.

Table 1. Test specimens for the ultrasonic investigations

		Number of test specimen	Total number of flaws	Realistic flaws	artificial flaws
Austenitic specimen	test	29	39	24	15
Dissimilar welds	metal	1	15	2	13
Cladded specimen		4	> 250	> 250	7

Different materials were chosen to cover different materials commonly used in nuclear facilities.

A picture of three different types of test specimens is shown in Fig 1. The austenitic test specimens are bar-shaped with a “wall thickness” of 32 mm, and contain different types of welds (different weld geometry). Some of the specimens are only base metal without weld. Most of the flaws are realistic flaws induced by inter-granular stress corrosion cracking or fatigue cracking [22].

Some of the test specimens contain EDM notches and some of them are completely intact. The flaws were documented using different reference tests like liquid penetrant testing, radiography and creation of etched macrographs.

The test specimen with a dissimilar metal weld consists of a circumferential pipe weld connecting an austenitic stainless steel pipe with a ferritic low alloy steel pipe, with austenitic cladding on the inner surface of the ferritic pipe section [23]. The test block has an outer diameter of 327 mm, a wall thickness of 29.5 mm, and a total length of 435 mm. The test block contains several EDM notches in the weld and in the base metal, as well as one axial and one circumferential crack.

The cladded test specimens consist of thick ferritic plates with an austenitic cladding, originally made to serve as pressure vessel test specimens, with wall thicknesses between 95 and 146 mm. The test specimens contain flaws that were created before, after and during the welding process [24]. A great number of flaws are natural underclad cracks, which were created by applying a wrong heat treatment after the welding process. This type of crack forms as inter-granular in the heat affected zone.



Fig. 1. Specimens used for the investigations

The number of flaws listed in table 1 is sufficient for a POD estimation. The distribution of flaw sizes is shown in Fig. 2.

It can be seen that the distribution is asymmetrical with the maximum at a flaw size of 2-3 mm. A flaw size distribution for POD estimations should cover the relevant flaw sizes and should contain a certain number of smaller flaws which are expected to be harder to detect. The size distribution of the available set of flaws meets these requirements.

Three teams of ultrasonic inspectors using two different types of measurement equipment examined the test specimens, using ultrasonic phased array (PA), phased array with transmitter-receiver probes (PATR) and ultrasonic sampling phased array (SPA) techniques. The details for the inspection are specified in a test procedure. The results of each team were compared to the actual flaws within the test specimens, known from reference tests. These results were then used as input for the POD determination. More details about the inspection procedure can be found in [25].

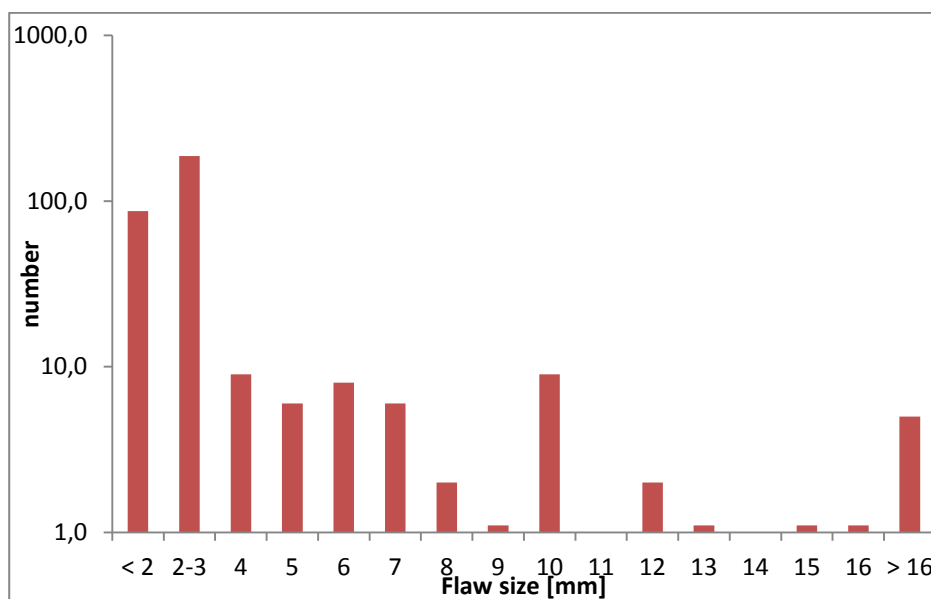


Fig. 2. Distribution of flaw sizes.

3. Results

3.1 Data Analysis

For an NDT inspection, a variety of factors can influence whether or not a flaw will be detected. The POD allows quantifying the reliability of NDT methods. The results can therefore have a direct influence on component design, maintenance and assessments. In principle, two related approaches to a probabilistic framework for the inspection of reliability data can be used [26]. One is based on the analysis of binary data, i.e. whether or not a flaw was found (hit/miss), and the other one is based on the correlation of signal response and flaw size ($\hat{a}$ vs. a). Analyzing binary data requires maximum likelihood regression to get the cumulative distribution function representing the chosen POD model. However, as is pointed out in [26], different results will be obtained when the two POD analysis methods are applied to the same data set. Furthermore, the maximum likelihood regression requires more data to achieve stable results [8, 26]. The main focus of the investigations presented here is laid on the $\hat{a}$ vs. a POD analysis.

POD determinations based on signal response data ($\hat{a}$ vs. a) have already been carried out for a variety of NDT methods. Therefore, the general applicability of ultrasound data for such a POD analysis was already proved. Ultrasonic data usually do not provide a direct correlation between UT signal and flaw size. In some cases, the reflected amplitude can be used as an estimate for the flaw size. A reasonably accurate determination of crack depths requires analysis techniques such as evaluation of crack tip signals. Crack lengths can be obtained from measurement of a registration length, using for example the 6 dB drop method or some other threshold value to define the end of the crack. In case of the Phased Array and SPA data, a registration length method was applied to measure the flaw size. For determination of flaw depths, flaw tip signals were used. This way, POD concepts can be applied to quantify the accuracy of the inspection in terms of a POD. If other than direct signal response data is used in the POD analysis, it has to be shown that this data is in accordance with the underlying POD models. In general, four conditions have to be fulfilled when measurement data is to be used for an " $\hat{a}$ vs. a " POD determination: a) linearity of the parameters, b) uniform variance, c) uncorrelated responses, and d) normal distribution of test flaw sizes. In our case, the determined flaw size represents the signal response data. Regarding an " $\hat{a}$ vs. a " POD analysis, the four conditions have to be valid for the data used. Otherwise the underlying POD model assumptions are invalid and incorrect results would be gained. The detailed proof that the four conditions for an " $\hat{a}$ vs. a " POD analysis are met by the ultrasound phased array data from the described inspections is given in [25]. There it is shown that the collected data which is based on the analysis of the registration length can be used for an " $\hat{a}$ vs. a " POD analysis.

A critical point in the POD analysis of signal response data is the definition of the decision threshold [8, 27]. This is the second major step in an " $\hat{a}$ vs. a " POD analysis. With the decision threshold the position of the POD curve relative to the flaw size axis is defined. In an analysis where the amplitude value of the signal is directly used, the noise level or the amplitude of a reference flaw can be used for selecting the decision threshold. In general, the decision threshold influences both the minimum size detected and the probability of false positive detections (probability of false alarm) [8]. Following the approach of [27], a criterion which considers noise and false positive detections in an indirect manner was chosen for the selection of the decision threshold [25, 28]. The difference between the true flaw size and the flaw size determined by from the UT data was taken and the standard deviation was calculated. This way, the accuracy of the measurements is considered and also the false positive indications which are an indication of the signal-to-noise ratio. Based on this approach, the inspection data was analyzed. The

results for artificial flaws and those for real cracks were considered separately because the real cracks show a different behavior. The relatively low number of artificial flaws does not lead to statistically verified results; however, the low scattering of the results of the artificial flaws allows stable POD calculations.

3.2 Artificial flaws

The austenitic specimens contain in total 15 artificial flaws. Fig. 3 shows the resulting POD curve for the analysis of the SPA inspection results of the tilted EDM notches (left) and the histogram of the deviation in flaw sizing (flaw depth). The term “tilted notches” refers to EDM notches with orientations between 10° and 50° relative to the surface normal. The depth of the tilted notches (2 mm, 4, mm, 10 mm) was determined with a relatively small deviation from their true depth. Fig. 3 (right) shows relative small values of a systematic overestimating the flaw depth. Therefore, the decision threshold was not calculated from the standard deviation of the difference between true flaw size and determined flaw size but a value of 2 mm was chosen. This value results from experiences gained throughout the project about realistic resolutions under optimal conditions for austenitic steels. It also corresponds to the reference flaw depth given in the Nuclear Safety Standards for in-service inspection.

The POD results show that a 90 % POD within a 95 % confidence interval ($a_{90/95}$) of 2 mm can be gained for artificial flaws in the austenitic specimens. However, it has to be pointed out here that this value will not be confirmed for real cracks.

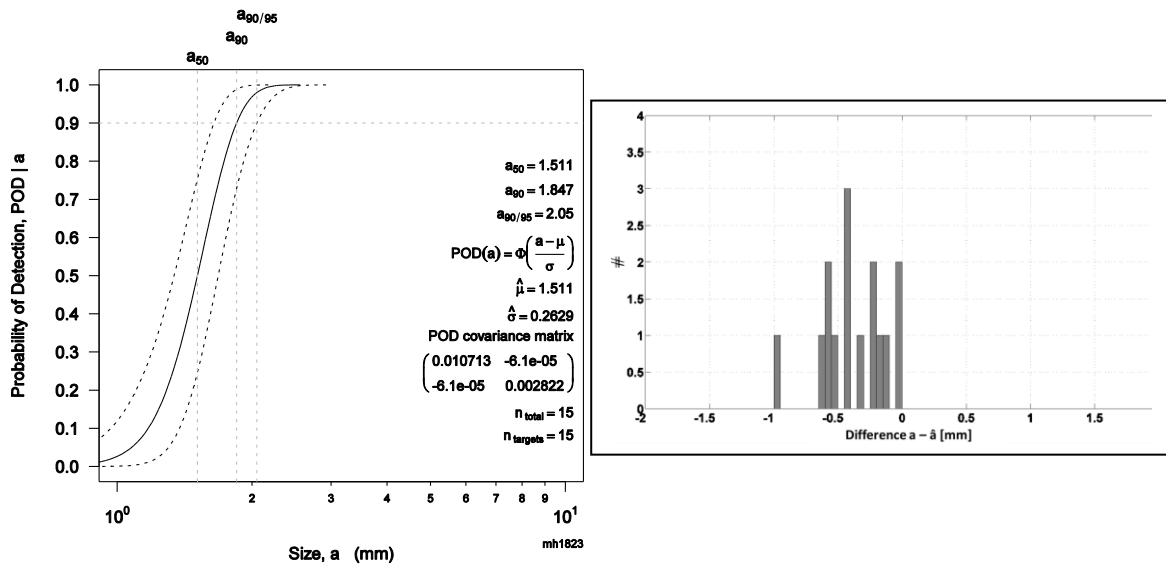


Fig. 3. Left: POD curve of the SPA inspection of the austenitic specimens with artificial flaws (tilted notches). The mh1823 software was used for POD calculations [28]. Right: histogram of the deviation in flaw sizing (flaw depth) of the SPA investigations.

Due to the material anisotropy in combination with the grain structure in the weld, austenitic steel welds and dissimilar metal welds are usually difficult to inspect using ultrasound techniques. Most of the flaws in the welded test specimens were located in the heat affected zone (HAZ). As a result, for artificial defects like notches with sharp geometric boundaries relatively good results can be obtained. The project results showed [29] that the SPA and PATR inspections lead to comparable results for the inspected specimens.

Dissimilar welds are usually even more difficult to inspect than austenitic steels, i.e. the $a_{90/95}$ value should be higher for dissimilar welds. This is confirmed by the results

shown in Fig. 4. The POD curve of the analysis of the PATR inspection results is based on the inspection of 8 notches longitudinally oriented relative to the weld direction (length and depth analyzed), and 1 circumferential as well as 1 axial crack (length analyzed). The decision threshold of 3.5 mm was determined as the standard deviation of the difference between true flaw size and determined flaw size (flaw depth and length). The resulting $a_{90/95}$ value is 5.1 mm, the confidence bounds indicate low scattering of the used data. As expected, the $a_{90/95}$ value for the notches of the dissimilar weld is higher than the one for the austenitic steel specimens.

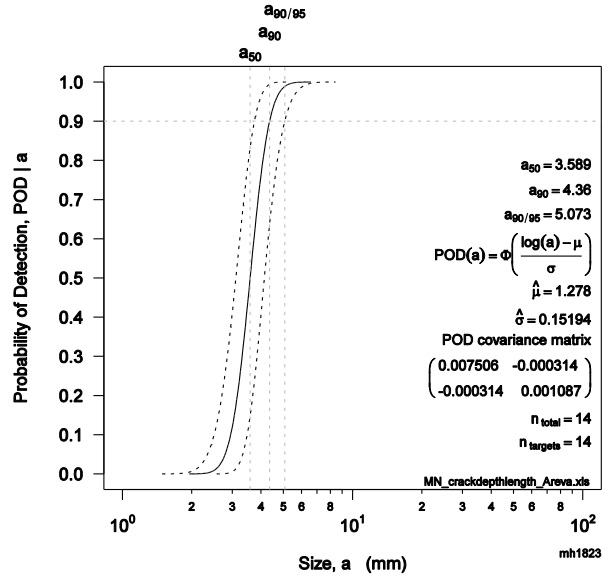


Fig. 4. POD curve of the results of the PATR inspections of the artificial flaws of the specimen with the dissimilar weld. The mh1823 software was used for POD calculations [28].

The cladded specimens contained 7 artificial flaws, for which flaw depth and length were evaluated. In addition, 3 large cracks were added to the dataset. Furthermore, two cladded specimens contained more than 200 under-clad cracks located in the ferritic part below the cladding. The results for these cracks are given in section 1.3. It has to be pointed out that the inspection techniques applied were not optimized for this type of flaw.

Fig. 5 shows the POD curves for the cladded specimens with artificial flaws obtain from the data of two inspection teams. The decision thresholds of 6.1 mm (PATR) and of 14.9 mm (PA Team II) were determined from the standard deviation of the difference between true flaw size and determined flaw size. The larger decision threshold value of the PA inspections was confirmed by the results of PA Team I, for which the standard deviation resulted in a decision threshold value of 14.5 mm. The reason for this difference in the decision values between PA and PATR is the significant difference in the accuracy with which the flaw length was determined. The resulting $a_{90/95}$ values are 9 mm (PATR) and 45 mm (PA Team II).

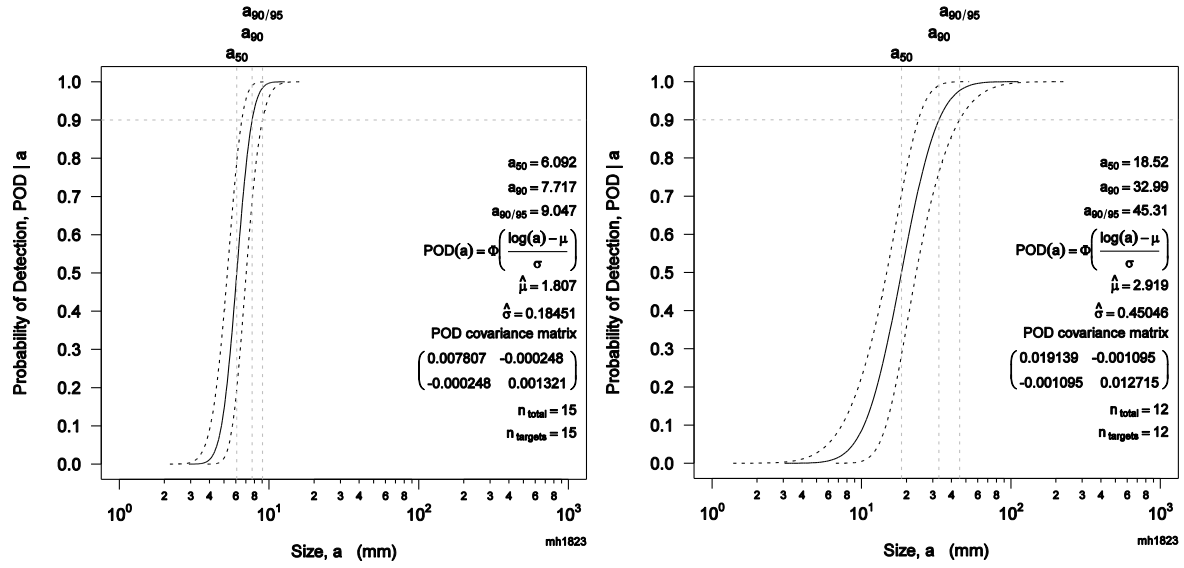


Fig. 5. POD curves of the inspection results of the artificial flaws of the cladded specimens. The mh1823 software was used for POD calculations [28]. Left: POD curve of the PATR inspection results. Right: POD curve of the PA inspection results of Team II.

3.3 Real Cracks in cladded specimens

The POD analysis of the inspection results of the artificial flaws in the cladded specimens already revealed that the $a_{90/95}$ value for the PATR results is higher than for the dissimilar weld specimen. The PA results show a significant inaccuracy in the determination of the flaw length. The analysis of the real cracks in the cladded specimens shows that a reliable detection is not possible while artificial flaws like notches are detected and sized correctly (Fig. 6). The number of detected flaws, as well as their positions and sizes differ much from the reference values. Therefore, a POD calculation was not possible for the real cracks of the cladded specimens.

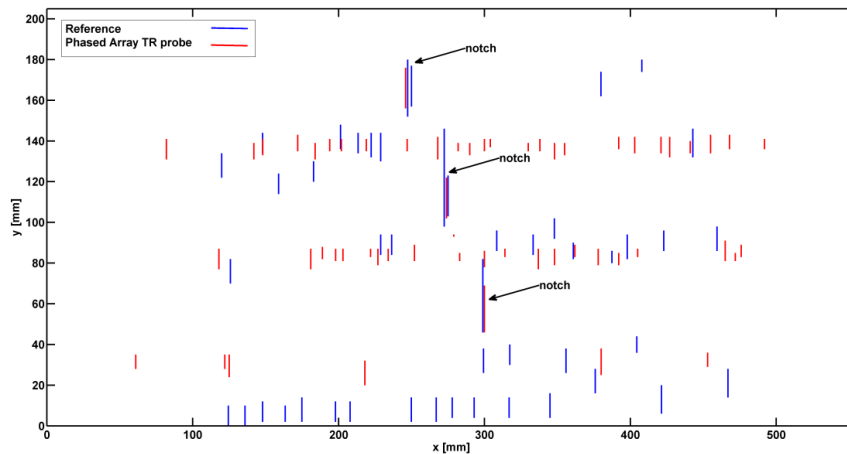


Fig. 6. Localization accuracy of real cracks in one cladded specimen also containing 3 notches. The actual crack positions are marked in blue and the results of the PATR inspection are marked in red.

3.4 Real Cracks in austenitic specimens

A crack depth determination for the cracks in the austenitic specimens was successfully done by all inspection teams. The 27 flaws considered here are realistic flaws (cracks). Fig. 7 shows the calculated POD curves from the results of the SPA and the PA (Team I) inspections. An overview of the results from the other teams can be found in [30]. The corresponding decision threshold value was calculated using the standard deviation of the difference between measured and actual size as described in section 3.1, resulting in values of 3.2 mm for the SPA data (Fig. 7, left) and 3.6 mm for the PA data (Fig. 7, right). .

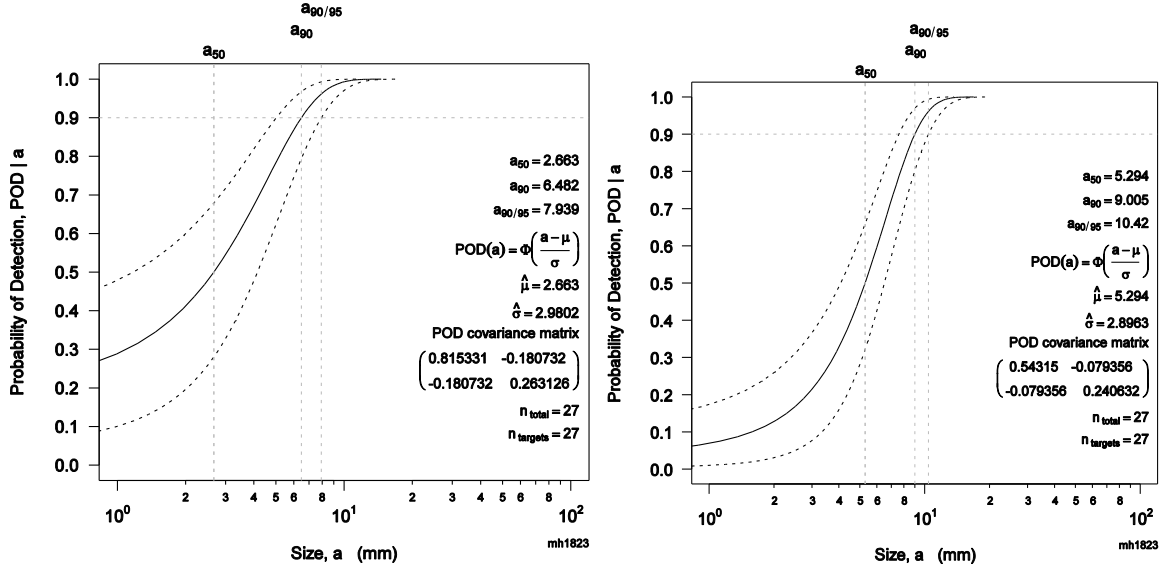


Fig. 7. POD curves of the inspection results of the real cracks of the austenitic specimens. The mh1823 software was used for POD calculations [28]. Left: POD curve of the SPA inspection results. Right: POD curve of the PA inspection results of Team I.

As expected, the accuracy in flaw sizing of the SPA technique is better compared to the results gained with the PA inspections. The phased array measurements lead to narrower confidence bounds, however, the a_{90/95} value of 10.4 mm is significantly higher than the 7.9 mm of the SPA POD. The a₅₀ value indicates the 50 % POD. This value represents the 50 % chance to detect a flaw. The values for a₅₀ are: 2.7 mm for SPA (Fig. 7, left) and 5.3 mm for PA (Fig. 7, right). These values correspond well with expert opinion regarding the detectability of cracks in anisotropic metals. Compared to the POD curve of the SPA results from the artificial flaws in the austenitic specimens (Fig. 3), the confidence bounds for the POD of the real cracks show more uncertainty for smaller crack sizes.

4. Conclusions

The experimental results presented in this paper were carried out in the frame of a German reactor safety research program. The main conclusions can be summarized as follows:

- NDT is an essential part during construction and maintenance of nuclear power plants. However, knowledge of the probability of detection (POD) curves of specific flaws in specific testing conditions using defined inspection methods is often not available.

- Quantifying NDT reliability with an experimental procedure needs appropriate specimens representing the required and realistic set of flaws.
- For many ultrasonic techniques, there is no direct correlation between signal amplitude and flaw size. The SPA technique in particular is an analysis technique which relies on determination of flaw depth and flaw length from crack tip signals and other characteristics in the ultrasonic images. In this case, e.g. flaw size has to be used as an equivalent parameter of the signal response for application of the “a vs. a” POD analysis principle. With no clear threshold for recording of flaws such as an amplitude level, the POD and the probability of false alarms will depend on the method used to determine whether an ultrasonic signal corresponds to an actual flaw.
- As expected, artificial flaws in the austenitic specimens were detected with a smaller $a_{90/95}$ value than for the dissimilar weld specimen.
- The cladded specimens show different results than expected. While artificial flaws could be detected and sized with good accuracy, in case of the real cracks, the number of detected flaws and their positions and sizes differ much from the reference values. Therefore, a POD determination was not possible for the real cracks.
- Artificial flaws and real cracks in the austenitic specimens could be analyzed. The resulting values for the PA inspections (Team I) correspond well with expert opinion regarding the detectability of cracks in anisotropic metals.
- The difference between the POD results for artificial flaws and real cracks is significant.

5. Acknowledgements

This paper shows results of a German reactor safety research project „Einbeziehung der Aussagefähigkeit zerstörungsfreier Prüfungen in probabilistische Versagensanalysen“, project no. 1501386. The financial support for this project by the German Federal Ministry of Economy and Technology (BMWi) is gratefully acknowledged.

References

- [1] Pudovikov S, Bulavinov A, Pinchuk R. Innovative Ultrasonic Testing (UT) of Nuclear Components by Sampling Phased Array with 3D Visualization of Inspection Results. In: Bièth M (Ed.): NDE in Relation to Structural Integrity for Nuclear and Pressurised Components, Berlin, Deutsche Gesellschaft für zerstörungsfreie Prüfung (DGZfP), 2011; Th.2.C.6 (10 p), (DGZfP-Berichtsbände 125-CD).
- [2] Dobmann G, Niese F, Willems H, Yashan A. Wall Thickness Measurement Sensor for Pipeline Inspection using EMAT Technology in Combination with Pulsed Eddy Current and MFL. Non-Destructive Testing Australia 2008; 45 (3): 84-87.
- [3] Maisl M. Nondestructive Testing, 2. Radiography. In: Ullmann's Encyclopedia of Industrial Chemistry. Weinheim, Wiley-VCH 2011: p 18.
- [4] Altpeter I, Tschuncky R, Hällen K, Dobmann G, Boller C, Smaga M, Sorich A, Eifler D. Early detection of damage in thermo-cyclically loaded austenitic materials. In: Rao BPC (Ed.), Jayakumar T (Ed.), Balasubramanian K (Ed.), Raj B (Ed.): Electromagnetic Nondestructive Evaluation (XV), ENDE 2011, Amsterdam, Washington, Tokyo, IOS Press (Studies in Applied Electromagnetics and Mechanics) 2012; 36: 130-139.
- [5] Dobmann G, Altpeter I, Wolter B. Materials Characterization a Challenge in NDT for Quality Management. In: UCTEA Chamber of Metallurgical Engineers: International Non-Destructive Testing Symposium and Exhibition. Istanbul 2008: 214-223.
- [6] Altpeter I, Dobmann G, Kröning M, Rabung M, Szielasko K. Micro-Magnetic Evaluation of Micro Residual Stresses of the IIInd and IIIrd Order. NDT&E International 2009; 42 (4): 283-290.

- [7] Schoefs F, Clement A, Nouy A. Assessment of ROC curves for inspection of random fields. *Structural Safety* 2009; 31: 409-419.
- [8]. Annis C. MIL-HDBK-1823A. Nondestructive Evaluation System Reliability Assessment. Department of Defense Handbook, Wright-Patterson AFB, USA 2009.
- [9] Georgiou GA. Probability of Detection (PoD) curves. Derivation, application and limitations, Jacobi Consulting Limited, Research Report 2006; 454.
- [10]. Kharin VV, Zwiers FW. On the ROC Score of Probability Forecasts. *Journal of Climate* 2003; 16: 4145-4150.
- [11] Sabelnikov A, Zhukov V, Kempf R. Probability of real-time detection versus probability of infection for aerosolized biowarfare agents: A model study. *Biosensors and Bioelectronics* 2006; 21 (11): 2070-2077.
- [12] Alvseike O, Skjerve E, Probability of detection of Salmonella using different analytical procedures, with emphasis on subspecies diarizonae serovar 61:k:1,5,(7) [S. IIIb 61:k:1,5,(7)]. *International Journal of Food Microbiology* 2000; 58 (1-2): 49-58.
- [13] Berens AP. NDE Reliability Data Analysis. In: *ASM Handbook, Nondestructive Evaluation and Quality Control*, 3rd Edition, ASM International USA 1994; 17: 689-701.
- [14] Brown J. POD test – It Takes More than a Statistician to do Probability of Detection. *Materials Evaluation* 2012;4: 421-426.
- [15] Lee JB, Park JH, Kim HD, Chung HS. Evaluation of ECT reliability for axial ODSCC in steam generator tubes. *International Journal of Pressure Vessels and Piping* 2010; 87: 46-51.
- [16] Spies M, Rieder H. Synthetic aperture focusing of ultrasonic inspection data to enhance the probability of detection of defects in strongly attenuating materials. *NDT&E International* 2010; 43: 425-431.
- [17] Carboni M, Cantini S. A “Model Assisted Probability of Detection” approach for ultrasonic inspection of railway axles. In: *South-African Institute for Non-destructive Testing: World Conference on Nondestructive Testing (18), WCNDT 2012*; paper 200.
- [18] Rummel WD, Matzkanin GA. *Nondestructive evaluation (NDE) capabilities data book*. 3rd Ed. Austin Texas: Texas Research Institute 1997.
- [19] Palanisami R, Morris CJ, Keener DM, Curran MN. On the Accuracy of A. C. Flux Leakage, Eddy Current, EMAT and Ultrasonic Methods of Measuring Surface connecting flaws in Seamless Steel Tubing. In: DE. Chimenti, DO. Thompson (Ed.), *Review of Progress in Quantitative Nondestructive Evaluation*, Plenum Press New York 1986; 5A: 215-223.
- [20] Moore, PO (Ed.), Workman GL (Ed.), Kishoni D (Ed.). *Ultrasonic Testing. Nondestructive Testing Handbook 7*, 3rd ed. American Society for Nondestructive Testing (ASNT) 2007.
- [21] Bulavinov A, Joneit D, Kröning M, Bernus L, Dalichow MH, Reddy KM. Sampling Phased Array A New Technique for Signal Processing and Ultrasonic Imaging. *Insight* 2006; 48 (9): 545-549.
- [22] Mletzko U, Zickler S. Verbesserte Bewertung der Aussagefähigkeit von zerstörungsfreien Prüfungen an Komponenten aus Nickelbasislegierungen oder austenitischen Werkstoffen. *BMU Vorhaben SR 2501, Technical Report (in German) 1.2*, Stuttgart 2007.
- [23] Mletzko U, Zickler S. Erste Bewertung der Aussagesicherheit von zerstörungsfreien Prüfungen an Mischschweißverbindungen auf Querrisse. *BMU Vorhaben SR 2501, Technical Report (in German) 1.1*, Stuttgart 2007.
- [24] Waidele H. Bewertung der Aussagefähigkeit von Ultraschall- und Wirbelstromprüfung austenitischer Plattierungen von Reaktordruckbehältern. *BMU Vorhaben SR2318, 2. Technical Report (in German) 8892 01 001*, Stuttgart 2003.
- [25] Kurz JH, Jüngert A, Dugan S, Dobmann G. Probability of Detection (POD) determination using ultrasound phased array for considering NDT in probabilistic damage assessments. In: *South-African Institute for Non-destructive Testing: World Conference on Nondestructive Testing (18), WCNDT 2012*; paper 329.
- [26] Berens AP. NDE Reliability Data Analysis. In: *ASM Handbook, Nondestructive Evaluation and Quality Control*, 3rd Edition, ASM International USA 1994; 17: 689-701.
- [27] Feistkorn S. Gütebewertung qualitativer Prüfaufgaben in der zerstörungsfreien Prüfung im Bauwesen am Beispiel des Impulsradars. *PhD Thesis (in German)*, TU Berlin 2011.
- [28] Annis C. Statistical best-practices for building Probability of Detection (POD) models. R package mh1823, Version 2.5.4.4 <http://StatisticalEngineering.com/mh1823/> 2010.
- [29] Kurz JH, Jüngert A, Dugan S. Consideration of non-destructive testing methods for probabilistic failure analysis. *Final Report (in German)*, Reactor Safety Research - Project No.: 1501386, 2013.
- [30] Kurz JH et al. Reliability considerations of NDT by probability of detection (POD) determination using ultrasound phased array. *Eng Fail Anal* (2013), <http://dx.doi.org/10.1016/j.engfailanal.2013.06.008>