A Plenary View on the Vigour of our NDE Reliability Models

Christina MÜLLER *, Marija BERTOVIC *, Ralf HOLSTEIN **, Daniel KANZLER *, Jorma PITKÄNEN ***, Ulf RONNETEG ****, Thomas HECKEL *
* BAM Federal Institute for Materials Research and Testing, Berlin, Germany
** DGZfP Ausbildung und Training GmbH, Berlin, Germany,
*** Posiva Oy, Eurajoki, Finland
**** SKB Swedish Nuclear Fuel and Waste Management Co., Oskarshamn, Sweden

Abstract. Using the Modular Reliability Model the three different main influencing elements, i.e. intrinsic capability (IC), application parameters (AP) and the human factors (HF), are, in the first instance, investigated separately. The intrinsic capability stands for the pure physical-technological process of the signal detection caused by the waves or the rays from a material defect in the presence of noise (driven by the material and the devices). This intrinsic capability is the upper bound of the possible reliability. Already when measuring this intrinsic capability for thick walled components the original one-parameter POD must be extended to a multi-parameter POD, where, in addition to the defect size, a number of additional physical parameters, such as the grain size distribution (or attenuation), defect depth, and angle or surface roughness, must be considered. For real life cycle assessments it is necessary to evaluate the signal response from real defects. The industrial application factors, e.g. coupling conditions, limited accessibility, heat and environmental vibrations, diminish the reliability. The amount of reduction can be determined quantitatively, if the underlying conditions are controlled. In case they are not controlled it is necessary to count for a fluctuation in the reliability in the field anyway. The third group of important influencing factors are the human factors, which do not only cover the individual performance capability of the inspectors but also the design of the working place,the procedure, the teamwork quality, interaction with systems, the organization, and finally, the relationship between the companies involved in the inspection process and to which extend the responsible parties are aware of it. When comparing an "ideal inspection" with a "real inspection" it is worth while to look how the existing practices, rules and standards support reliable testing and where the "delta" is. In the context of vigor, with respect to the industrial end user, it needs to be shown how the level of reliability of NDE, influenced by the different factors, has an impact on acceptance or rejection of safety critical parts.

Introduction and Current Status

The aim of this paper is not only to present the new paradigm in the field of NDE reliability but also to scrutinize the current vigour of our reliability models and methods for the final situation in the industrial application.

The insights from the first four workshops from the series of European-American Workshops on NDE reliability can be summarized as follows. The main achievement from the first workshop 1997 in Berlin was the conceptual model [1] (later referred to as

Modular Reliability Model [2]) in terms of the reliability formula which simply says that the total reliability of an NDE system is composed of the intrinsic capability, IC (which stands for the physical principle behind the defect indication and its technical realization as an upper bound), the application factors, AP (describing the realistic circumstances like UT coupling, limited access, noise of the surrounding …) and the human factors, HF, present in each application. While imperfect – e.g. the mutual interactions between the factors were not considered – the conceptual model helped to properly define the potential for performance optimization, and worked also as an assessment tool for the adequacy of open and blind trials. A main benefit from the Second European American Workshop on NDE Reliability, September 1999, Boulder, Colorado, USA [3], was the clear definitions of i) the NDE System as the procedure, equipment and personnel that are used in performing NDE inspection and ii) the NDE reliability as the degree that an NDT system is capable of achieving its purpose regarding detection, characterization and false calls. The main conclusion from 2002 from the third European American Workshop [4] was: We need to quantify the risk in NDE and demining! The fourth European-American Workshop in 2009 in Berlin again showed the progress in attempts to consider the reliability of NDE on a system level with the goal for integrated solutions in the industrial applications; the limitations of the original empirical methods were shown resorting to advanced and model assisted methods. Also, the first time an extra Human Factors session was launched with the result to recognize that the Human Factors are present in all stages of NDT activity while as targets of improvement the training, procedure, calibration, inspection and evaluation were assigned.

A view on the current international main activities, not claiming completeness but reflecting the historical development of dealing with NDE reliability is illustrated in Fig. 1 "The Reliability Flower Garden 2013":

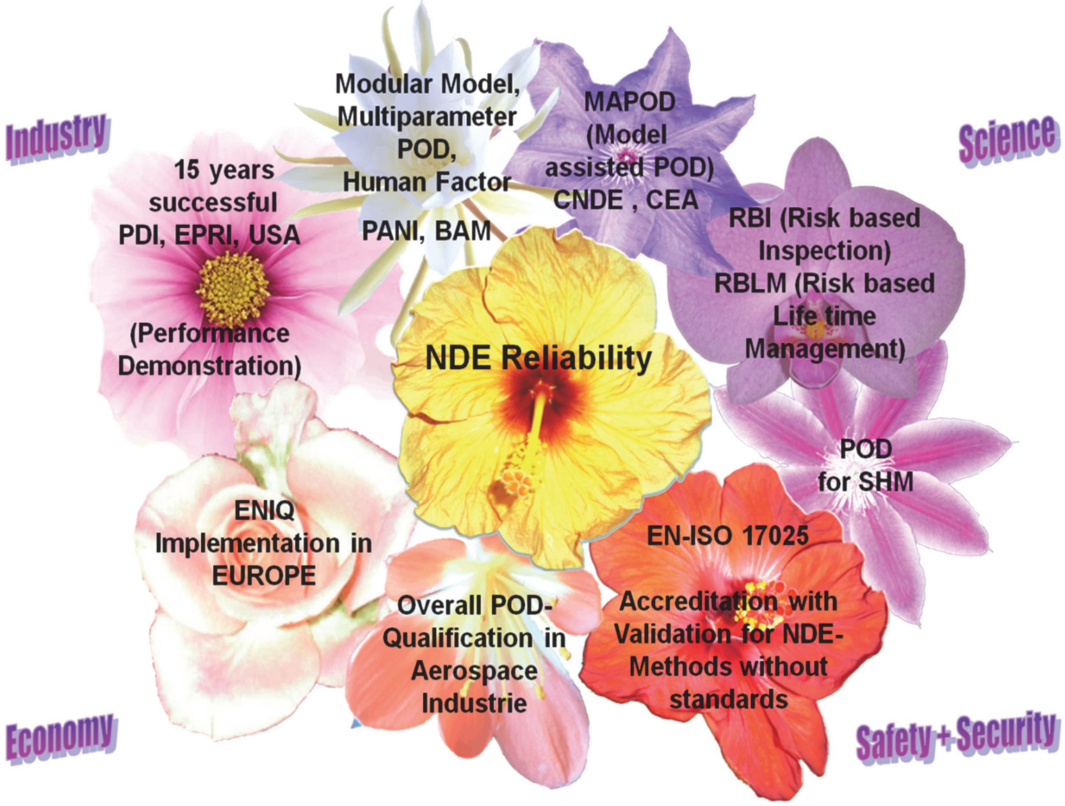

Figure 1: Today: The Reliability Flower Garden, 2013

where progress all over the world between the corner points of industrial demands and scientific progress on the one hand and between safety and security demands and economic conditions on the other hand are shown. What is the blooming item in each of the flowers? As indicated in the work from EPRI [5, 9] the PD (Performance Demonstration) is now well established in the US nuclear industry and with its procedure examination it drew nearer to the ENIQ (European Network of Inspection Qualification) approach, which was from the beginning composed of parameter assessment and open and blind trials. The National adaptations of the ENIQ methodology are also spread out over Western and Eastern Europe today [10]. The POD-qualification of NDE methods applied to airplane components is mandatory worldwide. For general NDE applied by accredited labs (EN-ISO 17025) the validation of quantitative and qualitative testing methods is required when no established standard exists. Also the new branch of SHM (Structural Health Monitoring) is going to evaluate the reliability via POD curves. RBI (Risk Based Inspection) and RBLM (Risk Based Life Time Management) is more and more applied to pressure equipment and nuclear facilities and the reliability of NDE is an important bridge between testing and the assessment of the remaining risk and life time. The youngest flower is the reliability assessment of Structural Health Monitoring with POD [11] and the most promising flower is the modelling assisted POD (MAPOD see Thompson, et. al. 2009 [12]) which helps to reduce costs for a high number of empirical experiments by substituting missing information by modelled results.

Last not least, for dealing with the human factors in an appropriate scientific way working psychological methods came on board firstly touched in investigations conducted in the USA (e.g. Harris, 1992 [13]), Finland (e.g. Norros & Kettunen, 1998 [14]), Sweden (e.g. Enkvist, 2003 [15]) and the UK (PANI program, McGrath, 2008 [16]) and later on investigations in demining (Gülle et.al, 2007 [17]) and NDE (Bertovic et.al. 2011 [18], Gaal et.al. 2009 [19]) by BAM. But even more insights from this field are to be taken into account when completing the Modular reliability model: the organizational context within which all the modules are embedded and also the interaction between different organizations as we learnt from Wilpert & Fahlbruch (1998, in [20]). The generalized Modular model (Müller et. al., 2009 [2]) as depicted in Figure 2 can also help to understand what issues are covered by the different "Flowers" of Figure 1.

Generally, when asked for an assessment of the reliability of an NDE system the following approach is helpful to consider: first, the actual level of safety demands have to be defined in order to adapt the thoroughness of investigation to the level of risk when the tested component would fail. Then, all the essential influencing parameters need to be listed and transferred to an appropriate design of experiments to determine the reliability in terms of a qualitative assessment for lowering the risk, or in terms of quantitative probability of detection (POD) or ROC (Receiver Operating Characteristics) curves.

We recommend here a new paradigm which considers the POD or reliability of the system as a matrix of input variables utilized for process optimization and not only for a final judgment. A subsequent element, and we believe advantage of this approach for the end user, is also to sample all single PODs to an integral "Volume POD" of a part including data fusion [21].

Among the influencing parameters, human factors are very important, but of course after the basics via IC are established. A systematic psychological approach should help to find out where the bottlenecks are, but as first priority, to provide the best possible working conditions for the inspectors.

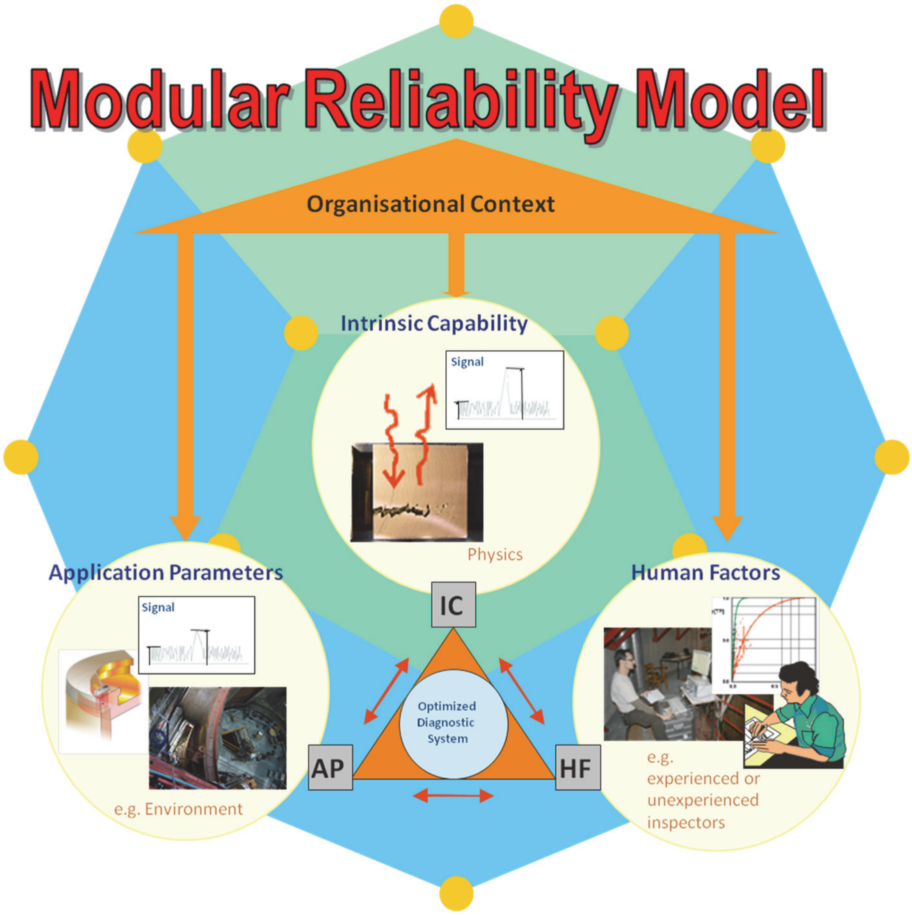

Figure 2: The Modular reliability model helps to understand and weight the different influences and which of them are covered in modelling, open or blind trials

Progress in Methods for Determining Intrinsic Capability, Application Parameters and Human Factors and Organisational Context towards the Vigour in the field

<u>Examples for new handling of Intrinsic Capability and Application Parameters</u>

Multiparameter POD

Figure 3 shows a typical result using the “â versus a” POD scheme as explained in [2, 8]. When applying this approach to industrial applications like copper canister components for radioactive waste (see e.g. [21], [22], [23], [24]) we have to expand this approach to real industrial conditions because defects in thick walled components need to be characterized by more parameters (like depth, angle, surface roughness) than just the size.

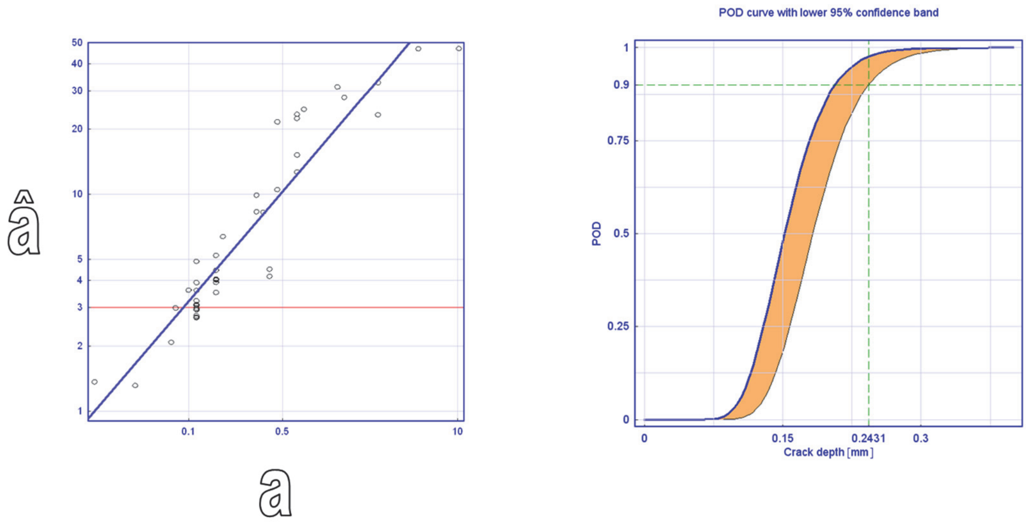

Figure 3: Probability of Detection (POD) – â vs. a approach

Specifically, we see a need for a multi parameter „a“ (depth, size, orientation, roughness…) and a data field „â“ (more than a maximum signal) for industrial applications with thick parts and complex defect shapes. This is shown for the amplitude dependence of PA UT echo field from FBHs (flat bottom bore holes) in Fig. 4.

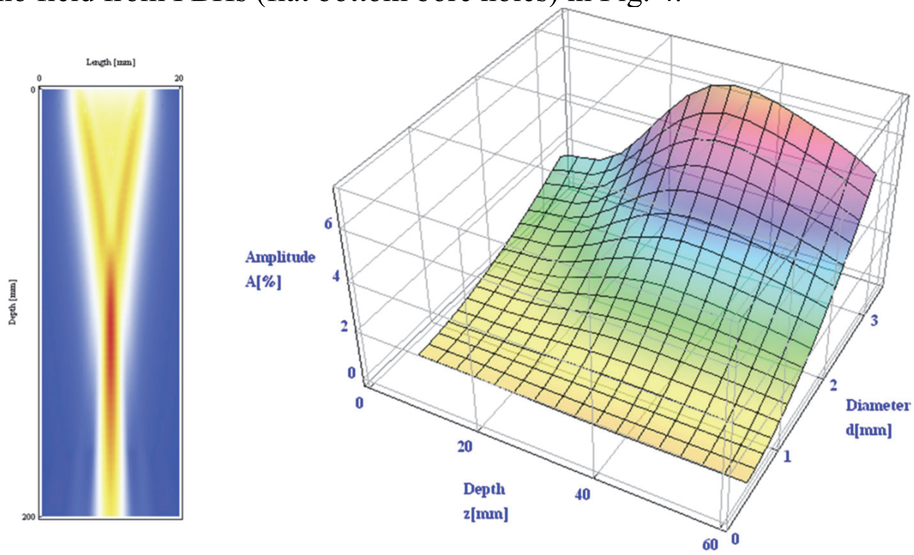

Figure 4: A = A(d, z, β, …)

To overcome this problem, a modelling assisted multi parameter methodology (Pavlovic et.al. 2008 [22], 2009 [21]) was developed and applied to the lid of the copper canister seen in Figure 5 and the different testing set up configurations in Fig. 6. The POD as a function of defect size (FBH diameter), defect depth and defect angle is presented in Figure 7(a), (b) and (c), respectively. The sharp decline of POD with increasing angle, for example, shows how important a comprehensive multi parameter consideration is.

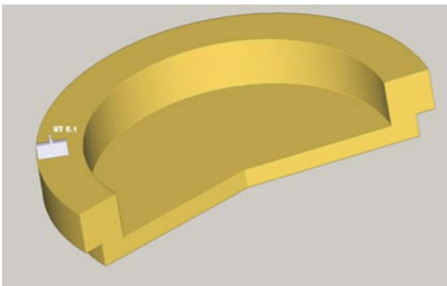

Figure 5: Copper canister for the deep deposit of highly radioactive nuclear waste (SKB/Posiva), canister parts planned for the NDT inspection: copper lid

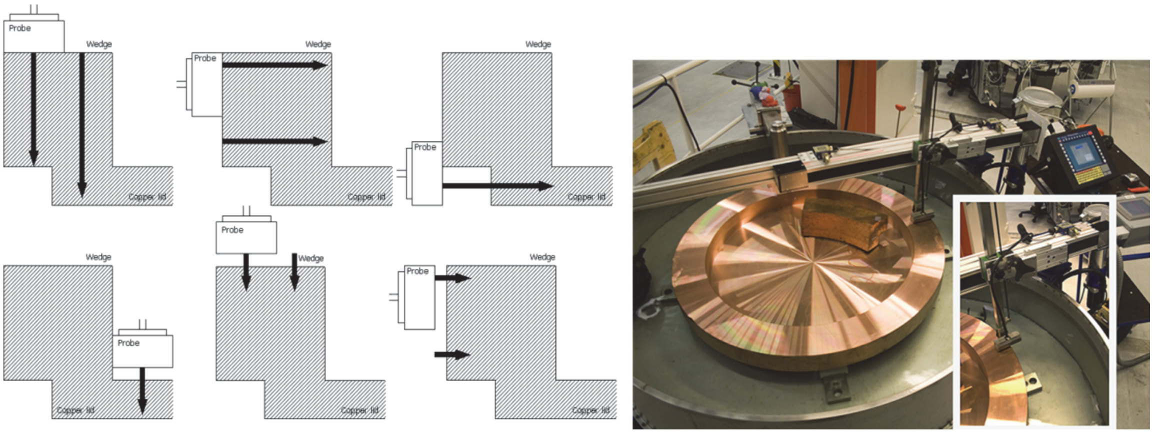

Figure 6: Multi-Parameter-POD; Copper canisters for deep deposit SKB/Posiva;
Canister parts planned for the NDT inspection: copper lid

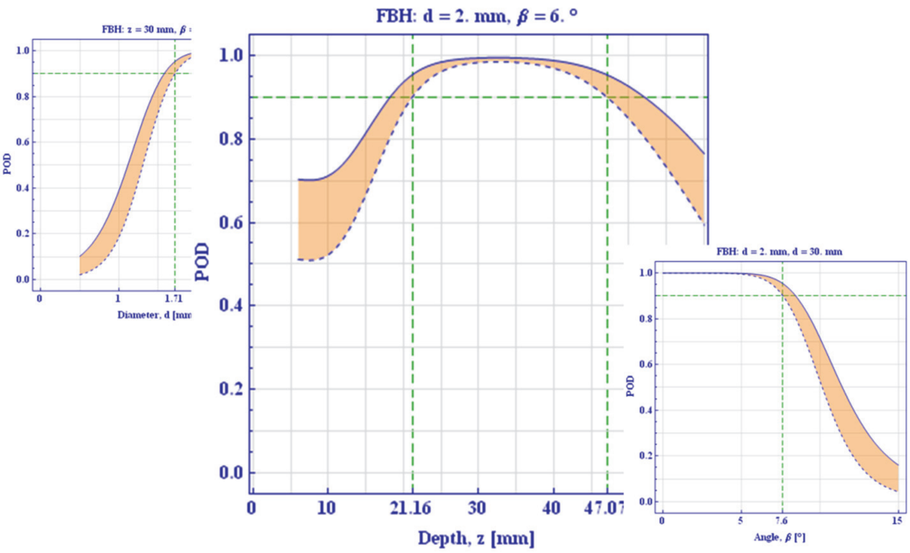

Figure 7: Multi-Parameter POD; POD as function of FBH-diameter (a), depth (b), and angle (c).
[Calculations from Mato Pavlovic]

Figure 8: Mechanized Inspection of the copper tube; UT equipment; Probe 128 elements, linear array, pitch 1 mm, width in passive direction 16 mm

Investigating the results of mechanized PA UT testing of the copper tube – as shown in Fig. 8 – the heterogeneous attenuation and relatively strong attenuation due to grain size distributions caused by the manufacturing process of the tubes has to be taken into account. The aim of the joint investigation of SKB, BAM, and Posiva[19]was to achieve a deeper insight into the effect of attenuation on the detectability of defects at larger depths in the copper tube. For this purpose, test samples with different level of attenuated material, containing flat bottom holes of various sizes, were manufactured. The samples were then inspected by the phased array ultrasonic technique developed by SKB. The results, including the amplitudes from the flat bottom holes and the multiple back wall echoes, as well as the frequency domain for both the surface and back wall echoes, were used as input for the POD calculations.

The data analysis showed that the attenuation and the low pass filtering of the ultrasonic signal, due to variations in the grain structure, varied along both the surface and the depth of the tubes. As a consequence, the difference between the second and the first back wall echoes were not sufficient to determine the local attenuation, which affects the signal response from the individual defect. These deviations were taken into account by a least square fit between calibrated modelled and measured data.

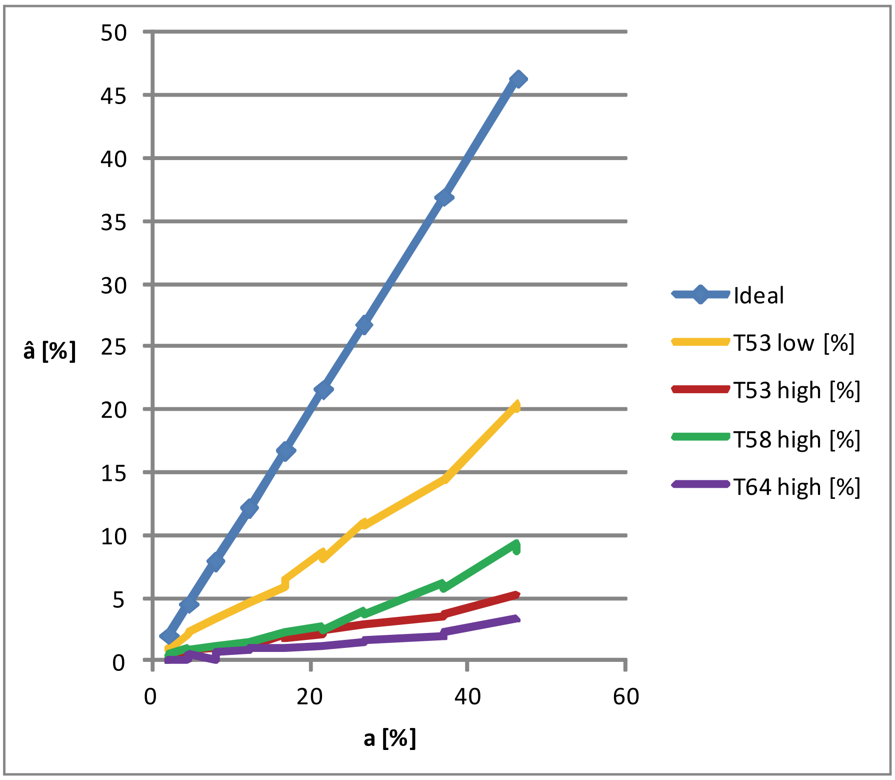

Figure 9: UT 11 measured amplitudes(containing attenuation) from FBH's versus modelled(without attenuation) echos (without α-renormalization)

Fig. 9 shows the results of the measured echo amplitudes as a function of the modelled ones without attenuation. The blue line represents the ideal curve where the modelled amplitude would be equal to the experiment.

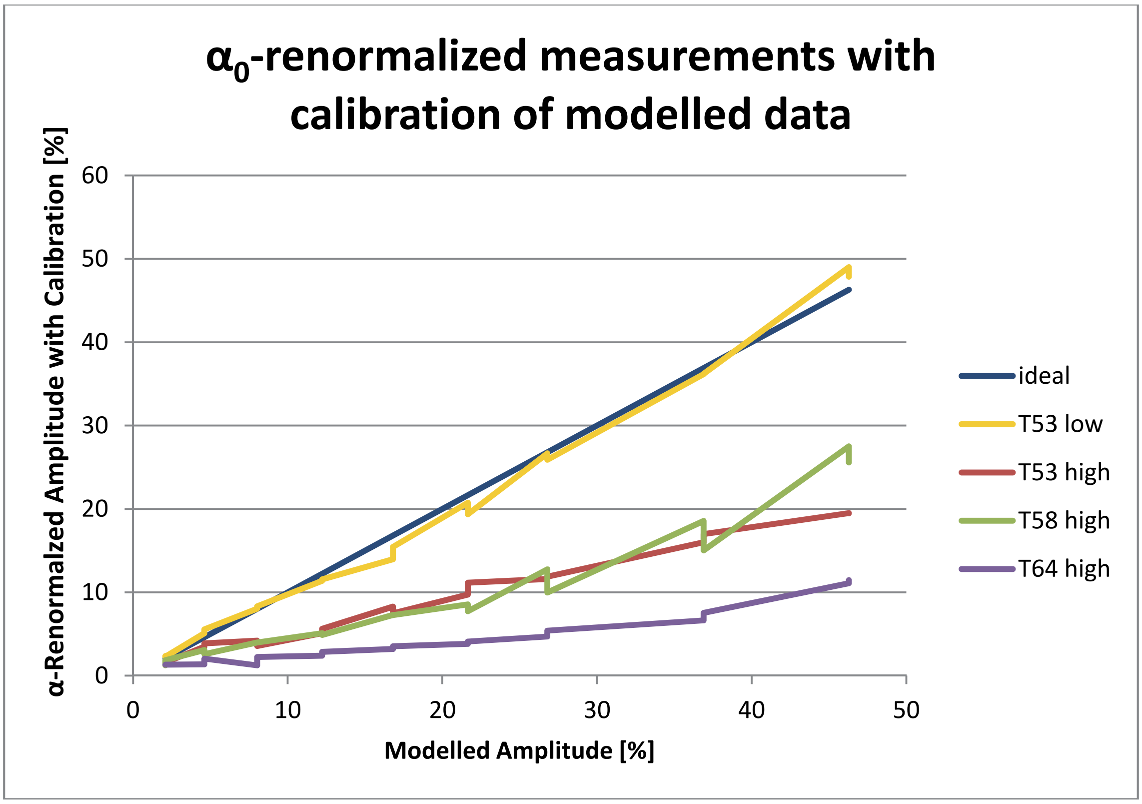

Figure 10: α renormalized measurements with calibration of modelled data.

Fig. 10 shows the relation between model and measurement after a re-normalization of the measurement values by the damping factor using the attenuation coefficient from the first and second back wall echo ratio. It is seen that there is a difference left due to the low pass filtering of the broad band signal in copper with high grain sizes. Those were determined from fitting the modelled curve to the measurements using this Δα.

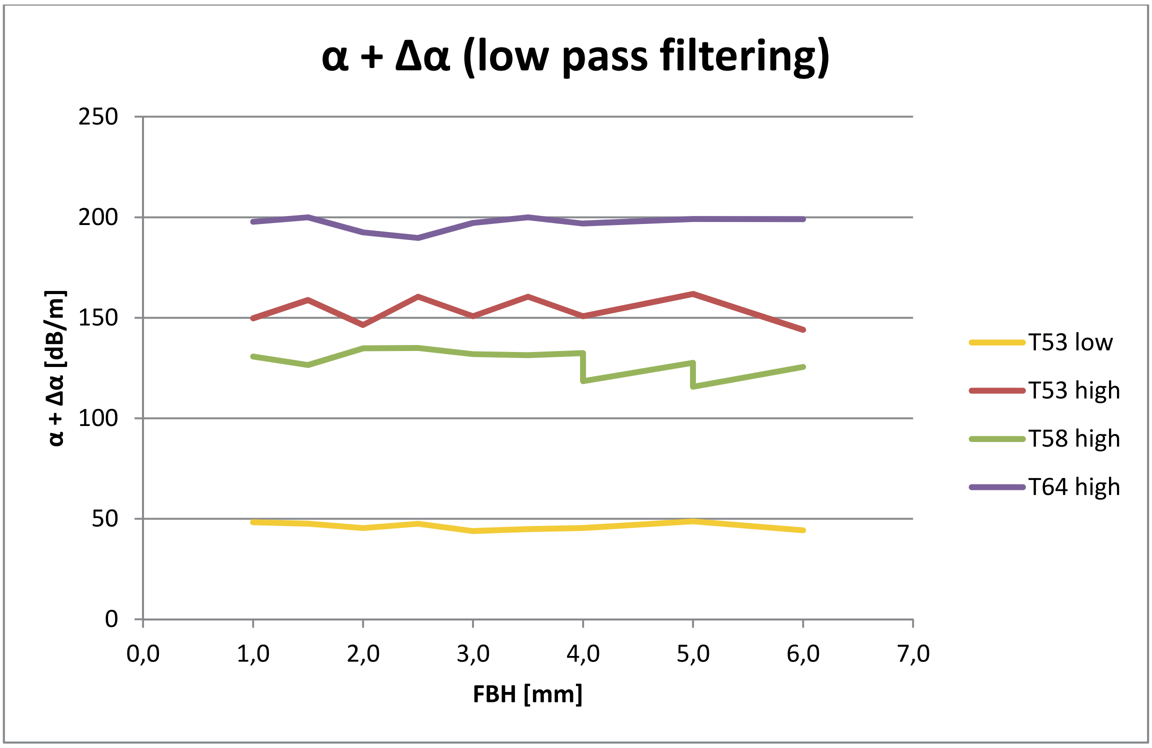

Figure 11: Final local α from BWE difference including corrections by beam spreading, transmission loss and low pass filtering.

The final α values for each tube material and each FBH surrounding are shown in figure 11. Finally, in figure 12 the POD curves as a function of the FBH diameter are seen for the different attenuation values. It is seen that the influence of attenuation is remarkable shifting the POD curves and confidence bounds to higher diameters the higher the α is. We can conclude that the attenuation is one of the most important influencing parameters here on the defect detectability.

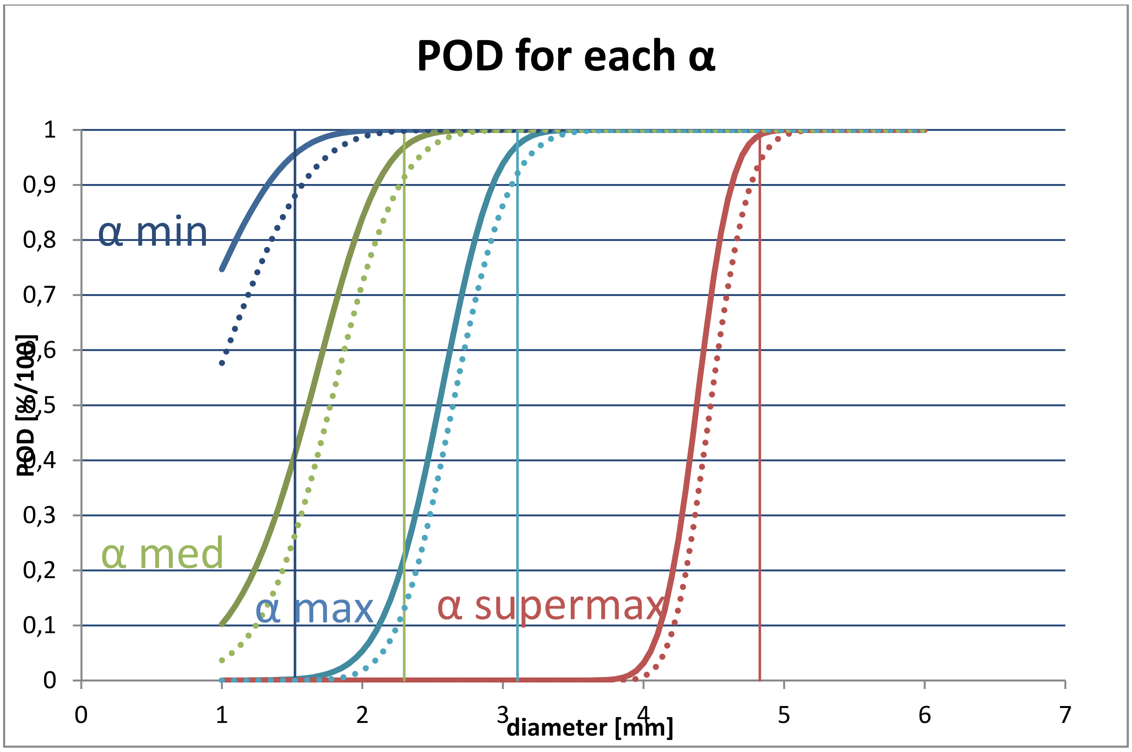

Figure 12. Summary of the calculated POD for α min, α med, α max and the postulated α supermax as a function of FBH diameter with indicated d90/95 values.

Table 1: Min, Med and Max and a fictive super-max attenuation values and corresponding detectable (d90/95) diameters and mean back wall echos; The POD and the detectable defect size change drastically with growingα due to grain size distribution in accordance with former investigations by T. Stepinsky[21].

	α [dB/m]	d90/95 [mm]	Mean BWE [%]
Minimum	46.1	1.5	43.43
Medium	141.4	2.2	12.17
Maximum	196.9	3.1	6.33
Super Maximum*	275	5.0	

It can be observed that the growing attenuation from about 50 dB/m up to about 200 dB/m is reflected in a highly diminished BWE-height, as well as in the growth of the detectable defect from less than 2 mm to more than 3 mm. This is still sufficient and the integrity of the tube is not affected. Utilizing the opportunity for scaling measured values and modelled amplitudes with arbitrary attenuation, a very high α of 275 dB/m was postulated and with this theoretical α a POD as a function of FBH diameter (d90/95) of 5 mm was calculated, The actual α variations are in the same order of magnitude, as mentioned by Stepinski et al. (2009), where the attenuation in the extruded copper parts varies from approx. 70 dB/m to approx 200 dB/m at 3.5 MHz. This shows that our mathematical-physical concept of $\alpha = \alpha 0 + \Delta\alpha$ is applicable for the current copper materials and the broad band phased array ultrasonic technique. Still, further confirmation by ultrasonic experiments and grain size measurements are desirable.

The influence of real flaws in real components

When the NDE engineer is asked to provide the probability of detection of critical defects as input for risk evaluation or life time management of real components, then it is necessary to provide information for real defects occurring in the component.

Usually, when starting with a new NDT-Problem, the method's capability it tested using well known artificial defects as shown in figure 13. To come then closer to the real problem artificially created real defects are used for the reliability tests. The "absolute" truth about the situation in the field can be found only under production, as indicated in figure 15, or in-service conditions. In this case it costs a lot of effort to identify and characterise the real defects with a statistically sufficient amount.

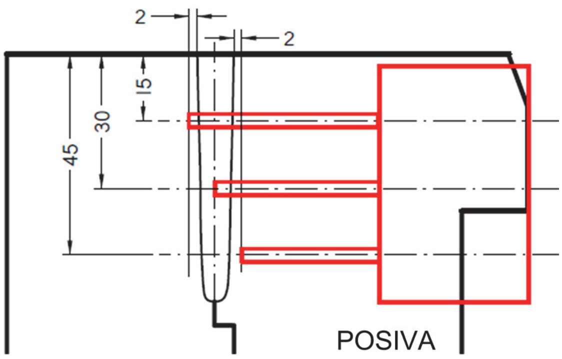

Figure 131: Artificial defects; ***The POD of artificial defects:*** evaluate the measurement process / the application factor => POD describes the limits with an ideal defect behavior.

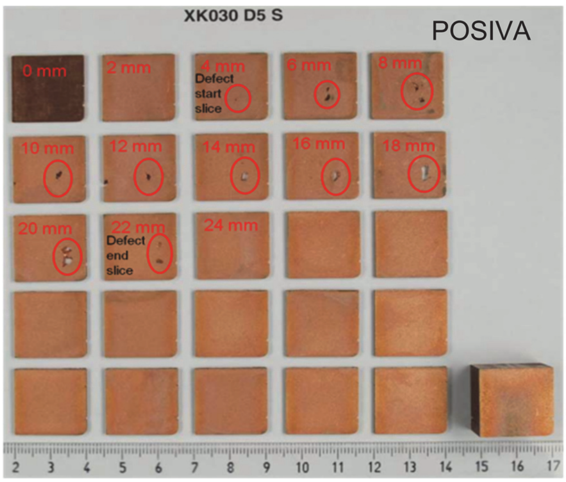

Figure 14: Realistic defects artificially induced; ***The POD of real defects:*** evaluate of the interaction with complex defects => necessary for an "Overall POD".

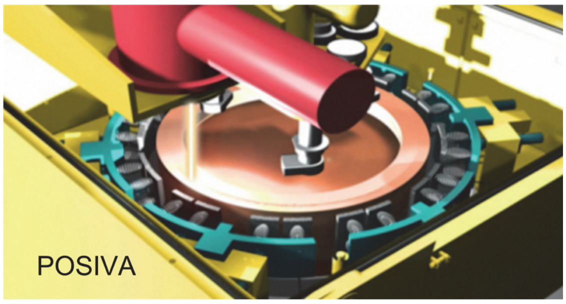

Figure 15: Production conditions

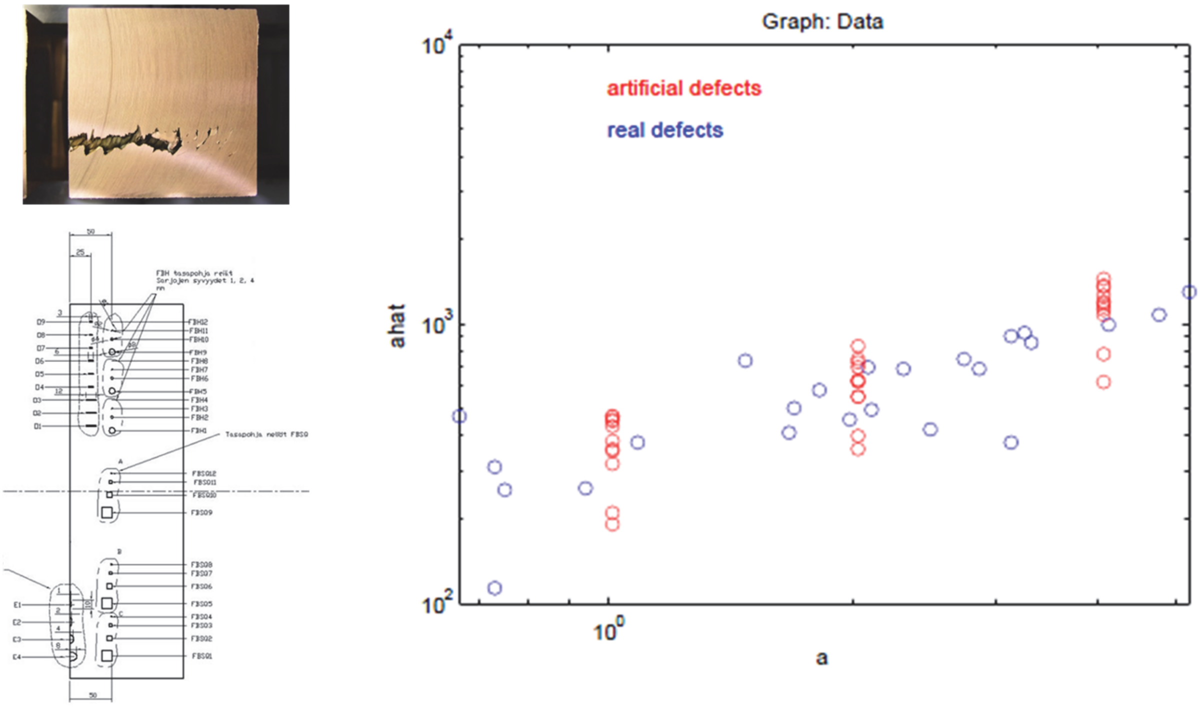

Figure 16: Data of artificial and real defects for RT for Cu-canister welds

Bayesian Approach for dealing with limited data from real defects

The Bayesian Approach provides a good solution to compute POD-curves in case of small amount of real defects without losing necessary information by incorporating data from artificial defects via a prior function while using information about the real defects in terms of a likelihood function. The posterior function composed in an appropriate way out of the latter contains needed information for the computation of POD-curves for real defects with an acceptable and sufficient amount of information, even for sparse amount of data as shown in the work of Kanzler, et.al. 2013 [23].

Examples for new Approaches for the improvement on Human Factors

When we started to co-operate with the field of working psychology we recognized that "Human Factors" means much more than a person's capability or the fact he might be tired.

According to the periods of safety research (see Figure 17), developed by Reason (1993; in [20] and expanded by Wilpert & Fahlbruch,1998; in [20]), the different influencing factors on the reliability of safety relevant work results range from technology over individual human errors to socio-technical and inter-organizational factors – and are well known in the human factors field.

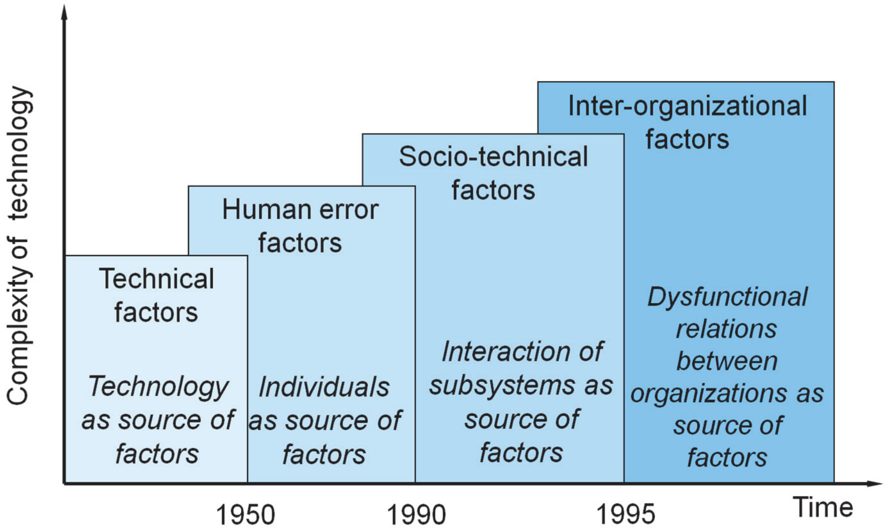

Figure 17: Periods of safety research

During the 4th EAW it was already reported about human factors' influences on manual ultrasonic inspection performance during in-service inspections (ISI) in nuclear power plants. A human factors model, which took into account up to date achievements in knowledge corresponding present safety demands, had been set up and a part of it tested (Figure 17 (see [18] for detailed results) and in general it could be concluded that for an improvement of the human performance attention needs to be drawn on:

- Demonstration task for training and performance confirmation
- Organization (good preparation)
- Written procedures and protocols
- Supervision

Further attempts to investigate human factors in the NDT field have focused on the mechanized inspection of the canister, to be used for the final disposal of nuclear waste in Sweden and Finland as was the subject of the multi-parameter POD as well mentioned above. A substantial number of factors that could affect the reliability of NDT methods have been identified and analyzed. Proposals for the compensation of varying human performance, according to the experts, include the implementation of human redundancy (known also as the 4-eye principle) or the semi-automation of the defect detection process. However, implementing human redundancy in critical tasks, such as defect identification, as well as using an automated aid (software) to help operators in decision making about the existence and size of defects, could lead to other kinds of problems, namely social loafing (excerpting less effort when working on tasks collectively as compared to working alone (Williams & Karau, 1991 [26])) and automation bias (uncritical reliance on the proper function of an automated system without recognizing its limitations and the possibilities of automation failure (a form of automation misuse) (Mosier & Skitka, 1996 [27])) that might affect the reliability of NDT in an undesired manner, when not taken into account adequately (as elaborated by Bertovic, et. al. 2012 [24]). Deeper understanding of the defect detection and sizing process, as well as an evaluation of the existing procedures, are a part of the continuing effort to understand the varying human performance, as well as to optimize current practices, if needed. In the following the main influencing factors on the human performance are summarized in fig. 17 and the conclusions from the Human Factors studies are summarized below.

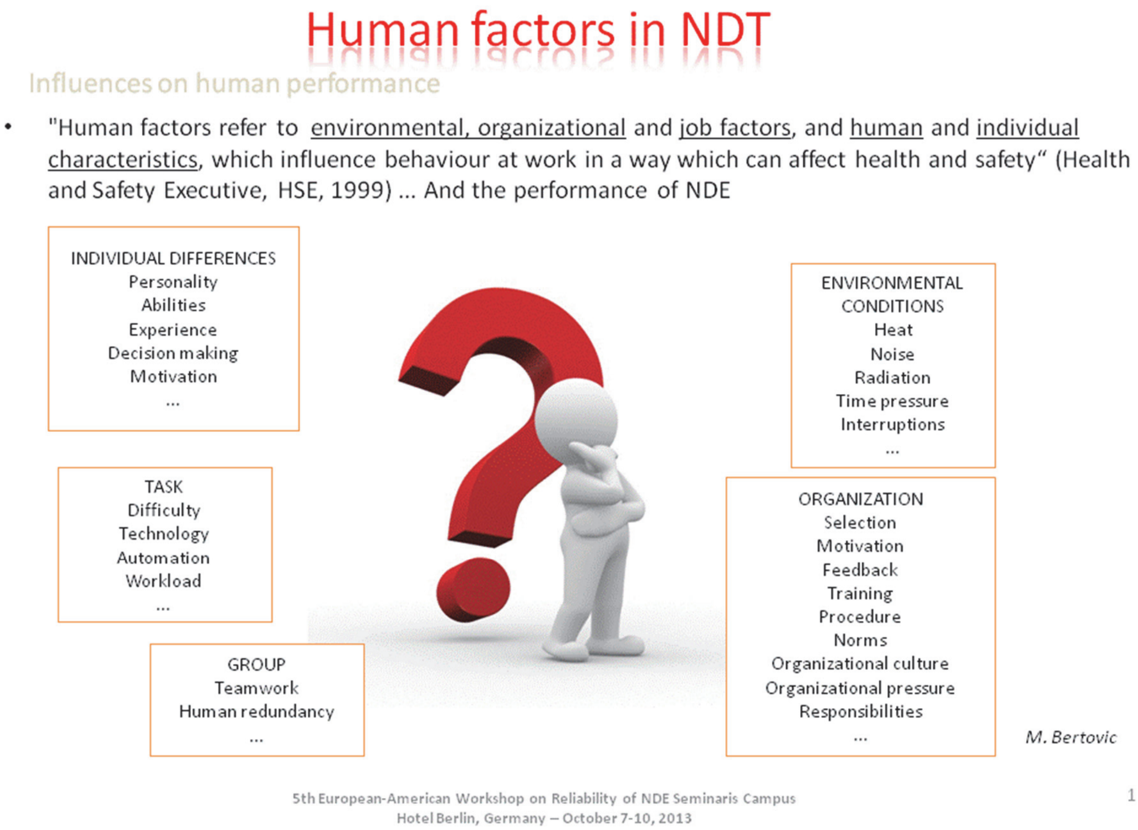

Fig. 18: Summary of influencing factors on human performance after Bertovic[24]

Main Conclusions from Human factors studies

WORKING CONDITIONS
- Time pressure has an effect on the quality of UT inspections
- Organizational context determines the way inspections are performed and therefore highly influences on the inspection quality

HUMAN REDUNDANCY
- Having two operators perform the same inspection might not necessarily lead to a more reliable inspection (problem of social loafing (relying on the other operator))
- It is important that operators conduct the task completely independently from each other and are aware of the effect of social loafing

AUTOMATION
- Implementing automation will not necessarily reduce human error
- As the role of the human operator changes - from doer into a controller of what the automation is doing - new error sources can arise
- Example: over trust in automation
- Operators should be trained to have appropriate trust in automation

INSPECTION PROCEDURE
- Relying only on norms and standards will not necessarily lead to a well written procedure
- The procedure has to contain all the relevant information and be user-friendly

Considering the flows of processes connected to industrial NDE applications Holstein et. al. [28] was able to identify three flows of processes immanent to NDE activities between a vendor and an end user: The business process, the information process and finally the technical NDT delivery process as depicted in Fig. 19. All three processes influence operator performance, which is unfortunately not well recognized neither to the technical experts nor to the economic experts. The amount of information requires *"Information Management"* in terms of a Communication Chain between *Customer, Contractor, Level III and the Operator* and should be carefully designed

. [

Figure 19: The three types of process flows incorporated in an NDE activity

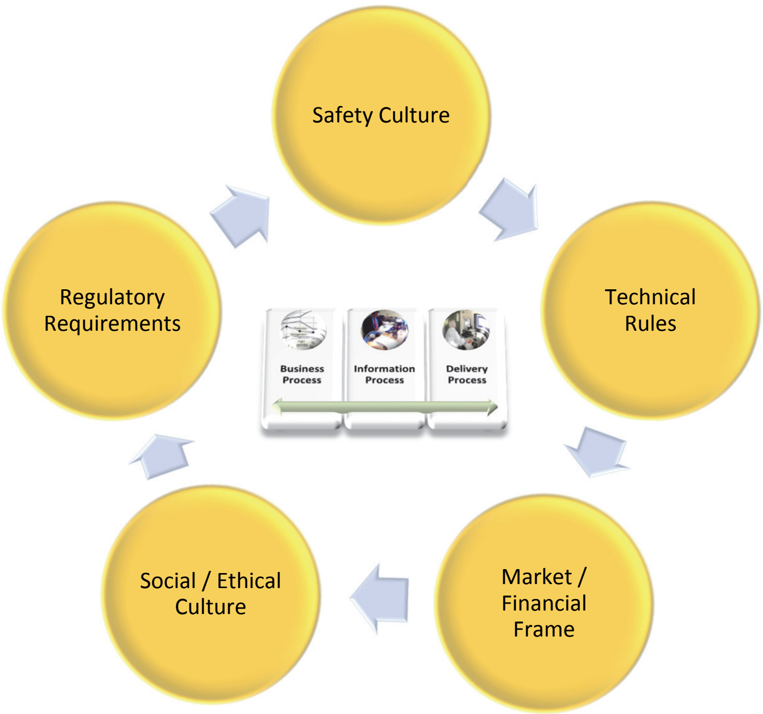

Figure 20: Process Environment

Furthermore, this communication chain is for each country or economic area embedded in a process environment composed of a safety culture, technical rules, market and financial frame, social/ethical culture and regulatory requirements. The question raises whether we need new tuning or controlling instruments here or simply an awareness to improve the reliability of our NDE processes?

Conclusion and Outlook

Dealing with the NDE reliability measurement effort the safety and economic requirements have to be taken into account and the qualitative and quantitative importance of each subtask (detection, false calls, characterisation, sizing) needs to be assessed. Multi-parameter POD and methods employing modelling assisted POD or data combination by Bayesian Approach appear promising to simultaneously fulfil safety and economic demands. The modular model helps to understand the impact of different sources of influences from parameters connected to physics, application or human factors – individual or organizational, respectively. Human factors need to be more deeply investigated by means of working psychology as ways to optimize the working conditions, training, inspection procedure and preparative actions for the operators.

To overcome the “delta” between the results from our current NDT reliability models and the real field situation the following issues have to be considered

- ROC and POD methods are adequate means measuring the reliability of NDE-systems for high safety demands BUT all influencing factors needs to be known and controlled
- POD should be used as an optimization tool
- A „Delta“ can occur when the part is not completely covered, actually influencing factors like attenuation are not well considered and from not knowing different response behaviour of artificial and real defects
- Importance of the Bayesian tool or other intelligent combination methods to make POD on real defects affordable
- Implementing human redundancy or automation to reduce human error can lead to problems of different nature – deeper understanding of these problems is needed, as well as consideration of these in the development of ways inspections should be performed
- A “Delta” can arise when we are not aware of the human factors or when think we are “on the safe side” with automation and redundant testing
- The inspection procedure needs not only to fulfil the requirements regarding the procedure content. "User centred procedure design" needs to be adopted in order to make the procedure more usable and ensure it is well understood
- A “Delta” can occur when the inspector has difficulties using the procedure and understanding its content on site
- Organizational context determines the way inspections are performed and therefore highly influences on the inspection quality.
- A “delta” can occur when the actual processes (business, information-, NDE-delivery) are not well recognized and tuned and the responsibilities of different parties are not clear enough
- The organizational context determines the way inspections are performed and therefore highly influences on the inspection quality.

References

[1] Nockemann, C. & Fortunko, C., "Summary of the workshop", Proceedings of the European-American Workshop on Reliability and Validation Methods on NDE, Berlin, Germany, June 18-20, 1997.

[2] Müller, C., Bertovic, M., Gaal, M., Heidt, H., Pavlovic, M., Rosenthal, M., Takahashi, K., Pitkänen, J. and Ronneteg, U., "Progress in Evaluating the Reliability of NDE. Systems – Paradigm Shift". Proceedings of the 4th European-American Workshop on Reliability of NDE, Berlin, June 24-26, 2009. DGZfP Proceedings BB 116-CD.

[3] ASNT Topical Conference Paper Summaries Book of the American-European Workshop on Non-destructive Inspection Reliability, September 21 24, 1999, NIST, Boulder, CO, USA, ISBN: 1 57117 041 3.

[4] Singh, R., Anderson, M. & Müller, C. „Summary & Consensus. Breakout session 5", Proceedings of the 3rd European-American Workshop on Reliability of NDE and Demining. DGZfP Proceedings BB-81-CD.

[5] ASME Boiler & Pressure Vessel Code, Section XI, Appendix VIII

[6] Nockemann, C., Heidt, H. & Thomsen, N., "Reliability in NDT: ROC study of radiographic weld inspections", NDT&E International, 24(5) 235-245, 1991.

[7] Metz, C.E., "Basic Principles of ROC Analysis", Seminars in Nuclear Medicine, 8(4), 1978.

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

[9] Dennis, M., Latiolais, C., Ammirato, F. & Heasler, P., "Development of POD and Flaw Sizing Uncertainty Distributions from NDE Qualification Data to Support Probabilistic Structural Integrity Assessment for Dissimilar Metal Welds, Proceedings of the 8th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurized Components, 29 September – 1 October 2010 – Berlin, Germany. DGZfP Proceedings BB 125-CD.

[10] Neuendorf, B., Seldis, T., & Gandossi, L., „Trends in Europe's NDE for the Nuclear Industry", Proceedings of the 8th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurized Components, 29 September – 1 October 2010 – Berlin, Germany. DGZfP Proceedings BB 125-CD.

[11] Aldrin, J.C., Medina, E.A., Lindgren, E.A., Buynak, C.F., Knopp, J.S., "Case Studies for Model assisted Probabilistic Reliability Assessment for Structural Health Monitoring Systems", QNDE 2011, AIP Conf. Proc. Vol. 13351589, 2011; doi: 10.1063/1.3592119.

[12] Thompson, R.B., Brasche, L. J., Forsyth, D., Lindgren, E., Swindell, P., Winfree, W. „Recent Advances in Model-Assisted Probability of Detection", Proceedings of the 4th European-American Workshop on Reliability of NDE, Berlin, June 24-26, 2009. DGZfP Proceedings BB 116-CD.

[13] Harris, D.H., "Effect of Decision Making on Ultrasonic Examination Performance. EPRI Report TR-100412, 1992.

[14] Norros, L. & Kettunen, J., "Analysis of NDT-inspectors working practices", Report STUK-YTO-TR 147, 1998.

[15] Enkvist, J., "A Human Factors Perspective on Non-Destructive Testing (NDT): Detection and Identification of Cracks". Doctoral dissertation submitted at the Stockholm University, Department of Psychology, 2003.

[16] McGrath, B. A., "Programme for the Assessment of NDT in Industry. PANI 3", Health and Safety Executive Research Report RR617, 2003.

[17] Gülle, D., Gaal, M., Bertović, M., Müller, C., Scharmach, M., and Pavlović, M. South-East Europe Interim Report Field Trial Croatia (Continuation of the ITEP-Project Systematic Test and Evaluation of Metal Detectors – STEMD). Berlin, March 2007. http://www.itep.ws/pdf/STEMD_Interim_Croatia_final.pdf.

[18] Bertovic, M., Gaal, M., Müller, C. & Fahlbruch, B., "Investigating Human Factors in Manual Ultrasonic Testing: Testing the Human Factor Model". Insight, 53(12), 673-676, 2011.

[19] Müller, C., Heckel, T., Brömel, P., Pavlovic, M., Ronneteg, U., „Attenuation dependent detectability at ultrasonic inspection of copper", SKB report within the licencing documentaion, 2014.

[20] Fahlbruch B, "Integrating Human Factors in Safety and Reliability Approaches", Proceedings of the 4th European-American Workshop on Reliability of NDE, Berlin, June 24-26, 2009. DGZfP Proceedings BB 116-CD.

[21] Stepinsky T, Engholm M, Olofsson T, 2009**.** Inspection of copper canisters for spent nuclear fuel by means of ultrasound, SKB TR-09-28. Svensk Kärnbränslehantering AB.

[22] Pavlovic M, Takahashi K, Müller C, Boehm R and Ronneteg U, "Reliability in Non-Destructive Testing (NDT) of the Canister Components. NDT Reliability – Final Report", SKB Technical report R-08-129, ISSN 1402-3091, Swedish Nuclear Fuel and Waste Management Co., 2008.

[23] Kanzler, D., Müller, C., Ewert U., Pitkänen, J. „Bayesian Approach for the Evaluation of the Reliability of Non-Destructive Testing Methods: Combination of Data from Artificial Defects and Real Defects", Proceedings of the 18th World Conference on Nondestructive Testing, Durban, South Africa, April 16-20, 2012.
[24] Bertovic, M., Fahlbruch, B., Müller, C., Pitkänen, J., Ronneteg, U., Gaal, M. Kanzler, K., Ewert, U., & Schombach, D. Human Factors Approach to the Acquisition and Evaluation of NDT Data, Proceedings of the 18th World Conference on Nondestructive Testing, Durban, South Africa, April 16-20, 2012.
[25] Fücsök F, Müller C and Scharmach M, "Reliability of Routine Radiographic Film Evaluation – an Extended ROC Study of the Human Factor", 8th ECNDT, Barcelona, 2002, CD-ROM.
[26] Williams, K. D. & Karau, S. J., "Social Loafing and Social Compensation: The Effects of Co-Worker Performance". Journal of Personality and Social Psychology, 61(4), 570-581, 1991.
[27] Mosier, K.L. & Skitka, L.J., "Human decision makers and automated decision aids: made for each other?" In: Parasuraman, R., Mouloua, M. (Eds.), "Automation and Human Performance: Theory and Application". Erlbaum, Mahwah, NJ, pp. 201–220, 1996.
[28] Holstein, R., Bertovic, M., Kanzler, D., Müller, C., "NDT Reliability in the Organizational Context of Service Inspection Companies", Proceedings of the 5th European-American Workshop on Reliability of NDE", Berlin, 2013, DGZfP Proceedings BB 144-CD