Summary of the "Open Space Technology Discussions" on Reliability of NDE

Marija Bertovic, Berlin, Germany, Pierre Calmon, Saclay, France, Luke Carter, Warrington, UK, Jay Fischer, San Antonio, USA, David Forsyth, Illinois, USA, Ralf Holstein, Berlin, Germany, Bernard McGrath, Warrington, UK, Christina Müller, Mato Pavlovic, Berlin, Germany, Ulf Ronneteg, Oskarshamn, Sweden, Ward Rummel, Littleton, USA, Frank Schubert, Dresden, Germany, Greg Selby, Charlotte, USA

Since the beginnings of the European American Workshops (EAW) in 1997, the aim was to gather the experts in NDE reliability and discuss burning topics with the aim of identifying crucial problems and suggesting ways to move forward. This was usually achieved during the so-called break-out sessions, in which predetermined topics were discussed. During the 5th EAW, held in Berlin in 2013, this approach was replaced by an Open Space Technology (OST) approach. The benefit of this approach is seen in the freedom of topic choices, i. e., the topics are not predetermined but rather arise at that moment in that space and by the participants' choice. The following topics arose: new reliability methods (Bayesian, MAPOD, ...), structural health monitoring, definition of requirements of NDE by customer versus provider, what value of POD is good enough?, human factors, manual versus automated inspection, and basic concepts of reliability of NDE. The participants were encouraged to walk from one session to another and openly express their opinions. These were in the end summarized by a chosen group of moderators and presented in this paper.

Article Information

Correspondence Address
Marija Bertovic
Fachbereich 8.3
Bundesanstalt für Materialforschung und –prüfung
Unter den Eichen 87, 12205 Berlin
E-mail: Marija.Bertovic@bam.de

Keywords
NDE reliability, POD, Bayesian approach, structural health monitoring, human factors, open space technology

The highlight of the 5th European American Workshop on Reliability of Non-destructive Evaluation (NDE) (5th EAW) was the "Open Space Technology"(OST) discussion. What was discussed in the previous 4 workshops in the so-called break-out sessions, in which a specific goal for each working group was defined beforehand, was this year replaced by an open space approach. The reason for this was not to restrict the participants to a certain number of topics, but rather to allow the topics to arise from the interests of the participants and the needs of both the researchers and the NDE practitioners. The deep aim of the 5th EAW was to shed light on the reasons, which create a "delta" between all the qualifications and reliability models and the real reliability in the field. It was our goal not to move on "old rails" when looking for an answer. This was achieved by setting up a wall of ideas. Each of the participants was asked to write down as many questions or topics as wanted, and pin them up on the wall throughout the duration of the entire workshop. Finally, when the participants gathered for the discussion, these ideas were classified into 6 main topics:

A. New reliability methods: multiparameter POD, MAPOD, Bayesian
B. Structural health monitoring
C. Applications in industry
D. Human factors
E. Integrated solutions and "delta"
F. Basic concepts of reliability of NDE

As in the spirit of the OST, in which the participants themselves determine the topics and decide how the discussion will go on, the participants from the topics C, D and E joined forces and discussed the topics as one group.

The groups separated and, under the lead of an expert, i. e., the group leader, the different issues were discussed. At the end of the day the groups were asked to present the content and the results of their work. Every group leader wrote a summary of the discussion. The purpose of this publication is to present the summaries of the held discussions and, therewith, share the main conclusions with the NDT community.

Open Space Technology

Open Space Technology, developed by Harrison Owen in the sixties, is one of the simplest methods to bring a big group of people together to discuss the topics that interest them [1, 2]. The choice of a topic, the duration, the participants, and the process of the discussion are in participants' hands.

The success of the discussion relies on the commitment and interest of everyone involved and follows simple principles and only one law.

The principles:

1. Whoever comes is the right person (people from different branches, people with different backgrounds, different positions, etc., are all invited to join)
2. Whenever it starts is the right time (spirit and creativity are not projectable)
3. Wherever it happens is the right place (in the discussion room or at lunch break, the discussion never stops)
4. Whatever happens is the only thing that could happen (the participants should let the discussion take its own flow, without trying to control it in any way)
5. When it's over, it's over. There are no rules how long a discussion should last. If it has reached its end, it is finished. Move on.

The so-called "law of two feet" relies on the premise that everyone carries responsibility for himself/herself. If the participant is not contributing or not learning from the discussion, then he/she should use his/her feet and move on to another place, which allows him/her to be productive.

Following these simple guidelines, the participants in the OST discussion held during the 5th EAW were instructed to do the to follow three items:

- Every participant is invited to attend any discussion topic he/she likes.
- Participants are free (and encouraged) to move from one session to another and openly discuss the topics with their peers.
- Each session will result in a short written summary of each topic discussed.
- The summary of each topic will be presented to the whole workshop audience at the end of the day.

As a result, in the following chapters the summaries of the discussed topics will be presented.

A: New Reliability Methods (David Forsyth & Pierre Calmon)

The planned purpose of Group A was a discussion of "New Reliability Methods: Multiparameter POD, MAPOD, Bayesian". The questions submitted by conference attendees during the week focused on "modeled" and Bayesian methods, and these were the subjects of the discussion.

The discussion began with Bayesian topics. The purpose of this is to combine data from multiple sources, including "modeled", experiments, and in-service inspections. Bayesian methods are one way to do this, and are flexible, robust, and well known. In particular, Bayesian inference is a method of inference in which Bayes' rule is used to update the probability estimate for a hypothesis as additional evidence is acquired [3].

The typical method proposed to do this is to choose one source of data as first step, defining the parameters of the POD estimate. Then we use the other sources of data to update the parameters of the POD estimate according to Bayesian methods. This leads to a few key choices for the responsible engineer or scientist:

1. What data shall serve as base in the first step?
2. How will the new data be weighted?
3. Will the new data be checked for consistency with the original assumptions?

The importance of using Bayesian methods instead of simply pooling the different data was discussed. The advantages of Bayesian methods compared to pooling are that the data can be weighted differently, and different distributions can be combined.

A key outcome of the discussion was the suggestion that an application guide including working problems would be a great benefit. This is a potential subject for the next workshop to undertake.

The next main topic of discussion was the use of models for POD estimation. There is general acceptance of this proposition, assuming that the models have been validated over the range of inputs associated with the specific problem at hand. By providing distributions on inputs to models, and execution of models in Monte Carlo type methods, a distribution of outputs can be found. One topic of discussion is how this modeled distribution of results relates to the confidence bounds on POD determined using empirical methods. It was noted that Bayesian methods may provide a solution to combine them.

Differently, the direct modeling of the complete inspection, it was noted that modeled or simulated data could be used to modify inspection data to get a better POD estimate in cases where it is difficult to make fully representative specimens. Noise due to different materials or lots or non-relevant signals due to extraneous geometries are examples for this.

In summary, this was a lively discussion marked more by agreements than disagreements. There are many open questions remaining regarding how to execute POD estimation using the different possible data sources and combinations thereof. Further research combined with sharing of experience through development of benchmark problems, guidebooks, worked examples, etc., will be beneficial for the community.

B: Structural Health Monitoring (Jay Fisher)

The main discussion areas were:

- The ways in which structural health monitoring (SHM) is different from conventional non-destructive evaluation (NDE)
- Reliability issues of SHM systems
- Issues concerning determination of reliability of SHM

Determination of POD for SHM will be different from that for NDE. In particular, for SHM there may not be a fixed threshold. One reason is that the noise level can vary over time, as well as other conditions. In NDE reliability determination, it is assumed that there is a fixed noise level, sensor performance level, and environmental conditions. These conditions cannot be counted on for realistic SHM. In particular, during the course of monitoring a structure, the mechanical loading and temperature can change. It was also noted that the performance of an SHM system will vary based on the relative location of the SHM sensors and flaws.

The possibility of making SHM fit into the maintenance system was brought up. It might even be possible to use information from SHM sensors to provide operational parameters to supplement existing system diagnostics. For example, an ultrasonic system can be used to obtain temperature information as well as defect information, based on changes in ultrasonic wave velocity.

Issues related to sensor reliability were discussed in detail. In order to have practical SHM, there should be a way to check the sensor performance. The sensors should be designed regarding reliability. Ways to include this would be to

- design with some level of redundancy,
- minimize the number of single-point failure modes
- include sensor self-check capability
- design a calibration reflector in the structure with the SHM system
- plan hardware and software support of installed SHM systems for a long term

In practice, one difference between SHM and NDE system implementation decisions is that in addition to sensor reliability, the SHM system will also have to balance weight and space against performance and cost.

It was decided that defining POD for various inspection methods are philosophically similar, but the method to introduce and locate flaws will be different.

The role of human factors was discussed. It was agreed upon that human interpretation is still important.

There should be demonstrations of SHM installations that cover small areas but have high payback, as it is a way to gradually move industry to accept SHM.

What needs to be done?

- Better ways to characterize reliability for SHM are needed.
- We need to determine how to define POD when conditions of the structure may change.
- Methods to qualify SHM systems should be developed.
- General methods to determine where to place sensors should be developed.

CDE: Definition of Requirements of NDE by Customer versus Provider (Christina Müller)

The group discussed the question of the definition of requirements on NDE from customer's (end user) point of view in contrast to the requirements seen from the NDE research and provider's point of view. A gap was discovered between both positions and means to overcome at gap were discussed.

What was discussed? Specifically the representatives of end users (e. g., power plants, aero-space industry, etc.) listed the aspects they are urged on when defining the requirements on NDE as:

- Safety margins
- Codes & regulations
- 100 % detection (of critical defects)
- Dependent on the number of defects and their characterization
- Difference between structural integrity and NDE
- POD including the human factors would be desirable for risk assessment

The NDE community (provider and research organizations) define:

- Detection capability regarding to parameters
- One step before the actual requirements
- Application of procedures and standards

What needs to be done? The gap between both positions could be found in an agreement on reasonable targets. This requires an adequate information management between both parties.

CDE: What Value of POD is Good Enough? (Luke Carter)

The general consensus of the group on the topic of what value of POD is good enough was that the necessary POD should be decided jointly between the structural integrity and NDE groups. Other comments and suggestions include:

- A higher POD value brings higher cost in demonstration and implementation.
- Refinement of procedures until required POD is achieved (Recursive POD).

CDE: Human Factors (Marija Bertovic & Luke Carter)

Another discussion topic that received a lot of attention during the workshop was the human factors (HF) topic. It appears that there is a lot of interest in the influence of human factors on the reliability of NDE, at least from the research community. However, there is a gap in the communication between the utilities and the service providers, causing problems in the transfer of knowledge and, hence, posing a difficulty to implement the findings in the field.

The first issue discussed was the shared responsibility of dealing with human factors between the NDE service provider and the customer. It was concluded that the human factors do not always receive the appropriate attention in the field, and that the communication between the customer and NDE provider is a big part of the problem. The customer often wants low cost and does not always take human factors into consideration. The inspection vendor, in contrast, tries to fulfil the requirements of the customer, sometimes at the expense of human factors and NDE reliability, in general. The question is: is the customer even aware of the human factors' influence on the POD and whose responsibility is to inform them? The existence of standards, codes, requirements for training and experience were acknowledged, as was the existence of HF tools (computer–based training, pre- and post-debriefing, self-checking, etc.), especially in the nuclear field, well established in the USA and followed slowly by the Europeans.

The Second issue discussed was of a more practical nature, i. e., how to keep an inspector vigilant, even when he never sees a flaw. Examples from different industries were given, e. g., railway axle inspection, where in the course of 5 years only 87 defects have been found. This was recognized as an important issue in different application branches and the need for solutions was expressed.

The Third topic was related to the practical needs for future research in the field. It seems there is a communication gap between what is known about human behaviour under difficult working conditions in psychology and what is known by the engineers.

What needs to be done? A broad discussion and raising of awareness is needed between the customer and the NDE service providers. Considering that the customer is no NDE expert, it is up to the NDE community to spread the word. However, the customer should take over the responsibility and consider the benefits and the costs of considering NDE reliability and human factors.

Re-qualification to refresh operator knowledge and skills for detection and characterization, recording geometric indications to keep operators occupied during their task, yearly practice on test components with realistic defects, engagement with staff in non-outage time (developing procedures, performing open trials) were some of the suggested solutions for the vigilance problem.

Relying only on the experience and qualification of the inspecting personnel has shown to be a flawed approach. Influences of quality and recency of experience, as well as loss of skills between certifications, have to be acknowledged. Further topics of interest for the future research include:

- The influences of working conditions on the inspectors (shift work, night shifts, vigilance, fatigue, heat, noise, etc.), or the transfer of that knowledge from the social sciences to the engineers and the inspecting personnel
- The extent of effects of social loafing on human redundancy and whether human redundancy is at all an appropriate approach
- Improvement of the interaction between the customer and the service provider
- Deeper insights into the human factors' issues in mechanized NDE, especially during data evaluation
- HF aspects of the management and the organization, in general

CDE: Manual vs. Automated Inspection (Ulf Ronneteg)

The group discussed among a number of topics the question – What advantages and disadvantages regarding the reliability can be foreseen in automated, mechanized, and manual inspection?

First of all it was determined that there is a clear difference between automated and mechanized inspections. In the automated inspections, data collection, data evaluation and even decisions are performed automatically while in the mechanized inspections normally the data evaluation and the decision making are done by the human inspector.

In general, it was concluded that the reliability is higher in mechanized inspection compared to manual inspection, but it was stated that "the mechanical system is not always as good as the best manual inspector but at least better than the average manual inspector".

Regarding mechanized inspections some advantages and some disadvantages were discussed and the most important are summarized below:

+ Usually better data coverage and especially better traceability of the data and coverage
+ The data evaluation can be done by different experts in the future (good if new design criteria arise).
- The data evaluator can miss the feeling by doing the inspection, i. e., if something unexpected occurs during inspection that the operator does not document or if the actual inspected part differs from the drawings. This might not be captured by the data evaluator.

There were also some discussions about fully automated inspections and the general remark about this is that the need for correct input data is extremely important. Otherwise the inspection could be very unreliable if something is wrongly set.

The final conclusion of the group is that independent of the level of automation, the human factor is to some extent always involved, i. e., human factors always exist with the risk of human errors.

What needs to be done? In general, it needs to be spread out that there are still needs to investigate the reliability of more or less automated inspection systems. Especially as it has been concluded that the errors that might occur often are different from errors in manual inspection. One of the most important tasks in mechanized inspection is the evaluation of data and its often interpretation of complex signals and/or images. There is a need for high quality procedures.

The general conclusion for what needs to be done is recommendations/guidelines for written procedures regarding format, structure, and content. It should be emphasized what should be considered in writing and reviewing of the procedure, involving the user in the process (user-centred approach). It should also be pointed out that the inspector should be informed about the scope of the inspection and that the procedure is well understood and accepted. Finally, this information needs to be available, well known, and accepted by the whole NDE "world". Valuable attempts into this direction have already been made within the scope of the HSE's Programme for the Assessment of NDT in Industry (PANI) project [4], as well within the scope of the cooperation between BAM, SKB, and Posiva on the NDT reliability project [5].

F: Basic Concepts of Reliability of NDE (Mato Pavlovic)

The group discussed, the basic concepts of reliability of NDE. The number of questions posted on the wall, as well as the number of participants in the discussion, suggests that even if the NDE reliability has gone a long way, there is still a need to discuss the fundamental principles that form the basis for all advanced concepts.

The first topic discussed was the confidence bands on the POD curve. It was stressed that the width of the confidence band comes from the experiment only and is determined by the sample size. Only the lower confidence band is of interest for determining the a90/95. Also, it has been noted that there is a difference in calculation of the confidence bands when calculation is performed as described by Berens [6] or according to MIL 1823 [7]. Consequently, this results in different a90/95s. It has been concluded that calculations are based on different mathematical models, hence the difference.

The Next topic was the applicability of POD curves on different kinds of defects. It was concluded that NDE reliability engineers have to work closely together with stress and materials engineers.

When considering the possibility of comparing POD data from different sensors, it was concluded that it is indeed possible and that POD can be used as a tool to compare performances of different sensors.

The question was followed by the discussion about how is risk included in the POD analysis. As described in MIL 83444 [8], first the crack growth rate is calculated. The time of the failure is derived from this rate and then the inspection interval is set at the half of this time.

The role of the threshold and who decides on it has also been discussed. As it is known, every system can have 100 % (low threshold) or 0 % (high threshold) detection rate, but what is changing is the false calls rate. It has been pointed out that even though stress engineers would like to have the threshold as low as possible, they should not be mating these decision.

What needs to be done? From the wide spectrum of questions one can conclude that there is a need for learning materials in the form of a text book and possibly organization of courses and tutorials on the topic of reliability of NDE. Interested parties would have the opportunity to learn in an easy and understandable way the elementary concept of reliability. These materials will give a jump start for those entering the field of NDE. On the other hand, as can be summarized from informal talks with the participants of the workshop, the materials would be a great help also for those that already have experience in NDE, but failed to understand properly some of the basic concepts of reliability.

References

1 R. Seliger: Einführung in Großgruppen-Methoden. Carl-Auer Verlag, Heidelberg (2011)
2 O. Harrison: Open Space Technology: Ein Leitfaden für die Praxis, Schäffer-Poeschel, Verlag (2011)
3 http://en.wikipedia.org/wiki/Bayesian_inference (accessed on April 11 2014)
4 B. McGrath: Programme for the Assessment of NDT in Industry, PANI 3, Report No. RR617, Health and Safety Executive (2008)
5 M. Bertovic, U. Ronneteg: User-centred approach to the development of NDT instructions, SKB Report R-14-06, Svensk Kärnbränslehantering AB, in review
6 A. P. Behrens: NDE Reliability Data Analysis, Metals Handbook Volume 17, 9th Edition: Non-Destructive Evaluation and Quality Control, ASM International (1988)
7 Department of Defense: MIL-HDBK-1823 Non-destructive evaluation system reliability assessment (2009)
8 USAF Damage Tolerant Design Handbook: Guidelines for the Analysis and Design of Damage Tolerant Aircraft Structures, MIL-A-83444 (1979)

Bibliography

DOI 10.3139/120.110604
Materials Testing
56 (2014) 7-8, pp. 602-606

ISSN 0025-5300

The Authors of This Contribution

Marija Bertovic, born 1981, received her diploma degree in psychology at the University of Rijeka in Croatia in 2006. Since then she has been working at the BAM Federal Institute for Materials Research and Testing in Berlin, Germany in division "Radiological Methods" in the area of "Reliability of Non-Destructive Evaluation Systems" as research scientist. The focus of her work has been on human and organizational factors affecting the reliability of NDT inspections. She is a doctoral candidate at the Berlin Institute of Technology (TU Berlin).

Luke Carter, born 1985, received his degree in physics at the University of Liverpool, UK. He joined Serco in 2007 (taken over by AMEC in 2012) and has been working as NDT consultant, providing assistance to a range of organizations in the UK and across the world. He is specialised in the application and development of ultrasonic techniques, specifically phased arrays.

Dr. Jay L. Fisher, born 1950, received his degree (ScD) in applied plasma physics at the Massachusetts Institute of Technology, Department of Nuclear Engineering, USA. He is responsible for program development and project management for non-destructive evaluation (NDE) systems and development of advanced electromagnetic non-destructive evaluation techniques. He holds a current position of a program director at the Sensor Systems and NDE Technology Department, Southwest Research Institute (SwRI).

Ralf Holstein, born 1961, received his degree in communication engineering from the Technical University of Dresden, Germany in 1993. In the same year he started working for the German Society for Non-destructive Testing (DGZfP) in Berlin, Germany, first as head of the personnel certification body, and then as head of the training department. Since February 2003 he is managing director of the DGZfP Education and Training Ltd. Since 2009 he is external PhD student at the VŠB – Technical University of Ostrava, Czech Republic. He is also elected chairmen of the DIN standardization committee "Qualification and Certification of Non-Destructive Testing Personnel".

Bernard McGrath, born 1959, is business Manager of AMEC's NDT Business at Birchwood Park, Warrington, UK. He is responsible for a team of NDT development engineers who undertake NDT design, development and capability projects. The business has a strong modeling capability to complement the experimental facilities. Bernard's project managed the PANI reliability projects on behalf of the UK's HSE and oversaw the production of the HSE's NDT Guidance Documents. Recently, he provided the inspection input to a framework for ageing management.

Dr. rer. nat. Christina Mueller, born 1953, studied physics at the Technical University in Dresden, Germany and finished her PhD in the field of theoretical solid state physics in 1982. After a phase of teaching and research at the university she worked in the practical development of material testing procedures in the electronic industry and in the development of ultrasonic probes for nuclear power plants in the former GDR. In 1988 she started to work at BAM Federal Institute for Materials Research and Testing in the division for NDT in the field of signal processing, modeling, and 3D data reconstruction. In international co-operation she developed basic principles for reliability assessment of NDT and is since 1992 head of a research group "Reliability of NDT".

Mato Pavlovic, born 1971, received his diploma degree in mechanical engineering at University of Zagreb, Croatia in 2009. After graduation he worked as research and development engineer on the development of flight simulator for Soko Z.I. in Zagreb, Croatia. From 2005 to 2012 he worked at the BAM Federal Institute for Material Research and Testing in Berlin, Germany on the application of existing and development of new NDT reliability models. Since 2013 he is working as scientific advisor for reliability questions for German Society for Non-destructive Testing (DGZfP) in Berlin, Germany.

Ulf Ronneteg, born 1970, received his university certificate in chemical engineering in 1993. After university he worked as development engineer in NDE in chemistry at ODAL, Norrkoping, Sweden and moved in 1995 to National Starch and Chemical Quality as engineer. From 1996 to 2008 he worked as development engineer in NDE at CSM Materialteknik (Bodycote) in Linkoping, Sweden. Since 2008 he serves as NDT manager at Swedish Nuclear Waste and Management Co (SKB), Oskarshamn, Sweden. He is responsible for the development of NDE for the components and welds in the canister for long-term storage of the Swedish spent nuclear fuel.

Dr. Eng. Frank Schubert, born 1966, received his diploma degree in physics at the University Duesseldorf, Germany and his doctorate at the Technical University Dresden, Germany. He received the Schiebold medal (1997) and Berthold award (2005) of the German Society for Non-destructive Testing (DGZfP). He holds a current position at the Field Support Technology Group as head of MD Simulation Center, Fraunhofer Institute for Ceramic Technologies and Systems, Branch Materials Diagnostics (IKTS-MD), Dresden, Germany.

Greg Selby received a bachelor of science degree in nuclear engineering from Oregon State University, USA. Selby started his career at Pacific Northwest National Laboratory in Richland, WA as a senior engineer in the Non-destructive Evaluation (NDE) group. In 1983, he joined the Electrical Power Research Institute (EPRI) as senior engineer in the NDE program. Selby's research has focused on developing ultrasonic techniques for examining reactor piping, pressure vessels, pressure vessel internals, bolting, and turbines. His research in NDE reliability led to his employment at EPRI and to his initial responsibilities there, developing training and qualification programs for NDE of reactor coolant piping systems. Selby is currently senior technical executive at the Electric Power Research Institute (EPRI). He holds key strategic leadership responsibilities within the NDE program and its staff of 50 members.

Abstract

Zusammenfassung der „Open Space Technology" Diskussionen. Seit den Anfängen der europäischen amerikanischen Workshops (EAW) im Jahr 1997 war es das Ziel, Experten auf dem Gebiet der ZfP-Zuverlässigkeit zusammenzuführen und aktuelle Themen zu diskutieren, um wichtige Probleme zu identifizieren und Wege für deren Lösung vorzuschlagen. Dies wurde in der Regel während der sogenannten „Break-out Sessions" erreicht, in denen vorgegebene Themen diskutiert wurden. Während des 5. EAW, der in Berlin im Jahr 2013 stattfand, wurde dieser Ansatz durch einen Open Space Technology (OST) Ansatz ersetzt. Open Space ist eine Methode der Gruppenmoderation, die sehr gut geeignet ist, um neue Ideen zu entwickeln und selbstständiges Arbeiten zu fördern, d. h. die Themen werden nicht vorgegeben, sondern ergeben sich in diesem Moment in diesem Raum und durch die Wahl der Teilnehmer. Die folgenden Themen wurden herausgearbeitet: neue Zuverlässigkeitsmethoden (Bayes, MA-POD, ...), Structural Health Monitoring, Definition von Anforderungen der Kunden gegenüber den ZfP-Anbietern, Welcher Wert der POD ist gut genug?, der Einfluss menschlicher Faktoren, manuelle gegenüber automatisierte Prüfung, sowie Grundkonzepte der ZfP-Zuverlässigkeit. Die Teilnehmer wurden aufgefordert, von einer Sitzung zur anderen zu wechseln und offen ihre Meinung zu äußern. Diese Meinungen wurden am Ende von einer ausgewählten Gruppe von Moderatoren zusammengefasst und werden in diesem Beitrag vorgestellt.