

# Production of Real Flaws in Probability of Detection (POD-) Samples for Aerospace Applications

Mika KEMPPAINEN\*, Iikka VIRKKUNEN\*  
\*Trueflaw Ltd., Espoo, Finland  
(+358 45 635 4414, mika@trueflaw.com)

**Abstract.** Assessment and demonstration of NDE reliability is important part of the inspection system. Capability is commonly assessed in terms of probability of detection (POD) as function of crack size (target size). For performance demonstration, number of cracks (with different size) is inspected and POD estimated from acquired inspection results. In order to provide reliable capability demonstration, the demonstration should closely mimic real life inspection. The cracks used in demonstration should closely resemble real-life defects. The component geometry, material and other features that may affect inspection reliability should also be representative of real inspection. Historically, it has been challenging to produce relevant test samples with high number of representative defects.

Trueflaw Ltd. produces real thermal fatigue cracks to NDE-applications. These cover different inspection techniques (UT, EC, X-ray, penetrant, magnetic particle, etc.) and whole range of inspection objects from nuclear to aerospace applications.

In this paper two aerospace cases are presented, where the clients needed real flaws in their POD-samples. Flaws were to be produced in-situ to the original samples.

The first case was produced for a jet engine manufacturer in Germany. In this case the sample was a turbine disk supplied by the client. Material of the disk was Inconel alloy. Totally 60 flaws were produced in the original sample by Trueflaw's in-situ crack growth process. Flaw sizes and locations were specified by the client and followed in the production. After delivery of the sample, the client used it in his POD-trials to assess the capability of his inspectors.

The other case was for Rolls-Royce plc, UK, requiring a range of defects in the actual component, as opposed to representative test pieces. By using the Trueflaw method the cracks could be placed accurately in the complex geometry. In total 14 flaws were produced in the original POD-sample by Trueflaw's in-situ crack growth process.

In this paper the successful results of flaw manufacturing in the two cases are presented and discussed.

## 1. Introduction

Assessment and demonstration of NDE reliability is an important part of an inspection system. Capability is commonly assessed in terms of probability of detection (POD) as function of crack size (target size). For performance demonstration, number of cracks (with different size) is inspected and POD estimated from acquired inspection results. In order to provide reliable capability demonstration, the demonstration should closely mimic real life inspection. The cracks used in demonstration should closely resemble real-life defects. The component geometry, material and other features that may affect inspection reliability should also be representative of real inspection. Historically, it has been challenging to produce relevant test samples with high number of representative defects.

The structural significance of flaws increases, in general, with increasing crack size. Thus, any available NDE system is developed to find as small cracks as possible as the value of the inspection increases with smaller detected cracks. At the same time, the acquired signal response and the ease of detection for many NDE techniques decreases with decreasing flaw size. Thus, the reliability of crack detection and hence the value of the inspection decreases, in general, with decreasing flaw size. Consequently, it is of interest to find the smallest flaw sizes, where the inspections can still provide sufficiently reliable results. Many NDE methods are used to detect flaw sizes in a range, where the probability of detection is still high, but significantly below one. This is where the POD curves and, in particular, the lower limit POD curve is used to plan inspections.

Significant amount of data is needed to reliably estimate the POD as function of crack depth. Data is needed both in terms of repeated inspections and in terms of different flawed samples inspected. Both the inspections and cracked samples are costly. Thus the main challenge in POD determination is to extract maximum amount of information from limited data and to get realistic lower limit POD curve.

The best practices for attaining a POD curve are well established in the aerospace industry and are thoroughly documented in MIL-HDBK-1823 [1]. The MIL-HDBK-1823 approach is based on seminal work by Berens [2] and relies on rather sophisticated statistical methods to attain the lower limit POD curve for a given data. These methods significantly reduce the amount of needed data, as compared to simpler, previously used, methods. Nevertheless, preferably 60 and no less than 40 cracked samples with various crack sizes are needed to reliably estimate the POD curve. Due to this high number of samples required (among other things), the MIL-HDBK statistical approach for POD determination has not so far found widespread use outside the aerospace industry, albeit there's been recent increase of interest for it's application in the nuclear industry [3].

In addition to the sheer number of cracks, determining POD curve contains other challenges: the validity of the obtained POD curve depends on the representativeness of the used test samples and inspection procedures. Manufacturing representative test pieces has traditionally been quite problematic. The test specimens should reflect the structural types that the NDE process will encounter in application with respect to geometry, material, part processing, surface condition, and, to the extent possible, target characteristics. In practice, producing realistic cracks in samples with realistic geometry and material has rarely been achieved.

Dye penetrant inspection and eddy-current inspection are commonly used for the detection of service-induced cracks in aerospace applications. Flaws to be detected are very small, as compared, e.g., to the nuclear applications. Both dye penetrant and eddy-current inspections are slow to perform and suffer from subjective interpretation of the results. Weekes et al. [4] show the results of POD studies for dye penetrant and eddy-current inspections.

New techniques based on active thermography, such as eddytherm and thermosonics (sonic IR, ultrasound-stimulated thermography), have the advantage of rapidness and automation in the interpretation of the results. Both of these techniques have shown high sensitivity in detecting small cracks. Weekes et al [4] reports the probability of detection (POD) for eddytherm and compared the results to similar POD studies with dye penetrant inspection, eddy-current inspection and thermosonics (see Figure 2, where the tabulated results of Weekes et al have been drawn to show the  $POD_{90}(a)$  curves).

The aim of a POD study is to determine the lower limit probability of detecting cracks as a function of crack size. Weekes et al [4] showed that, in addition to that eddytherm is very rapid, it has very high sensitivity to small cracks. The results show that, when applied and analyzed automatically, eddytherm can detect sub-millimeter cracks with a high degree of confidence, as shown in Figure 1. Furthermore, the comparison results of Weekes et al [4] show that both eddytherm and thermosonics have quite similar capability to detect sub-millimeter cracks (Figure 2). Hence, the choice between these two techniques is recommended to be based on their respective practicalities (e.g., area to be inspected, requirement for non-contacting inspection, etc). Comparison to more conventional NDE methods; dye penetrant and eddy-current inspections (see Figure 2), showed that dye penetrant can detect even smaller cracks than eddytherm, but may miss bigger ones due to over-washing. In eddy-current inspection the sensitivity can exceed eddytherm's sensitivity in tightly controlled inspections, but a manual inspection (most of the practical cases are performed manually) is typically less sensitive than eddytherm.

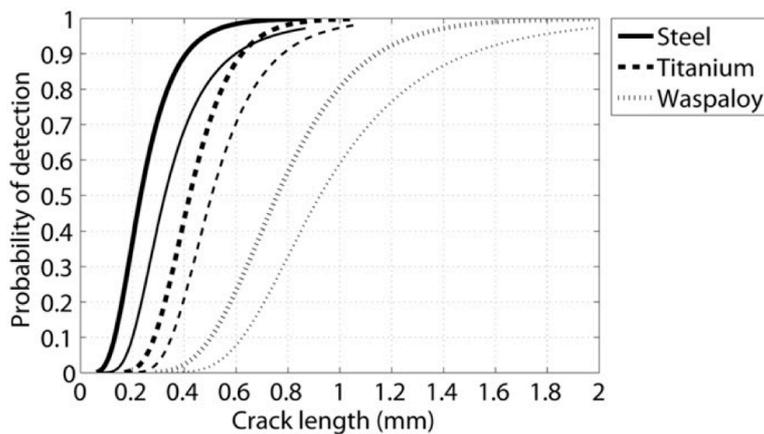


Figure 1 POD curves of Eddytherm for Steel, Titanium and Waspaloy. [4]

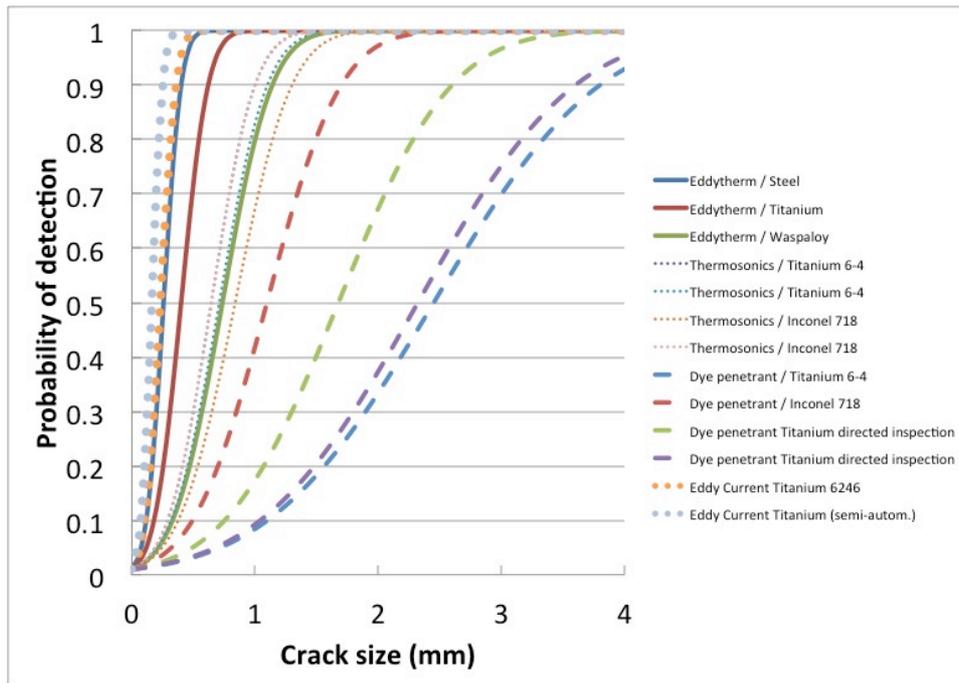


Figure 2 Curves estimated based on  $a_{90}$  values given by Weekes et al. [4]

Trueflaw's unique technology has been used widely in training and qualifying inspectors for inspecting nuclear power plants (NPP's). Trueflaw's technique has been used to produce true thermal fatigue cracks to samples provided by clients. In each case the flaw characteristics (length, depth, opening, location, skew, tilt etc.) were determined beforehand. Also the tolerances used have been tight and they have been successfully reproduced in all cases. Samples have been full-scale mock-ups, actual samples from NPP's, and different simplified training samples. In all samples, flaws were needed without any additional disturbances to the ready-made samples than the natural crack. The published cases (see e.g., [5]) show that Trueflaw's technique fulfils the requirements in the nuclear engineering applications.

Trueflaw's technique has been used only in few development cases for aerospace applications, but never in aerospace POD samples. The aim of this study was to determine if Trueflaw could produce flaws to full-scale aerospace POD samples. Normally, the POD studies are done using small flat samples. Those samples do not include the challenge provided by the actual geometry. Hence, Trueflaw technology would allow aerospace companies do POD studies including the demanding challenge of actual geometries.

## 2. Materials and Methods

In this paper two aerospace cases are presented, where the clients required real flaws in their POD-samples. Flaws were produced to real samples supplied by the clients. These samples were actual components that were similar to the ones used in the actual location.

Trueflaw Ltd. has a unique process where the flaws can be placed accurately, in-situ to the customer's complex samples. See more detailed description in [5, 7, 8]. Trueflaw utilizes patented technology where the flaws are induced by using high frequency induction heating and water spray cooling. By repeated heating – cooling cycles, the flaws are created to the ready-made samples without any additional preparation or modification of the samples. The

flaws produced are real thermal fatigue cracks with accurately controlled location and flaw characteristics (length, depth, opening, etc.). Flaws are used in different NDE-applications from nuclear to aerospace industries covering different inspection techniques (UT, EC, X-ray, penetrant, magnetic particle, etc.). Flaws are used to assess the performance of the technique used in inspecting or monitoring the true condition of a component.

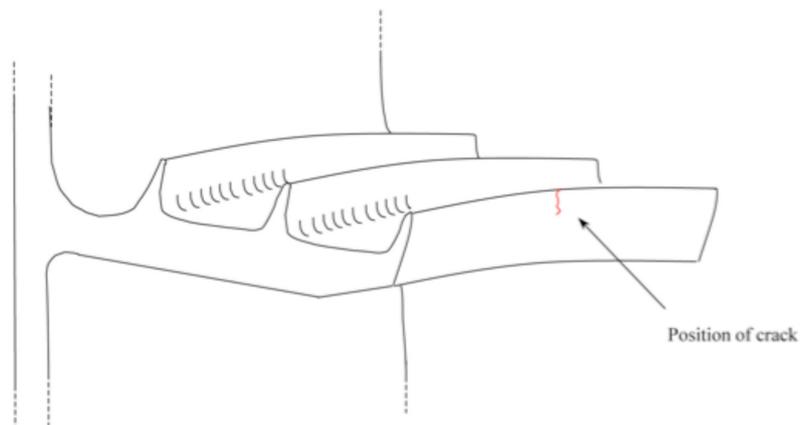
### 2.1 Sample A - German Jet Engine Manufacturer

The first case presented is flaw production in the POD sample of a German jet engine manufacturer (it is not allowed to release the name of the client). The sample was a turbine disc made of Inconel alloy. Flaws were produced in this sample by Trueflaw's in-situ crack growth process. Flaws were produced in the locations and with characteristics specified by the client (see Table 1).

### 2.2 Sample B – Rolls-Royce Plc. UK

The other case was for Rolls-Royce plc., UK, requiring a range of defects in the real component, as opposed to representative test pieces. Rolls-Royce Plc. was developing a novel inspection technique for the inspection of defects under coatings. Normally, POD studies are carried out with tens of flat test pieces with mechanical fatigue cracks. These samples are small and do not represent the actual geometry. So, as the normal way is to use mechanical fatigue flaws in very simple samples, Rolls-Royce plc. saw here the opportunity to have real flaws in their actual component. They saw that the advantage of Trueflaw's technique is the ability of placing flaws freely into any geometry and therefore carry out the POD study taking in account geometric issues.

Rolls-Royce is developing techniques for inspecting through coatings, including thermal NDE techniques. For developing the new technology, they needed a sample with known crack population. The inspection target was a seal fin specimen of a turbine disc, which is covered with a wear coating while in use. Cracks were needed under the coating in the same way that they would appear during use (see Figure 3).



**Figure 3** A schematic illustration of the seal fin region [5].

Due to complicated geometry, it would have been impossible for Rolls-Royce Plc. to use mechanical fatigue to generate cracks in the real component. Furthermore, in this case

using flat test pieces would have been so different from the actual geometry that it would have yielded any POD study meaningless [6]. By using the Trueflaw method the cracks could be placed accurately in the complex geometry. Flaws were produced in the actual POD sample by Trueflaw’s in-situ crack growth process to allow inspection of all interesting locations.

**Table 1 Samples and number of flaws produced in this study.**

Sample code	Number of flaws	Location of flaws	Flaw characteristics
Sample A	60	Including hole corners, fillet radii, etc.	Range of crack sizes suitable for POD demonstration, representative crack characteristics
Sample B	14	Several locations of the sample	

**3. Results**

*3.1 Sample A - German Jet Engine Manufacturer*

Flaws were produced in different locations of the disc sample. This Inconel alloy disc sample is used in POD studies to determine the performance of NDE inspectors. The results of the flaw production are reported in the following.

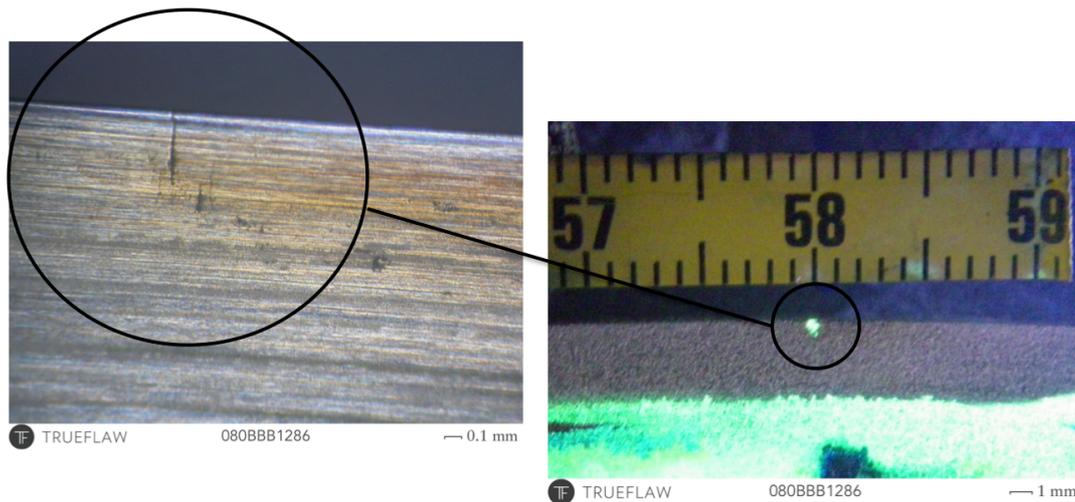
Trueflaw’s capability to produce required number of flaws in to real specimen was shown in this work. An example of the produced flaws is shown in the following Figure 4. The produced POD sample is used for assessing the performance of the inspectors.



**Figure 4 An example picture of one of flaws in the sample A (fluorescent dye penetrant image).**

*3.2 Sample B - Rolls-Royce plc. UK*

Previously, Trueflaw manufactured flaws for Rolls-Royce plc., UK, for their development sample [5]. In that development work, it was shown that Trueflaw’s technology has the potential of producing flaws to actual components. Rolls-Royce plc. used this first development sample in series of trials while developing their new inspection technology. The current POD sample builds on the earlier work. The POD sample is used to assess the performance of the new inspection technology and inspectors. The following Figure 5 shows a typical example of produced flaws.



**Figure 5** An example picture of one of the flaws in the sample B (microscopic image on the left, fluorescent dye penetrant image on the right).

#### 4. Conclusions

Trueflaw placed the flaws in the planned locations with specified flaw characteristics. The specified flaw characteristics were flaw sizes (length, depth) and openings varying between different flaws (releasing detailed flaw characteristics is prohibited). Trueflaw placed the flaws in the actual components supplied by the clients. Flaw manufacturing for both of the cases described above were successful.

The results show that, the problem of not being able to produce real flaws to true geometries and size of inspection samples (as indicated, e.g., in ref. 4) is overcome. It was shown that flaws could be produced to wide range of geometries, in different flaw locations and crack sizes.

Results show, that the Trueflaw method can be used to manufacture cracks to real components for use in POD studies. This allows more realistic POD curves to be determined that include the effect of sample geometry.

These flawed samples are now used in assessment of probability of detection performance of different inspectors and inspection techniques.

#### Acknowledgements

The authors want to acknowledge the support from the industrial partners and their permission to publish results in this paper.

#### References

- [1] MIL-HDBK-1823, 1999. Department of Defense handbook: Nondestructive Evaluation System Reliability Assessment. MIL-HDBK-1823, 30th April 1999.
- [2] Berens, A. P. NDE Reliability Data Analysis in: Metals Handbook, 9th ed. Vol 17.
- [3] Gandossi, L. & Annis, C. 2010. Probability of Detection Curves: Statistical Best-Practices, ENIQ report No 41. European Commission Joint Research Centre Institute for Energy, EUR 24429 EN, ISBN 978-92-79-16105-6.

- [4] Weekes, B., Almond, D.A., Cawley, P. and Barden, T., 2012. Eddy-current induced thermography – probability of detection study of small fatigue cracks in steel, titanium and nickel-based superalloy. *NDT&E International* 49 (2012) pp. 47-56.
- [5] Virkkunen I., Kempainen M., Ostermeyer H., Paussu R. and Dunhill T., 2009. Grown cracks for NDT development and qualification. *Insight* Vol 51 No 5 May 2009, 5 p.
- [6] Oral communication with client representative, 2012-05-25.
- [7] Kemppainen, M., Virkkunen, I., Ostermeyer, H. and Gribi, M., 2009. Development and production of test specimens to evaluate an inspection issue. *Materialprüfung mechanischer Komponenten in Kernkraftwerken, Nuklearforum Schweiz*, November 2009, Olten. 2009-11-19.
- [8] Kemppainen M., 2006. Realistic artificial flaws for NDE qualification – novel manufacturing method based on thermal fatigue. Dissertation for the degree of Doctor of Science in Technology, Espoo, Finland, (Available online from: <http://lib.tkk.fi/Diss/2006/isbn9512282631/>)