

SIMULATION SUPPORTED RELIABILITY ANALYSIS IN RADIOGRAPHY

Gerd-Rüdiger JAENISCH, Carsten BELLON, and Uwe EWERT

FEDERAL INSTITUTE FOR MATERIALS RESEARCH AND TESTING, Berlin, Germany

INTRODUCTION

Efficient and reliable non-destructive evaluation techniques are necessary to ensure the safe operation of complex parts and constructions in an industrial environment. Radiography is one of the classical non-destructive testing techniques widely applied in industry. Over the years modelling became more and more important in modern NDT. It is increasingly used to optimize techniques for complex applications, to support the preparation of written procedures, and to be introduced into the education process of NDT personnel. Hence, computer modelling is nowadays able to handle all significant properties of a NDT system with sufficient accuracy. In case of radiographic applications the model includes the radiation source, the interaction of radiation with material, the detection process, and the geometrical description of parts or constructions as well as realistic flaws. As known from practice parts and flaws can be very complex and require a description that allows the handling of arbitrary geometries. The link between NDT models and CAD provides the ability to quantitatively evaluate complex inspection procedures. Depending on the formulated inspection problem or the influencing factors that should be accessed by modelling an appropriate physical model has to be chosen to describe the underlying interaction mechanisms. The developed simulation capabilities are used to assess the reliability of the radiographic inspection of thick section welds for a special application and to determine the probability of detection (POD) for flaws of different size and shape.

Radiography using X-ray energies of 1 MeV or greater are commonly considered being high-energy applications. Today, mostly linear accelerators (LINAC) or betatrons are used as high-energy X-ray sources. The original requirement for high-energy applications was for the inspection of thick metal castings and welds. Additionally modern technologies in the nuclear field and rocket motor industry have established requirements for higher energies and more X-ray output, as can be provided by X-ray tubes. Also CT applications in the high-energy regime are known.

NDT reliability means the degree that an NDT system is capable of achieving its purpose regarding detection, characterization and false calls, according to the 1999 American European Workshop on NDT Reliability. Looking up the Metals Handbook [1] NDT reliability is a quantitative measure of the efficiency of the NDT procedure in finding flaws of specific type and size. For the reliability assessment various approaches are available:

- Probability of Detection (POD) correlates the probability that a certain flaw can be detected depending on the size of the flaw. No information about the costly false calls provided.
- Probability of False Alarm (POF) provides information about the misinterpretation of an indication and is important for economic considerations.
- Receiver Operating Characteristic (ROC) correlates the probability of detecting a flaw with the probability of false alarm for a given flaw ensemble: POD vs. POF.

This paper discusses the reliability assessment based on POD by means of simulating the system response for high energy radiographic techniques to a flaw of given size and shape. For this purpose the \hat{a} vs. a approach [1] is applied as illustrated in Figure 1.

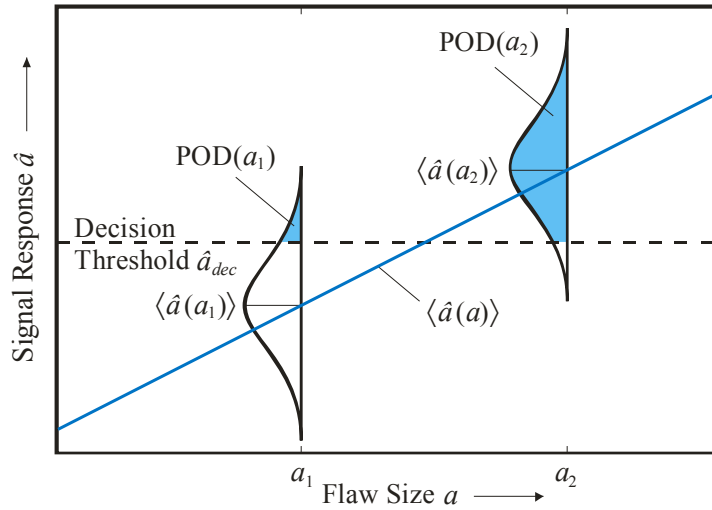


Figure 1. POD assessment: \hat{a} vs. a approach.

Consider a quantitative NDT system. As a result of the investigation of a discontinuity having the size a , it generates a signal \hat{a} . If the signal exceeds a certain decision threshold \hat{a}_{dec} , the system registers flaw detection. As the NDT system is influenced by uncontrolled factors, discontinuities of the same size can cause signals of different strength. For this reason the strength of the signal \hat{a} to the discontinuity of size a is considered as a random value. The relation between a and \hat{a} can be expressed as follows

$$\hat{a} = \langle \hat{a}(a) \rangle + \delta \quad (1)$$

with the random error term δ . According to [1] a linear relation between $\ln(\hat{a})$ and $\ln(a)$ is assumed

$$\ln(\hat{a}) = \beta_0 + \beta_1 \ln(a) + \delta \quad (2)$$

where δ is normally distributed with zero mean and constant standard deviation σ_δ independent of a . Accordingly, the $\text{POD}(a)$ function is calculated as

$$\begin{aligned}
\text{POD}(a) &= \Pr[\hat{a} > \hat{a}_{dec}] \\
&= 1 - \Phi \left\{ \frac{\ln(\hat{a}_{dec}) - (\beta_0 + \beta_1 \ln(a))}{\sigma_\delta} \right\} \\
&= \Phi \left\{ \frac{\ln(a) - (\ln(\hat{a}_{dec}) - \beta_0) / \beta_1}{\sigma_\delta / \beta_1} \right\}
\end{aligned} \tag{3}$$

where Φ is the standard normal distribution function with mean and standard deviation given by

$$\begin{aligned}
\mu &= \frac{\ln(\hat{a}_{dec}) - \beta_0}{\beta_1} \\
\sigma &= \frac{\sigma_\delta}{\beta_1}
\end{aligned} \tag{4}$$

A maximum likelihood method is used for estimating the intercept β_0 , the slope β_1 , and the standard deviation σ_δ from the \hat{a} vs. a data together with the 95% confidence bounds as described in detail in [1].

SYSTEM CHARACTERIZATION

Radiographic Setup

The radiographic set-up for the X-ray evaluation is shown in Figure 2. A 9 MeV LINAC of the company Varian is used as X-ray source. For detection a collimated line detector with 2048 pixels (pixel size 50 μm) of the company BIR is positioned outside the canister made from cooper. The angle of incident is chosen to 35° to ensure single wall penetration. For scanning the full weld, the canister is rotated by 360°. Data acquisition and LINAC pulses are synchronized. Standard image correction is applied. The measured radiographs are evaluated by trained personal.

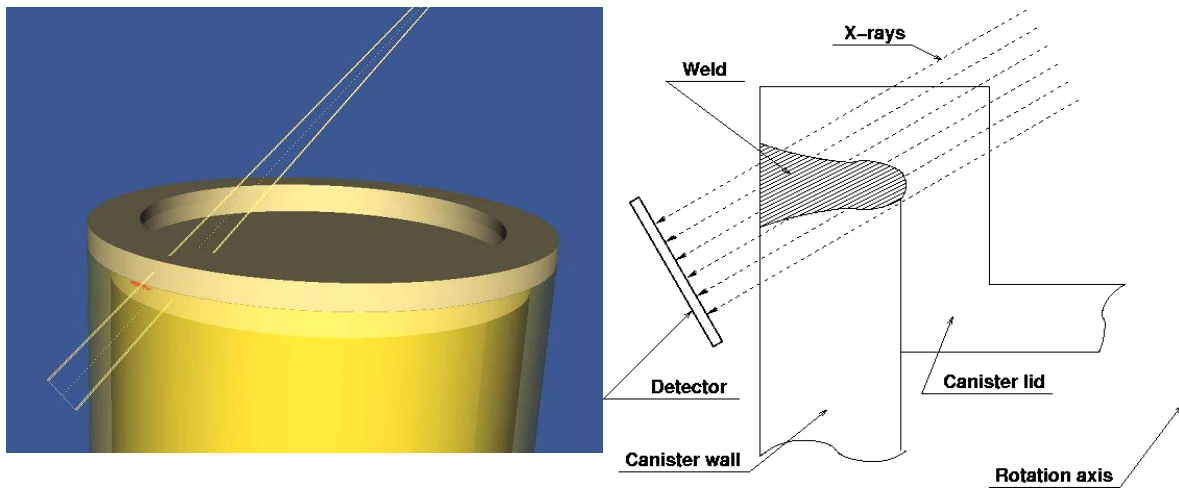


Figure 2. Radiographic set-up: virtual scene (left) and schematic view (right).

X-Ray Spectrum

For the characterization of the X-ray source two methods are used: (i) Monte Carlo simulation of the Bremsstrahlung production applying the Monte Carlo N-particle transport code MCNP [2] and (ii) a specially developed technique to estimate the photon spectrum from attenuation measurements.

Monte Carlo calculations have been performed to describe the energy spectra of the linear accelerator. These spectra can be used to describe the radiation source in the X-ray simulation program called aRTist [3] (analytical RT inspection simulation tool). For the calculations the Monte Carlo code package MCNP was used. The energy spectra of photons have been calculated resulting from 9 MeV electrons impinging on a tungsten target. Target and filter thicknesses have been chosen as follows: 28 mm tungsten, 15 mm water, 25 mm copper. Figure 3 shows the calculated photon spectrum after penetrating 15 cm of copper. The graphs of attenuated spectra show the total, primary, and scattered spectra.

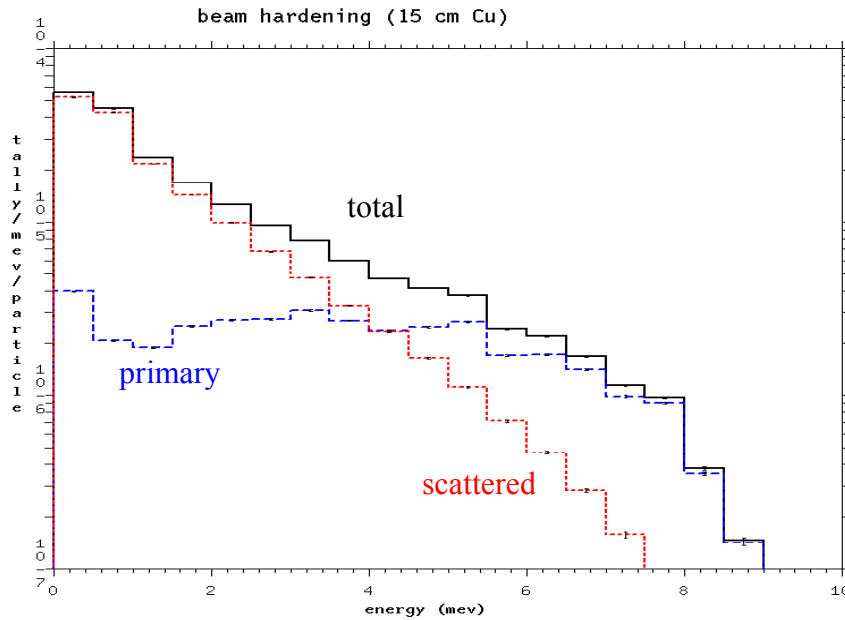


Figure 3. Monte Carlo results for spectrum evaluation (MCNP output).

A second attempt was undertaken to gain information about the spectrum for high-energy X-ray sources if the spectrum is not directly accessible by measurements. To estimate the source spectrum in terms of an inverse problem, measurements of the total intensity or the dose are evaluated: If the measurement set-up is known, the Green function describing the response of a system to a delta pulse can be calculated. Assuming a linear operator A , the Green function $G(E, E')$ is given by

$$G(E, E') = A\delta(E - E'). \quad (5)$$

The transmitted spectrum $f(E)$ follows with the Green function to

$$f(E) = \int dE' f_0(E') G(E, E') \quad (6)$$

with the initial spectrum $f_0(E)$. If scattering is neglected, the attenuation law gives the linear operator A . Otherwise the operator is given by the Boltzmann equation. The Green function can be calculated by Monte Carlo simulation. Once knowing the Green function it can be used for reconstructing the initial spectrum $f_0(E)$ from measurements of the transmitted dose. Combining dose measurements for aluminium, copper, and lead, Figure 4 shows the result of spectrum reconstruction. The estimated average photon energy is about 3.5 MeV, which agrees with the experimentally determined integral attenuation coefficient (see discussion below).

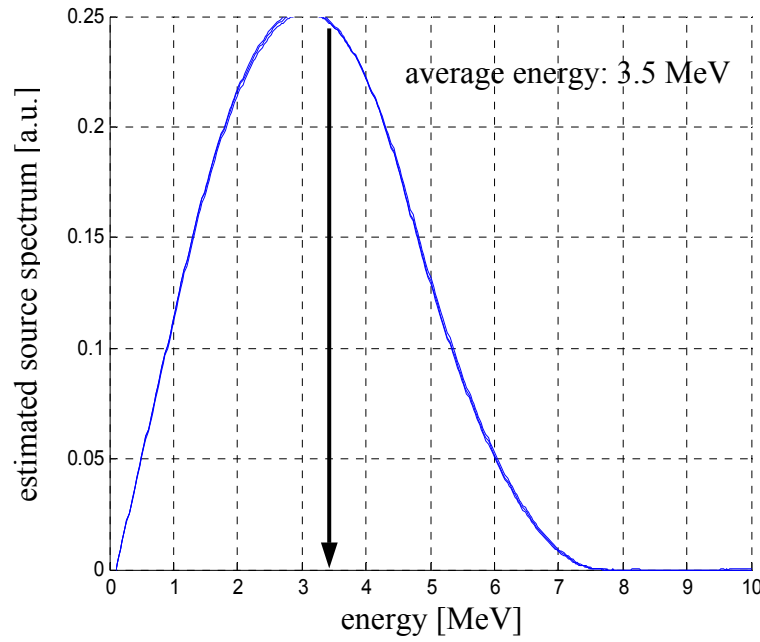


Figure 4. Estimated spectrum from measured dose data.

Experimental Estimation of Attenuation Coefficient

Experiments have been performed to determine the attenuation coefficient for copper within the full application range, i.e. between 50 mm and 165 mm. For the experiments an absorber is being placed in front of the detector. The collimation of the detector reduces the contribution of scattered radiation by about 95%. The measurements have been evaluated in two means: estimation of the (i) integral and (ii) differential linear attenuation coefficient based on the attenuation law. To calculate the integral attenuation coefficient an exponential fit over the full range was used. As result, an attenuation coefficient of 0.299 cm^{-1} was found (see Figure 5). Note that the multiple R^2 value is close to 1 describing the correlation between the data and the regression. It can be concluded from this fact that the data fit well to the exponential attenuation law with a constant linear attenuation coefficient over the full application range between 50 mm and 165 mm thickness of copper. The corresponding average energy of the source spectrum was determined to 3.8 MeV, which is in good agreement with the average energy found from the estimated spectrum (compare Figure 4).

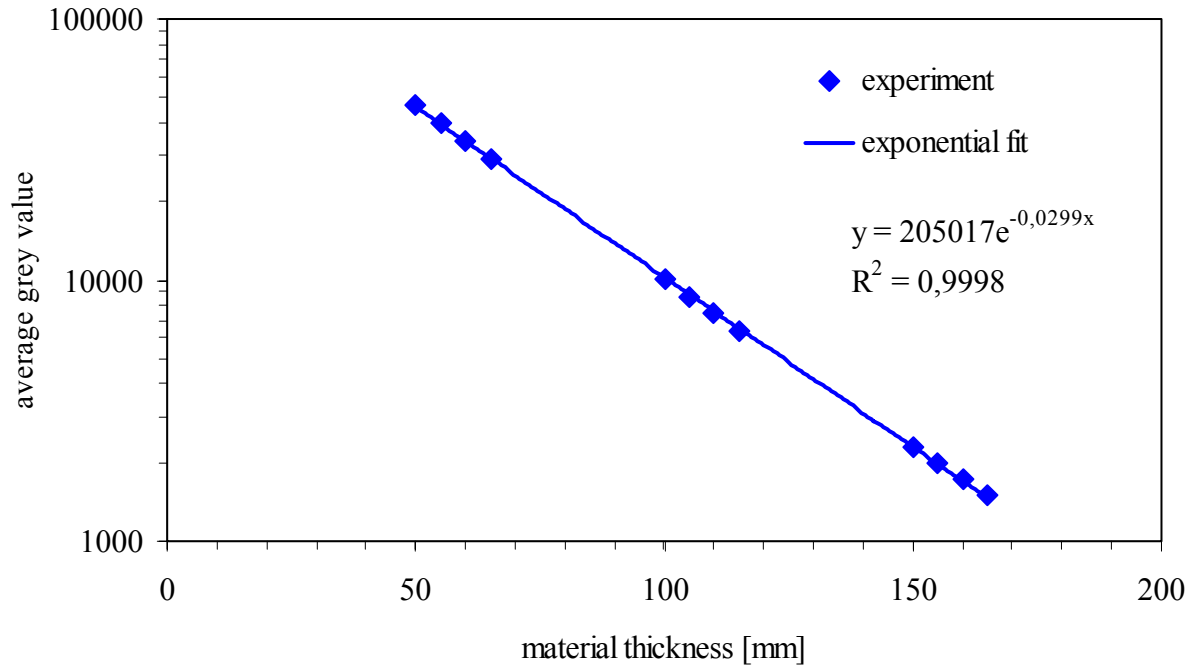


Figure 5. Result of exponential fit for the integral attenuation coefficient: $\mu=0.299 \text{ cm}^{-1}$.

The differential linear attenuation coefficient was calculated from two neighbouring data points according to

$$\mu = -\frac{1}{\Delta d} \ln \left(\frac{I(d + \Delta d)}{I(d)} \right) \quad (7)$$

with the thickness d and $d+\Delta d$ of neighbouring data points. The error propagation law gives the uncertainty of the calculated attenuation coefficients. Figure 6 shows the tendency that the linear attenuation coefficient decreases with increasing material thickness. This behaviour can be understood by the hardening effect, i.e. with increasing penetrated material thickness the average energy of the transmitted spectrum increases. Taking into consideration the error of estimation, as shown in the figure, this tendency of decreasing differential attenuation coefficient is significant. The same tendency was found by earlier dosimeter measurements given in the figure by the black points.

The differential attenuation coefficients decrease of about 5% within the thickness range between 50 mm and 165 mm which corresponds to a hardening of the spectrum of about 1.4 MeV (average energy at 50 mm thickness is found to be 3.0 MeV; at 165 mm thickness as to 4.4 MeV). Also this result does not agree with the spectrum and its characteristics calculated by MCNP.

As a result of this investigation the integral attenuation coefficient is used for predicting the system response by means of simulation.

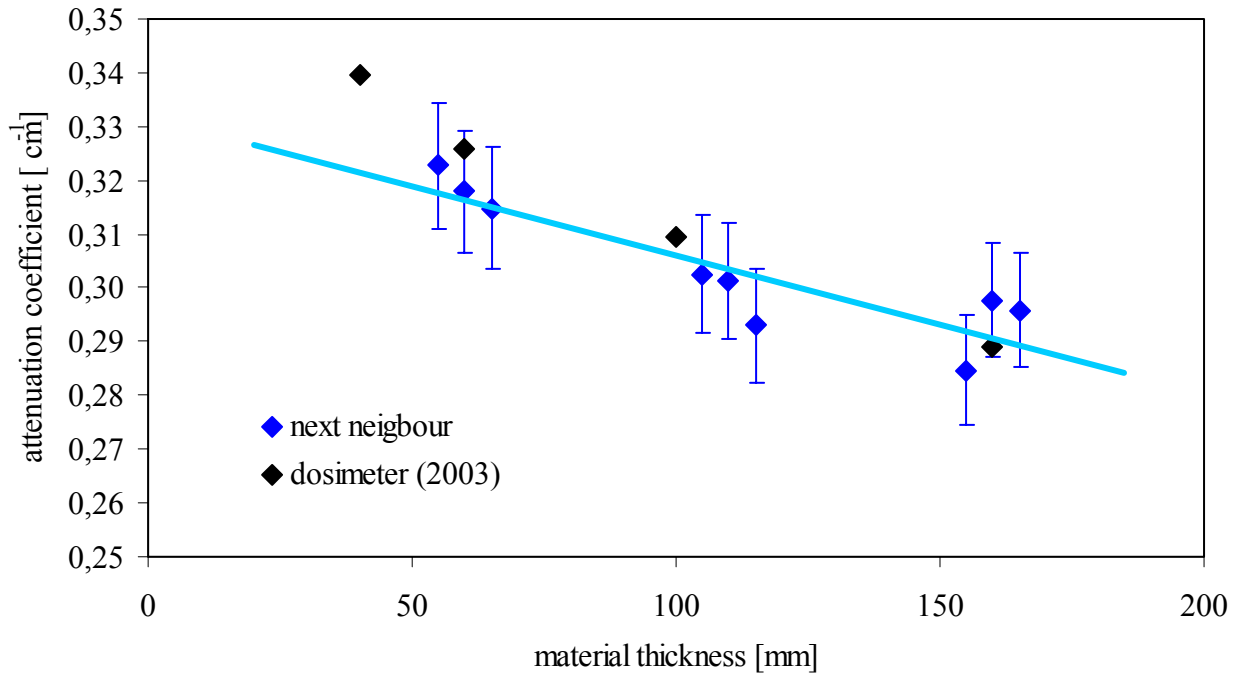


Figure 6. Estimation of differential attenuation coefficient.

Comparison of Modelling Results with Experiments

A test specimen (see Figure 7) with defined voids was designed to verify the measured radiographic and geometrical parameters of the RT set-up and the simulation tool aRTist. Holes with diameters between 2 mm and 32 mm were introduced to investigate the radiographic projection. The test specimen was placed on top of the 50 mm copper canister wall with support of an additional segment to have the holes at weld position.

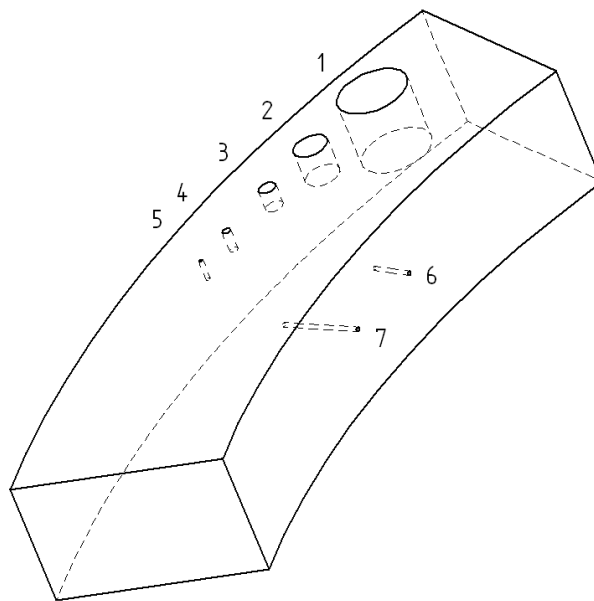


Figure 7. Test specimen for model verification.

The comparison of the resulting simulated image with the experimental radiograph gives a good visual agreement (compare Figure 8). There are visual divergences in the images originating from disturbances of the radiographic detection process, which are not covered by modelling. This is among others the horizontal striped background of grey value shifts from scan line to scan line. Blooming effects arise in line direction at high contrast steps. Also a dark vertical stripe is seen in the experimental image, which arises from a misalignment of the test specimen and the supporting segment. This misalignment is considered in the virtual model by a small opposite translational displacement of the two segments such that the stripe could be reproduced by the simulation.

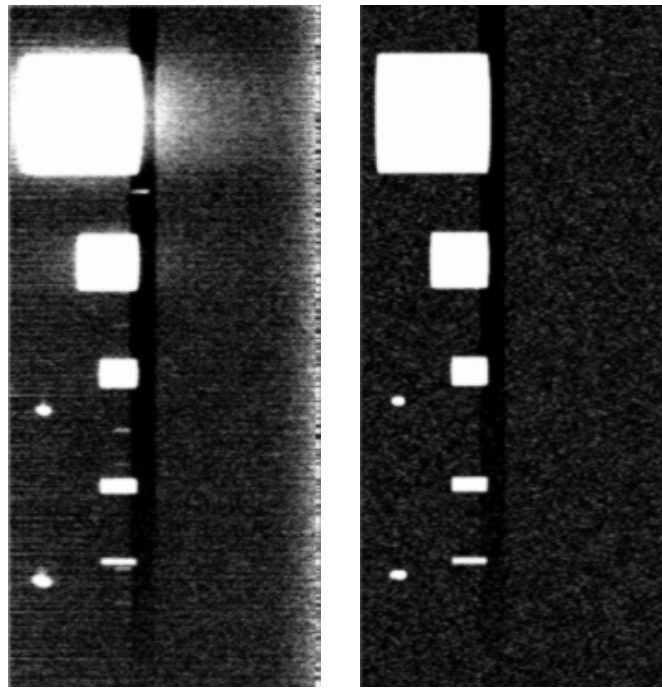


Figure 8. Experimental (left) and simulated (right) radiograph of the test specimen.

The quantitative evaluation is focused on flat bottom holes numbered 1 to 5 in Figure 7. Table 1 gives the results from the measurements for the experimental radiograph and the corresponding simulation. First, there are given grey values for different penetrated lengths referred by the flaw number. The first line (without flaw number) gives the values at full wall thickness for completeness. All other grey values are measured at maximal flaw depth, which is given by the hole's diameter. Further, Table 1 lists the flaw depth or maximal reduction of penetrated length due to a flaw. Besides the real flaw depth as reference, there are values calculated from the radiographic contrast using the attenuation law. For the large flaws a good agreement with the true flaw size ($< 2\%$) is found. As the real flaw size decreases, the extracted flaw depth is increasingly underestimated. If the flaw size approaches the unsharpness of the system, the contrast of the indication is reduced due to unsharpness, i.e. the contrast is not proportional to the flaw depth any more. Hence, the radiographic

contrast does not reflect the attenuation law any more. Therefore the extracted flaw depth is too small.

Table 1. Comparison of grey values and flaw depths from experimental and synthetic radiographs.

flaw no.	gray value at flaw depth		flaw depth [mm]		
	experiment	synthetic	real	experiment	synthetic
-	10000	10000	0,0	0,0	0,0
1	26649	26898	33,0	32,6	33,0
2	15750	15968	15,6	15,1	15,6
3	12544	12692	8,0	7,6	8,0
4	11258	11296	4,2	4,0	4,1
5	10566	10586	2,0	1,8	1,9

POD ESTIMATE BY MEANS OF SIMULATION

General Remarks

The POD assessment follows the \hat{a} vs. a model described in detail in [1] assuming a linear relation between $\ln(\hat{a})$ and $\ln(a)$ with normally distributed deviations (\hat{a} – system response, a – flaw size). To apply this model to radiography the following assumptions are made: (i) the contrast ΔI of an indication can be directly correlated with the size of the flaw, particularly with the wall thickness difference in beam direction

$$\begin{aligned}
 \Delta I &= I_2 - I_1 \\
 &= I_0 \left(e^{-\mu(d-\Delta d)} - e^{-\mu d} \right) = I_0 e^{-\mu d} \left(e^{\mu \Delta d} - 1 \right) \\
 &\approx I \mu \Delta d
 \end{aligned} \tag{8}$$

(ii) the contrast of an indication contains all information necessary to interpret the image, (iii) the noise in the radiographic image is normally distributed with zero mean. For the following investigation pores, described as spherical voids of different diameter, have been placed in the copper canister and the radiographic images have been simulated. The contrast is measured as the difference between the maximum gray value at the flaw indication and the average gray value at the same position without a flaw present. Figure 9 demonstrate the influence of the considered system unsharpness. If unsharpness is excluded from the simulation the largest indication contrast is found. By adding separately the measured unsharpness the contrast decreases. The difference between the shown 3 cases decreases with increasing pore size. Figure 10 shows the \hat{a} vs. a data that have been used for the POD estimate together with the corresponding decision threshold \hat{a}_{dec} for SNR=2.7 according to the measured system characteristics.

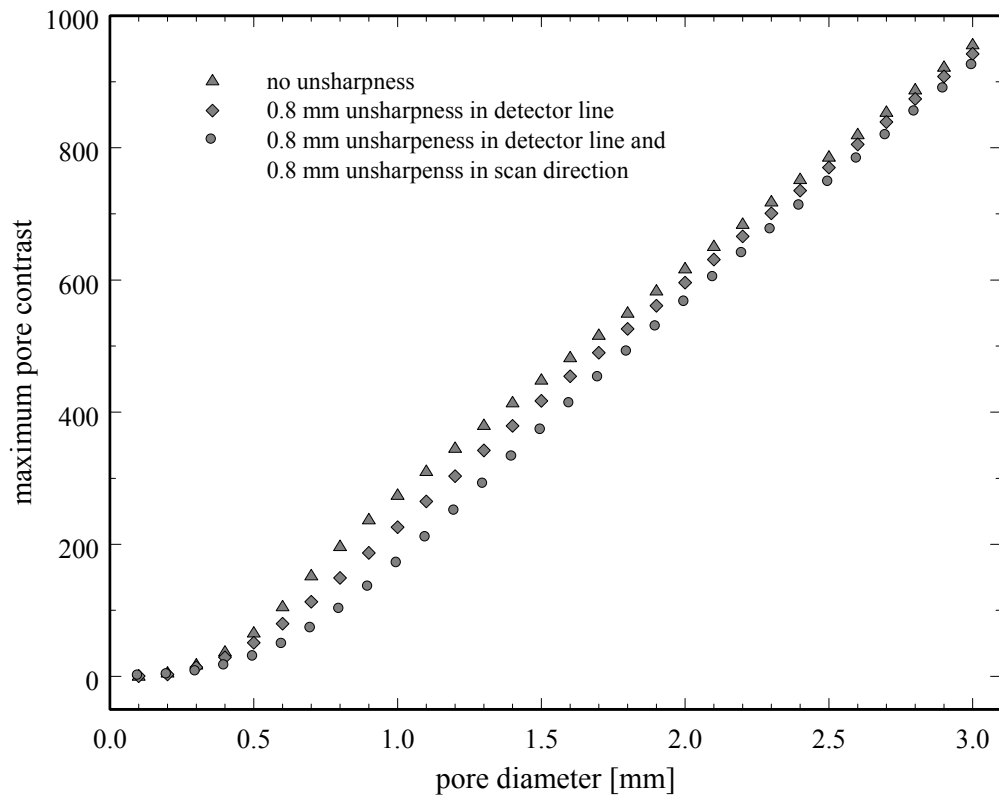


Figure 9. Influence of system unsharpness on indication contrast.

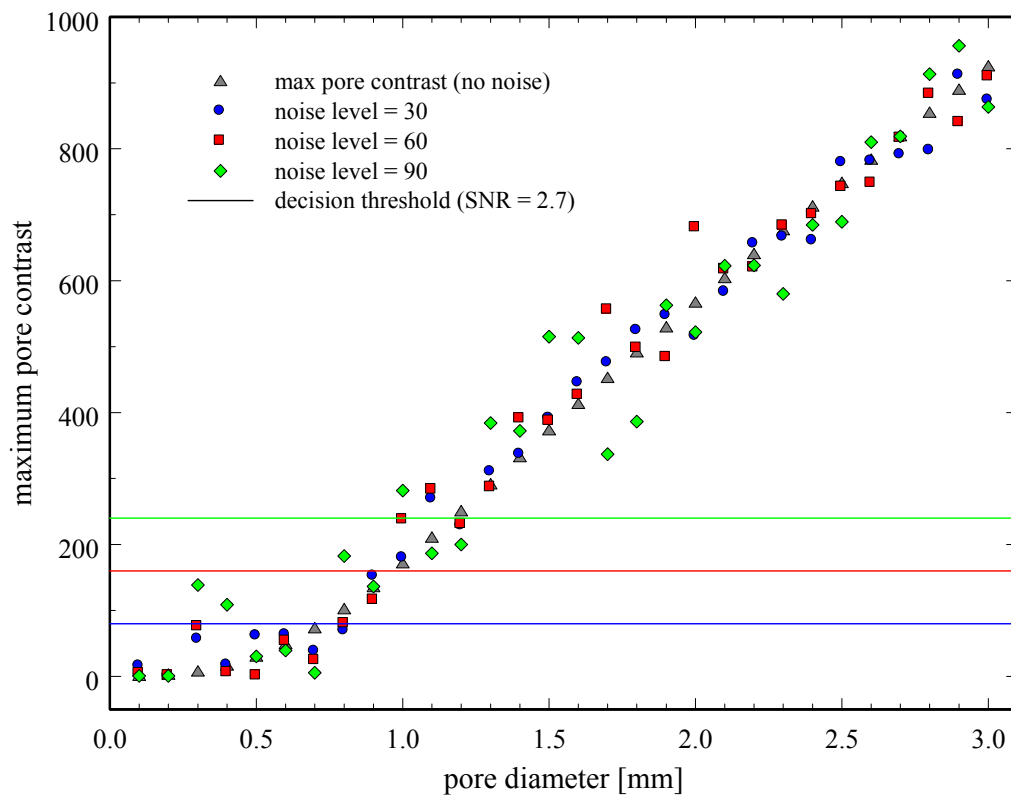


Figure 10. \hat{a} vs. a data for POD estimate for different noise contributions.

POD Assessment

The POD according to [1] is given in Figure 11 for different decision thresholds \hat{a}_{dec} reflecting a SNR of 2.7 for different noise contribution in the image. Additionally the $a_{90/95}$ values together with the 90% POD value are listed in Table 2. The calculated a_{90} and $a_{90/95}$ values increase with the amount of noise in the image. Anyway, indications from volumetric flaws larger than 2 mm, which corresponds to a wall thickness change of about 2%, are detected with high reliability in any case. This result agrees with radiographic practice. It has to be pointed out, that the presented POD evaluation gives a conservative estimate of the system reliability to be discussed in more detail.

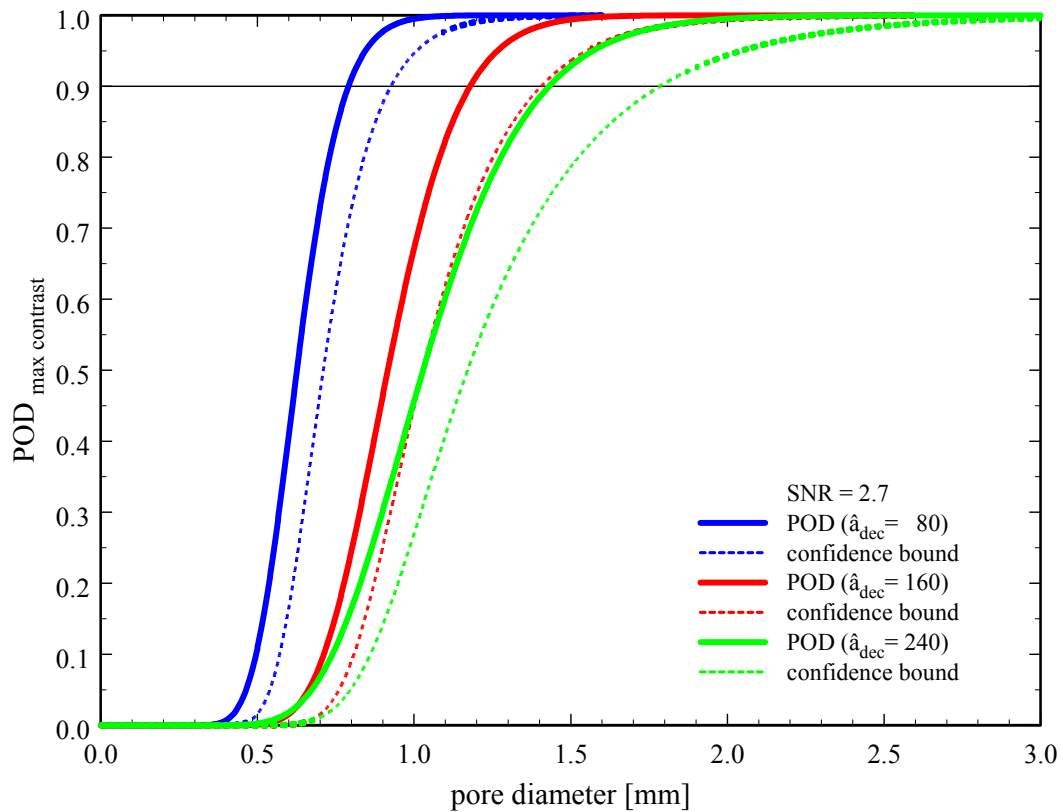


Figure 11. Theoretical POD investigation for different decision thresholds \hat{a}_{dec} .

Table 2: Theoretical POD investigation.

decision threshold (SNR=2.7)	a_{90} [mm]	$a_{90/95}$ [mm]
80	0.79	0.93
260	1.18	1.41
240	1.43	1.79

Remember the assumptions made to apply the POD model [1] to radiography. It was assumed that the contrast of an indication contains all information necessary to interpret the image. But in reality, the human operator draws his decision about the presence of a flaw indication in a radiograph from more than only one feature, i.e. not only the contrast of an indication is used, but also its shape, orientation and others. Especially the fact is neglected, that a crack indication with SNR smaller than one is clearly detectable by a human operator. The POD model [1] does not reflect this fact. Hence, if applying this model to radiography using only the contrast of an indication as signal \hat{a} , the resulting POD gives a conservative estimate of the system reliability because it neglects important features for image interpretation, i.e. flaws larger than the $a_{90/95}$ value – flaw size at which the 95% confidence bound to the POD curve exceeds a POD of 90% – are detected with much higher probability. On the other hand, the lower part of the POD curve does not contain sufficient information to define the application limits of the technique or to optimize the technique for detecting small flaws. Finally, the POD model does not give any information about false calls or how to reduce them. Because the operator evaluates not only the contrast of an indication but also other features to decide, if a flaw is present or not, he would reject even high contrast indications whereas those would be count as flaw in the POD investigation.

Therefore, the POD model proposed in [1] should be generalized to a multidimensional response function, which is not necessarily a normally distributed multidimensional random variable but includes all features essential for the interpretation of images.

CONCLUSIONS

The paper reviewed results for the applicability of modelling for high-energy radiography including reliability investigations. It has been demonstrated, that all essential parameters influencing the result of high-energy radiography for a special application can be described by modelling. The discussion has been shown, that the presented model is capable to predict experimental data with high accuracy even if some characteristics of the system are difficult to access like the X-ray spectrum of a LINAC.

Additionally, it has been shown that the model results can be used for the assessment of the system reliability in terms of a signal response POD applying the method proposed in [1]. In case of radiographic applications this investigation gains a conservative estimate of the system reliability because the POD model does not consider all features that are evaluated by a human operator while interpreting a radiograph. For radiographic techniques or other imaging NDT applications a POD assessment should be developed counting for multiple parameter evaluation.

Moreover, the presented approach should be included into a more general methodology such as developed within the model assisted POD activities (MAPOD) [4], because simulation can not cover

all aspects of an in-field NDT application, e.g. environmental influences or human factor. Therefore, (i) controlling factors have to be identified whose influence can be simulated using physics based models, (ii) appropriate models have to be developed and validated, (iii) the developed simulation tools are used to predict the response to well understood factors, (iv) the un-controlled or un-modelled factors are to be experimentally assessed. The computation of the resulting POD combines simulation results as well as experimentally assessed data (see Fig. 12).



Figure 12. Simulation supported POD assessment.

ACKNOWLEDGEMENT

The authors are grateful to Swedish Nuclear Fuel and Waste Management Co. (SKB) for financial support, and H. Ryden, U. Ronneteg, Ch. Müller, B. Redmer and U. Zscherpel for valuable discussions and experimental support.

REFERENCES

- [1] A. P. Berens, “NDE reliability data analysis”. In Metals Handbook, Volume 17. ASM International, 9th edition, 1989.
- [2] J. F. Briesmeister (ed.), MCNP-A General Monte Carlo N-Particle Transport Code, LANL Report LA-13709-M, Los Alamos, 2000.
- [3] G.-R. Jaenisch, C. Bellon, and U.Ewert, “aRTist – Analytical RT Inspection Simulation Tool for Industrial Application.” Proceedings of the 17th World Conference on Non-Destructive Testing, Shanghai, China, International Committee on NDT, CDrom paper 64 (2008)
- [4] Model Assisted Probability of Detection Working Group, Center for Nondestructive Evaluation, Iowa State University: www.cnde.iastate.edu/research/MAPOD