ABSTRACT

When performing a non-destructive inspection a certain amount of uncertainty has to be associated with the result. For instance, fluctuations may be observed for signals generated by nominally identical flaws or alternatively, for signals obtained by repeated scans of the same flaw. These fluctuations result from the large number of influencing factors such as small variations of the material and geometry of the specimen, the location, orientation and geometry of the flaw, the ability of the inspector. The actual values of these factors may somehow vary from their nominal values and thus bring some uncertainty to inspection results. Presence of noise during the inspection is an additional factor that must be taken into account when looking to signal fluctuations. More importantly, noise sources may cause false alarms. This paper proposes a methodology based on NDE simulation tools to obtain Probability Of Detection (POD) curves and Probability of False Alarms (PFAs).

CONTEXT

Inspection reliability is one of the key issues in ensuring safety of critical structural components. Among the various methods dedicated to NDE performances evaluation, probabilistic approaches have aroused interest because of their ability to naturally account for uncertainties. A probabilistic criterion called Probability Of Detection (POD) relates in particular the detectability of a flaw to its size and to other features. As POD curves were before exclusively obtained through expensive and time-consuming experimental campaigns, simulation-based POD calculations have recently been introduced in the NDE community.

CIVA physical models have been developed for years and are mature enough to support simulation intensive approaches like POD calculations. Dedicated statistical tools have been added to the software in order to switch from the usual deterministic framework to an approach based on uncertainty. Therefore, some model inputs are first defined using statistical distributions instead of deterministic nominal values. These distributions describe the random variations of the actual values of these parameters when simulating an inspection. Thus, they provide a quantitative representation for uncertainty. Noise sources such as UT inspection structural noise are also taken into account through the use of dedicated algorithms. Uncertainties are then propagated through the model using a sampling procedure. The objective here is to determine how the fluctuations of the influencing factors of the inspection procedure affect the uncertainty in the inspection result. This step requires the computation of the physical model for a set of random values for the input parameters. Finally, a functional form is assumed for the POD curve and the related parameters are estimated from the uncertainty propagation results.

This paper presents the statistical approach as well as the choice of distributions and POD functional forms. It also describes two applications of this methodology to realistic inspection setups using UT and EC techniques, respectively.

METHOD

Probability Of Detection (POD), Probability of False Alarm (PFA) and Relative Operating Characteristics (ROC)

POD curves relate the detectability of a flaw to its size (or to another geometrical characteristic). For a specific inspection method the POD expresses the probability that a flaw of a certain size will be
detected, provided that this flaw is actually present in the inspected component. This probabilistic approach to the assessment of nondestructive inspection (NDI) performances may be used in a wide range of circumstances [0]. For instance, it provides a basis to prove that a certain detectability level has been achieved in particular applications. It is also widely used during the design of a component with a damage tolerance philosophy. This particular approach, which stands in opposition to the safe-life design philosophy, has been used by many aircraft builders in the last decades.

Following a general procedure, POD curves are estimated from NDI reliability data that come mostly from dedicated round-robin inspection programs. In these programs, the same samples or mock-ups, which contain defects of known sizes and locations, are inspected by several technicians under realistic conditions (laboratory type or in-service). The whole process and the large amount of data required makes the determination of a reliable POD curve very expensive and time consuming. Based on a statistical analysis of the inspections results, the usual approach consists in assuming a functional form for the POD curve and in estimating the parameters of the function, as well as its associated confidence bound [0].

Inspection results may be recorded in two different formats. In the so-called “hit/miss” format, binary information is recorded by the inspection system, indicating whether or not a defect has been detected: data located above a given threshold are declared ‘hits’ (or 1), whereas the others are declared ‘miss’ (or 0). In the format known as “signal response”, the amplitude of the signal is recorded; hence more information is contained in the data. The statistical analysis applied to the data depends on this recording format.

In the case of hit/miss data, the probability of detection is estimated for each crack size by computing the ratio of hits to the number of inspections (hits + misses) for that crack size. Thus, for each flaw size, a detection probability is estimated. The objective is then to estimate a functional form for the POD that fits at best the estimated detection probabilities. Based on empirical data, it has been shown that the log-odds function provides an acceptable model. This model can be written as:

$$POD(a) = \left[1 + \exp\left(\frac{\pi}{\sqrt{3}} \left(\frac{\ln a - \mu}{\sigma}\right)\right)\right]^{-1}$$ (1)

where \(a\) is the flaw size. The parameter \(\mu\) is the natural logarithm of the crack size for which there is 50% detectability. The parameter \(\sigma\) is a scale parameter that determines the flatness of the POD function. The parameters \(\mu\) and \(\sigma\) are computed by maximum likelihood estimation using the estimated detection probabilities.

In the case of signal response data, the analysis is first made by interpreting the system response to the presence of a defect. In the usual approach, it is assumed that the logarithm of the signal response amplitude is linearly correlated to the logarithm of the flaw size. The random error between the signal response amplitude predicted using the linear relationship and the measured one is supposed to be normally distributed with a standard deviation which does not depend on the flaw size. Based on these assumptions, it can be shown that the POD curve can be modeled using a cumulative log-normal distribution function. This function can be expressed as:

$$POD(a) = \Phi\left[\ln a - \mu\right]/\sigma\right]$$ (2)

where \(\Phi(z)\) is the standard normal cumulative distribution function.

Due to the fact that the POD model is estimated from a finite size data sample, there is a sampling uncertainty on the estimates of the model parameters. This uncertainty is usually characterized by placing confidence bounds on the POD function. Characterizing the sampling uncertainty is especially important for certain applications where quantities such as the inspectable flaw size are used. This quantity, usually noted \(a_{XY}\), reflects the size of the defect that will be detected with a probability \(X\%\) and a confidence of \(Y\%\). For instance, a probability of 0.9 with a 95% confidence is used to express the inspectable flaw size which will thus be noted \(a_{0.995}\). In the approach described in this paper, the computation of the lower confidence bound is based on asymptotic properties of the maximum likelihood estimate [0,0].

Other probabilistic criteria such as the Probability of False Alarm (PFA) or the Relative Operating Characteristics (ROC) curve are also useful when assessing NDI reliability [0]. For
instance, PFA is the probability that the detection threshold is exceeded when no flaw is present (for instance, due to a strong background noise). Thus, the PFA will strongly impact the definition of the detection threshold. A good knowledge of the noise properties is required to assess the PFA and fix an adequate threshold. In this paper, we will show how a specific algorithm describing the scattering of an ultrasonic wave by a polycrystalline material can be useful to determine a proper detection threshold when strong structural noise occurs during an ultrasonic inspection.

Finally, the ROC curve is a convenient way to display simultaneously POD and PFA information. A ROC curve makes it possible to compare inspection methods without having to fix a particular threshold value. Even though ROC curves are of great interest when evaluating NDI reliability, this paper will focus on the assessment of POD curves and PFAs based on numerical simulation results and having good knowledge of uncertainty and noise sources.

**Modelling uncertainty and noise sources in the CIVA software**

Simulation codes developed at CEA aim at providing fast and cost-effective tools to predict the results of inspection methods carried out in realistic configurations. These simulation tools are typically used to conceive or to optimize inspection characteristics and probes, to qualify inspection procedures and to provide help in the interpretation of experimental measurements. Those simulation tools include beam propagation and flaw scattering models which have been previously detailed [0-0]. They are based on semi-analytical approaches where both numerical algorithms and analytical models are used. These approaches make it possible to simulate fairly complex problems in a fast and accurate way, provided that the adequate assumptions are made concerning the underlying physics involved in the problem.

**Modelling uncertainty sources using statistical distributions to describe the input parameters**

When performing an inspection, the probe signal response due to a flaw will be affected by three types of factors. Factors of the first type are related to the NDI system (transducer, scan plan, electronic device), factors of the second one are related to the part (geometry, material properties, surface roughness). Finally some factors are related to the flaw (size, shape, orientation, position). These factors are chosen to be either deterministic or random in the simulation.

![Image](image_url)

**Figure 1 - On the left:** Definition of the uncertainty sources using the graphical user interface. In this example, uncertainty of the skew value of a notch is defined using the normal distribution with a standard deviation value of 3°. On the right: Definition of the dependency parameter of the POD curve. Here, the notch height as been selected with values ranging from 1 to 8 mm.

A signal-determining factor is taken to be deterministic if either the factor can be controlled in the inspection operation or if it is desired to estimate POD as a function of the level factor (flaw size for instance). The factor is thought to be random during the production/field inspection if there is insufficient knowledge related to it, if it is not well controlled or if it implies physical phenomena with inherent randomness. Among random factors, one can cite:

- the detail of metallurgical microstructure that will lead to structural noise,
- the flaw position with respect to the probe
the flaw morphology; For a fixed size, various parameters such as shape, orientation, position, elastic properties, and conductivity can vary in a random manner;

- probe lift-off and/or orientation variations during a manual inspection using an eddy current pencil probe.

In addition, one can think of factors related to detection recognition by human operators. These human factors are not addressed in the approach described in this paper.

Some tools that have been added to the software CIVA in order to compute simulation-supported POD curves will now be described. First, random number generators for various statistical distributions have been introduced, in order to describe different kinds of parameters uncertainty. To date, distributions proposed in the software are the uniform, normal, log-normal, Rayleigh and exponential ones. The graphical user interface dedicated to uncertainty characterization through statistical distributions is shown in Figure 1. POD curves are expressed in terms of a dependency parameter which is frequently a geometrical dimension of the flaw. This particular parameter, which can be considered as deterministic in our approach, must be fixed during the process. This is done by defining its domain of variation (the numerical values for which the inspection will be conducted) and the number of inspections that will be simulated for each flaw size (i.e. the number of inspectors that will perform the inspection of the specimen containing the particular flaw). This step is also illustrated in Figure 1.

For particular sources of uncertainty, the response depends in a more complex manner on the actual properties of the setup and component. Thus, specific algorithms have been developed in order to evaluate for instance, the impact of the metallurgical structure of a cast stainless steel on the inherent variability of the signal [0]. The objective of these specific algorithms is to predict signal fluctuations due to complex physical phenomena. Thus, these algorithms complement the description of uncertainties using statistical distribution functions to describe input parameters of the model.

Modelling ultrasonic noise using a specific algorithm

A specific algorithm has been developed in order to model inherent ultrasonic noise linked to the inspection of coarse grained metallurgical structures [0]. As an ultrasonic wave propagates into a polycrystalline material, a part of its energy is scattered in all directions of space by individual crystallites. This interaction between the ultrasonic field and the microstructure of the material generates a set of random signals known as structural noise (or grain noise). Since this noise superimposes on defect related signals, it can be seen as an additional complex source of signal fluctuations. Besides, a strong noise can be the source of false alarms. It is therefore fundamental to take account of it when setting up the detection threshold.

CIVA 9 provides specialized tools for the simulation of structural noise observed in UT. The noise computation is carried out by a noise generator consisting in a set of randomly distributed point-like scatterers located into a synthetic structure. Each one of them is characterized by a scattering coefficient which is picked from a normal distribution. Two parameters must be adjusted to mimic the grain noise measured for a given material: \( \rho \) (the number of scatterers per volume unit) and \( \sigma \) (the standard deviation of the normal distribution from which scattering coefficients are drawn). The model used to compute noise is similar to the one developed by Gustafsson and Stepinski [0], which is based on a single scattering assumption. Thus, the response of a set of scatterers to an ultrasonic pulse is the sum of the individual contributions of each scatterer of the synthetic structure. In our approach, the incident ultrasonic field is computed at each scatterer site using a pencil approach, which is an extension of a ray theory [0]. As the synthetic structure remains the same for all positions of the transducer, a same scatterer is apparent for successive positions of the transducer. This property of the noise generator ensures that a generated B-scan has the same coherent properties (i.e. the same speckle texture) as the measured one. A detailed description of this algorithm and of ongoing work at the CEA on the noise generator can be found in [0].

Results presented in Figure 2 were obtained for the inspection of a forged stainless steel coupon. The inspection was performed using a probe with a 2.25 MHz central frequency and generating shear waves with a 45° refracted angle in the specimen. The simulation was performed using the following values for the input parameters: \( \rho = 0.1 \text{ mm}^{-3} \) and \( \sigma = 50 \text{ SI} \). As it can be observed
on the B-scans, the experimental SNR is well reproduced in simulation using these numerical values for the inputs of the noise generator. The textures of measured and computed B-scans have similar aspects.

By performing a first order statistic study of the ultrasonic noise, it is possible to estimate a probability density function (PDF) describing the probability of occurrence of noise amplitude values. The PDF can be estimated from a B-scan by picking all the values recorded at a specific time $t_0$ and by ranging them into bins in order to construct a histogram. An illustration of PDF estimated on either simulated or measured B-scans is given in Figure 2. These results show very close statistical properties of both experimental and simulated envelope amplitudes of the grain noise: they seem to follow a Rayleigh distribution.

Typical UT inspection results are the maximum amplitude of the rectified signal or the peak-to-peak amplitude within a time gate. Hence, it is interesting to determine the statistical properties of such quantities for the grain noise. This is equivalent to assess the statistical properties of the C-scan pixels (the C-scan displays the maximum value of the rectified signal within a time gate). Determining this PDF is useful for several purposes. One of them is to estimate a Probability of False Alarm, given a particular detection threshold. The PFA is simply defined as the probability that the noise level exceeds the threshold. Another approach would consist in first fixing an acceptable PFA and then determine the corresponding threshold.

![Probe](image1.png)

**Figure 2 -** Measured and computed B-scans and C-scans for a specimen of austenitic stainless steel. Lower left plot: Statistical properties of envelope amplitudes measured at a specific recording time $t_0$ on the B-scan: Rayleigh distribution (blue curve), simulated noise (red curve) and experimental noise (black curve). Lower right plot: Statistical properties of the C-scan pixel values (the maximum amplitude of the time gated noise): simulated noise (red curve) and experimental noise (blue curve). Shaded areas correspond to the PFA.
The results shown in Figure 2 illustrate the skewed nature of C-scan noise distributions. Several methods were proposed to establish closed-form parametric distributions of the C-scan noise statistical properties in the same way as the statistic of the envelope amplitudes can be described by a Rayleigh, Nakagami or K distribution [0]. This complex task goes beyond the scope of this paper. Here, we will assume that the properties of the C-scan noise distribution and more specifically those of the tail area of the distribution (low probability values), are well reproduced by the noise generator. Assuming that the simulated C-scan is large enough, one can estimate the PFA with a limited impact of the sampling uncertainty. The proportion of the total area under the noise curve to the right of the decision threshold is equal to the probability of false alarm.

Simulation of inspection responses in presence of uncertainty and noise sources

Once the inspection setup is defined (fixing the input parameters of the defect, part and probe with their nominal values), and the uncertainty and noise sources are modeled (using adequate statistical distributions and specific algorithms such as the ultrasonic grain noise generator), simulations can be performed. The objective of this step is to obtain realistic inspection outputs that will be further processed to compute a statistical estimation of the POD. The objective of this process being to assess how uncertainty and noise sources impacts the simulation results, this step will be designated as the “propagation of uncertainty” step.

For the time being, a simple Monte Carlo approach is proposed. It is a sampling method that consists in randomly generating values for the uncertain entry parameters according to the statistical distributions selected by the user. Then, the model is computed for each n-tuples of values of the uncertain inputs (in the case where n unknown inputs have been selected by the user). Practically, for each flaw size a set of values for the entry parameters is fixed. Once the execution of the model is achieved, the variable of interest is automatically extracted and displayed in a specific window. A set of options proposes the user to:
- switch from the “signal response” data format to the “hit/miss” one,
- if “signal response” is chosen, select between two different assumptions on the linearity between the signal response and the flaw size,
- select between two different assumptions on the linearity between the log of the odds of the probability of detection and the flaw size.

Depending on the selected assumption, the simulation results are displayed on a plot using the adequate scale (linear or logarithmic), thus giving the possibility to the user to qualitatively validate the linearity assumption. The POD curve is then automatically computed using the adequate model (based on the approach of Berens [0] previously described in the paper) and plotted in the lower part of the window. Examples are given in the next section.

APPLICATION CASES

Ultrasonic inspection of a pipe containing a notch

In the first application case, a stainless steel pipe containing notches is inspected using a standard contact probe (2.25 MHz, diameter of 6.35 mm). The inspection is performed using shear waves refracted at a 45°angle in the specimen. During this inspection, several input parameters may have uncertain values and thus introduce random fluctuations on the result. These uncertainty sources can be either due to insufficient knowledge related to the parameters, or to poor control over them. Here, four uncertain parameters were defined: the notch skew, tilt and length, and the probe orientation. These random parameters are defined in Figure 3. Their statistical properties were described using univariate normal distributions with standard deviations of 1° for the notch skew and tilt, 1° for the probe orientation, and 3 mm for the notch length.

Two different synthetic microstructures were defined. They correspond to stainless steel materials exhibiting microstructures with different mean grain sizes. As it is illustrated in Figure 4, the
simulation results for these various synthetic structures show grain noise levels that vary significantly from one case to the other. The relative noise levels were expressed in terms of signal-to-noise ratios (SNR), with the maximum amplitude of the corner echo on a 10 mm-height notch used as a reference value and the absolute noise levels assessed by averaging the maximum amplitudes of the noise signals measured for each probe position.

By computing the statistical distribution of the maximum amplitudes of the gated rectified noise signals, a detection threshold can be fixed for each synthetic structure. For instance, this can be done by fixing a probability of false alarm to 0.05. The values of the detection thresholds obtained thru this process are closely related to the SNR values estimated previously.

Figure 3 - Inspection of a stainless steel pipe section containing notches located at the inner surface (upper illustration). Description of the uncertain parameters (lower illustrations): the notch skew, length and tilt, and the probe orientation.

![Figure 3](image)

\[ \sigma = 125 \rho = 0.1 \text{ mm}^3 \quad \text{SNR} = 14 \text{ dB} \]
\[ \sigma = 250 \rho = 0.1 \text{ mm}^3 \quad \text{SNR} = 8 \text{ dB} \]

Figure 4 - Simulation of the inspection for various noise levels. The signal-to-noise ratios are expressed with respect to the maximum amplitude of the corner echo on the 10 mm-height notch and by averaging the maximum amplitude of the noise measured for each probe position.
Simulation results presented in figure 5 were obtained with 5 inspections per notch size. The height of the notches ranged from 0.50 mm to 6 mm with steps of 0.5 mm, giving a total of 12 x 5 = 60 values. The only difference between the two results is the noise level which is 6 dB higher for the right plots than for the left ones. Thus, the detection threshold has different values for the two cases. The probability of detection is expressed in terms of the notch height (red curves on the lower plots) and the lower 95 % confidence bound is also shown (blue dotted curves).

**Eddy current inspection of a tube with an axial probe**

As a second application case, perturbed EC signals are calculated in simulation in the case of an axial probe testing a non magnetic steam generator tube of inconel (conductivity 1 MS.m$^{-1}$, inner and outer diameters of 9.84 mm and 11.11 mm, respectively). The probe consists here in two identical coils functioning in differential mode, each coil having a height of 1 mm and inner and outer diameters of 17 mm and 18 mm, respectively. Simulated EC signals correspond to the response of this probe at the frequency of 300 kHz to a cylindrical notch with a depth 0.51 mm and an angular extension of 30 degrees.

POD calculations carried out in this case aim at describing the effect on the signal amplitude of the flaw opening, which varies in a range from 0.1 mm to 1 mm, when taking into account a probe wobble during the inspection. This particular perturbation is simulated using a normal distribution (with a mean of 0 and a standard deviation of 0.1) of the radial probe position and a uniform distribution of the probe angular orientation. In this way a normal variation of the probe position around the tube axis is modelled, as illustrated in figure 6.
Figure 6 - EC inspection of a steam generator tube at the frequency of 300 kHz. A simulated POD curve (in red) with respect to the flaw opening is obtained using the “hit/miss” approach. A small probe wobble is taken into account as a perturbation parameter defined by a normal distribution of the probe position.

After hundred simulations (ten samples of the radial position per opening value and ten opening values in the range [0 1] in millimetres), Hit/miss data corresponding to a threshold of 1.2 mV are generated from EC signals amplitudes and used to estimate the POD curve, as well as the corresponding 95 % confidence interval, see figure 6. Given this detection threshold and the statistical description of the probe wobble, simulated results predict a high probability of detection of the flaw if its opening is bigger than 0.6 mm.

The semi-analytical model used in these simulations is based on the Volume Integral Method [16] that provides fast results when modeling canonical geometries. This time efficiency is especially appreciable for applications like POD that require a lot of calculations. As an example, hundred simulations were carried out for this application case in less than three hours.

CONCLUSION

An approach is proposed in this paper in order to study inspection reliability using simulation results. The UT and EC physical models used in the simulations are those proposed by the software CIVA. The uncertainties on the input parameters are described using parametric statistical distributions. Complex noise sources such as ultrasonic grain noise are computed using specific algorithms. Then, by performing an uncertainty propagation study based on a simple Monte Carlo sampling method, the fluctuations of the signal are estimated.

The Probability of False Alarm can be estimated using the grain noise algorithm to assess the statistical properties of the noise. Inversely, the detection threshold can be determined by first fixing a value for the PFA. The POD curve and the lower confidence bound are then computed using the Berens method.
In the near future, NDE simulation tools such as those proposed in the CIVA software could be an effective and economic way to replace or complement the empirical data that is required today to determine POD curves.

ACKNOWLEDGMENTS

This work was supported by the French National Research Agency (ANR – The SISTAE Project). We are grateful to N. Dominguez from EADS-IW for providing us his numerical implementation of the Berens method for the estimation of the POD curve and its lower confidence bound.

REFERENCES

2) “The Use of In-Service Inspection Data in the Performance Measurement of Non-Destructive Inspections”, RTO-TR-AVT-051, NATO Research and Technology Organisation, March 2005
6) A collection of papers concerning CIVA can be found at http://www-civa.cea.fr
13) Gustafsson M.G. and Stepinski, T, “Studies of split spectrum processing, optimal detection, and maximum likelihood amplitude estimation using a simple clutter model”, Ultrasonics, 35, 31-52
14) Gengembre, N and Lhémyer, A, “Pencil method in elastodynamics: application to ultrasonic field computation”, I, 38, 495-499