ICOSSAR 2021
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021), June 21-25, 2021, Shanghai, P.R. China
J. Li, Pol D. Spanos, J.B. Chen & Y.B. Peng (Eds)

Rethinking the quality of NDE within a unifiying framework

Elizabeth Bismut[a,*] and Daniel Straub[b,*]

[*] *Engineering Risk Analysis Group, Technische Universität München*
[a] *E-mail: elizabeth.bismut@tum.de.*
[b] *E-mail: straub@tum.de.*

ABSTRACT: Probability of detection (POD) curves and receiver operating characteristic (ROC) curves are the two most common measures of the quality of non-destructive evaluation (NDE) in inspection and monitoring applications. These models of NDE performance have historically been developed independently within different research fields. We show with a unifying framework how these and other models are connected to a base model in which both the condition of the structure and the NDE signal are continuous. This framework highlights that the models derived from the base model depend on the experimental design and calibration choices, thus they cannot reflect the base model without losing some information. We show that this information loss, as well as overlooking the impact of experimental design, can lead to a misinterpretation of NDE results, erroneous reliability estimates and suboptimal decision making. The latter can be quantified by the value of information resulting from the application of NDE. We illustrate these effects through a numerical application.

1 MODELS OF NDE QUALITY

Non-destructive evaluation (NDE) is central to monitor the progress of deterioration in aging structures and infrastructure. The information provided by the NDE is typically a partial and inexact reflection of the state of the system. Probabilistic models have been developed to link the *observed signal* output by the NDE and the state of the system (or *condition*) (e.g., Sarkar et al., 1998).

An NDE quality model is defined by the conditional probability $\Pr(observed\ signal|condition)$, which in statistics is known as the likelihood function. This relationship can be derived empirically by performing a number of tests (Packman et al., 1968; Berens and Hovey, 1981; Berens, 2000). The two most commonly used models of NDE quality are probability of detection (PoD) curves and receiver operating characteristic (ROC) curves (see section 2).

Through these models, the performance of NDE techniques can be evaluated (a) with Bayesian analysis to assess the accuracy of the information provided; and (b) with decision analysis to understand the impact of this information on the optimality of mitigating actions (see section 4).

A number of indices (e.g., the Youden index for ROC curves (Youden, 1950)) have been proposed as metrics of the quality of NDE to answer point (a) above. Similarly, the interpretation of the observed signal is typically linked with fixed decision criteria to address point (b) above. However, these indices and criteria do not fully reflect the decision context and the decisions taken on this basis may turn out to be sub-optimal.

In this contribution, we give an overview of the existing types of models for NDE quality, and show how they are connected within a unifying framework (see section 3). Within this framework, we highlight how the models are derived, what choices the analyst face and how they can be optimized, such that better models of NDE are adopted for a specific decision setting. The framework is demonstrated on a case study in section 5. This paper is a summary of (Bismut and Straub, 2021).

2 FOUR MODEL TYPES AND THEIR USE

In the NDE literature, one can distinguish four main types of NDE quality models. They differ with regards to the continuous or binary (discrete) nature of the random variables representing the measured signal and the condition. Table 1 summarizes these model types:

Table 1. Monitoring models for binary or continuous signal and condition

		Condition	
		Continuous X	Binary Y
Signal	Cont. S	(1) $f_{S\|X=x}(s)$	(3) ROC curve
	Bin. I	(2) $PoD(x)$ curve	(4) PoD / PFA

Cont.: continuous. Bin.: binary

Several NDE methods relate a continuous condition with a continuous observed signal, and could be described with the model (1). For example, ultrasonic testing (UT) detects discontinuities inside a metal plate by emitting a high frequency ultrasonic pulse towards the plate and recording the echo. The amplitude of this echo relates to the thickness of defect-free material (Lavender, 1976).

However, PoD curves and ROC curves are predominantly used in the literature instead of the base model, to describe the quality of NDE techniques such as UT, magnetic particle inspection, impulse radar or flooded member detection (e.g. Hovey and Berens, 1988; Sarkar et al., 1998; Feistkorn and Taffe, 2011; Visser, 2002). We refer the reader to Bismut and Straub (2021) for an extensive literature review on the types of model for NDE quality and their origins.

3 THE UNIFYING FRAMEWORK

We have formulated a framework unifying the different models of NDE summarized in Table 1 from the following observation: The connection between the models comes from the fact that *the binary/discrete variables are the result of imposing one or more thresholds on the underlying continuous variables*.

The mathematical formulation of this unifying framework is summarized in sections 3.1 to 3.4. Importantly we show that the links between models requires an understanding of the population of defects in the experimental design. This affects the validity of the NDE models outside of the experimental setting with which they were learned. This has ramifications on the optimal interpretation of data and eventually the decisions taken based on NDE.

Figure 1 shows how the models are linked.

3.1 *Model (1): Base model - S continuous, X continuous*

In the configuration of model (1), the NDE system is fully characterized by the probability distribution of the observed signal S given the condition X, through the conditional probability density function (pdf), $f_{S|X}(s|x)$, or the associated cumulative distribution function (CDF), $F_{S|X}(s|x)$. Model (1) of Figure 1 gives an example of such a conditional pdf such that $S|X$ follows the lognormal distribution with parameters $[\ln\left(2X^3+X^2+10^{-2}\exp(-1/2)\right), 1]$.

Typically, the probabilistic model is obtained from experimental data (experimental test-blocks), if possible in different experimental settings. A traditional approach to find such a model is called "$\hat{a}$ *vs* a", where $\hat{a}$ is the continuous observed signal and a the continuous condition (Berens, 2000). Simulation and meta-models of NDE process have also

ICOSSAR 2021
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021), June 21-25, 2021, Shanghai, P.R. China
J. Li, Pol D. Spanos, J.B. Chen & Y.B. Peng (Eds)

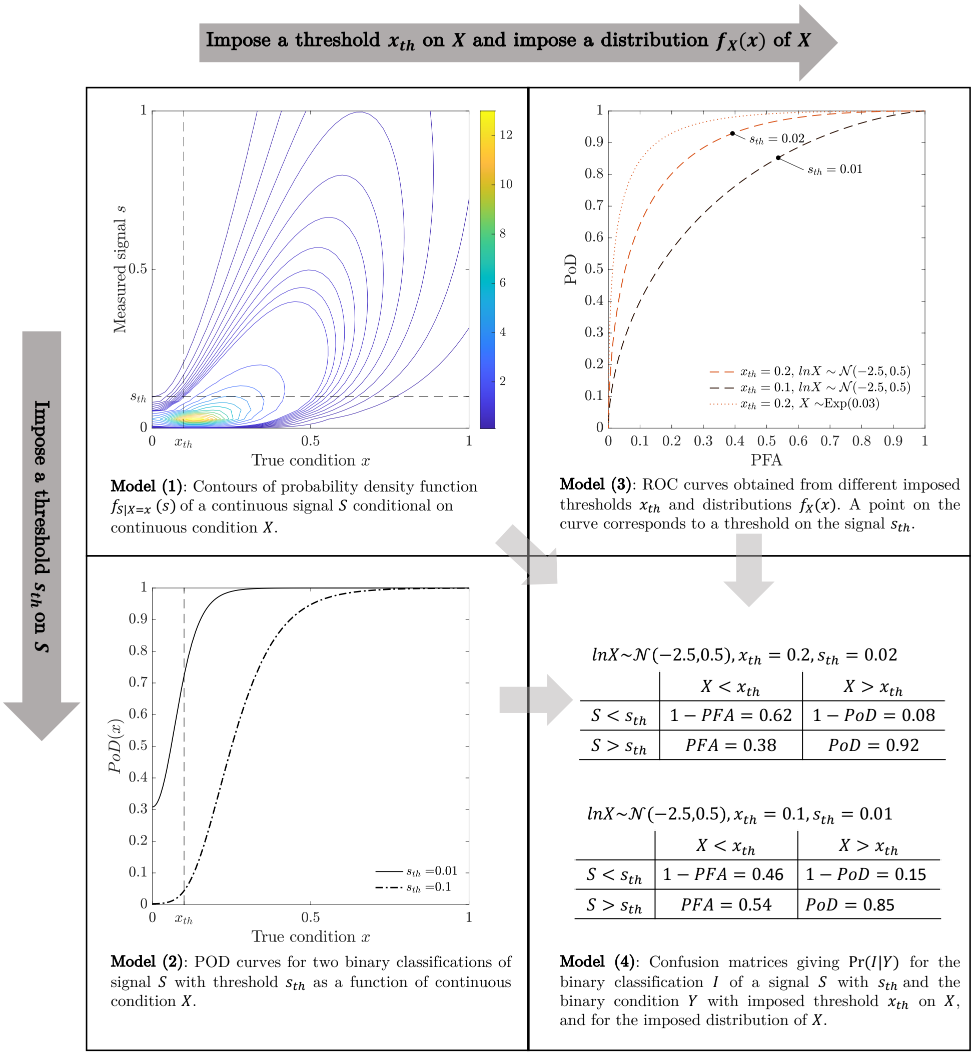

Figure 1. Unifying model of quality of inspection and monitoring. Models (2-4) with a binary signal or condition can in principle be linked to model (1). This link is established by fixing a threshold either on the signal, s_{th}, or on the condition, x_{th}, to classify the continuous signal or condition into binary states. The link between models (1-2) to (3-4) requires additional information on the distribution of the continuous condition X.

given rise to model-assisted PoD (MAPOD) (e.g. Aldrin et al., 2013).

3.2 *Model (2): PoD curves - I binary, X continuous*

The probability of detection curve, or *PoD* curve, is

$$PoD(x) = \Pr(I = 1|X = x). \tag{1}$$

As previously mentioned, one can interpret I as a classification on a continuous signal S by fixing a threshold s_{th}, i.e. $\{I = 1\} = \{S > s_{th}\}$. The PoD function in Equation (1) can thus be written as a function of the continuous/continuous model (1).

$$PoD(x) = \Pr(S > s_{th}|X = x) \tag{2}$$
$$= \int_{s_{th}}^{+\infty} f_{S|X}(s|x)\mathrm{d}s = 1 - F_{S|X}(s_{th}|x)$$

By changing the threshold s_{th}, the PoD curve changes; at the limit, it is $PoD(x, s_{th} = -\infty) = 1$ and $PoD(x, s_{th} = +\infty) = 0$ for any value x.

We note that Equation (2) does not preclude the PoD curve from taking a non-zero value when $x = 0$. A PoD curve for which $PoD(0) = 0.3$ is depicted in Model (2) of Figure 1 by fixing $s_{th} = 0.01$. It includes the possibility of false detection, which is needed to quantify the impact of unnecessary repairs (Heasler and Doctor, 1996; Straub, 2004).

ICOSSAR 2021
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021), June 21-25, 2021, Shanghai, P.R. China
J. Li, Pol D. Spanos, J.B. Chen & Y.B. Peng (Eds)

3.3 *Model (3): ROC curve - S continuous, Y binary*

This model is commonly used to describe how the continuous observed signal S from NDE can be interpreted to discriminate between the absence ($Y=0$) and presence ($Y=1$) of a flaw.

In this framework, the discrete condition Y is defined by setting a threshold on the continuous condition X such that $\{Y=1\}=\{X>x_{th}\}$ and $\{Y=0\}=\{X\leq x_{th}\}$. The two conditional pdfs of the signal S associated with NDE quality model type (3) can be derived from Model (1):

$$f_{S|Y=1}(s)=\frac{1}{1-F_{X,exp}(x_{th})}\cdot\int_{x_{th}}^{+\infty}f_{S|X}(s|x)f_{X,exp}(x)\mathrm{d}x, \quad (3)$$

$$f_{S|Y=0}(s)=\frac{1}{F_{X,exp}(x_{th})}\cdot\int_{-\infty}^{x_{th}}f_{S|X}(s|x)f_{X,exp}(x)\mathrm{d}x. \quad (4)$$

This model is commonly visualized by plotting the corresponding receiver (or relative) operating characteristic (ROC) curve (see Figure 1). This curve is parametrized by a threshold on the signal s_{th}, also called the cut-off point (Fluss et al., 2005). The *PoD* and *PFA* in function of s_{th} are

$$PoD(s_{th})=\Pr(S>s_{th}|Y=1)=\int_{s_{th}}^{+\infty}f_{S|Y=1}(s)\mathrm{d}s \quad (5)$$

$$PFA(s_{th})=\Pr(S>s_{th}|Y=0)=\int_{s_{th}}^{+\infty}f_{S|Y=0}(s)\mathrm{d}s \quad (6)$$

One can alternatively express the *PoD* and *PFA* as a function of the conditional CDF $F_{S|X}$ from model (1):

$$PoD(s_{th})=\Pr(S>s_{th}|X>x_{th})=\frac{1}{1-F_X(x_{th})}\cdot\int_{x_{th}}^{+\infty}\left(1-F_{S|X}(s_{th}|x)\right)f_X(x)\mathrm{d}x, \quad (7)$$

$$PFA(s_{th})=\Pr(S>s_{th}|X<x_{th})=\frac{1}{F_X(x_{th})}\cdot\int_{-\infty}^{x_{th}}\left(1-F_{S|X}(s_{th}|x)\right)f_X(x)\mathrm{d}x. \quad (8)$$

Equations (3), (4), (7) and (8) show that the *PoD* and *PFA* on the ROC curve are a function of the distribution of the condition X. This signifies that even if the ROC curve is evaluated directly from experiments, it is only strictly valid for the distribution of the defects from which it is derived. Therefore, the ROC curve for the same NDE method can vary when applied to different situations. Model (3) in Figure 1 illustrates how different ROC curves can be obtained from the same base model (1).

3.4 *Model (4): PoD/PFA - I binary, Y binary*

This is the most elementary NDE quality model associated with an NDE, which identifies whether the system is in a certain state or not. The associated likelihood is described by a confusion matrix, involving the operating *PFA* and *PoD*, as presented in Figure 1 Model (4).

The transition between model (3) to model (4) corresponds to calibrating the NDE system to an operating point on a ROC curve, by fixing the threshold s_{th}. The *PoD* and *PFA* are obtained from Eqs. (5) and (6). A system calibrated with a too low threshold will catch most failures but will also provoke a lot of false alarms. Conversely, a monitoring system calibrated with a too high threshold will have less alarms overalls but will likely miss a lot of component failures.

The transition from model (2) to (4) is obtained by recognizing the expression for the PoD curve of Eq. (2), $PoD(x)$, in Eqs. (7) and (8). The *PoD* and *PFA* for model (4) are

$$PoD=\Pr(I=1|X>x_{th})=\frac{1}{1-F_X(x_{th})}\cdot\int_{x_{th}}^{+\infty}PoD(x)f_X(x)\mathrm{d}x \quad (9)$$

$$PFA = \Pr(I = 1|X < x_{th}) = \frac{1}{F_X(x_{th})} \cdot \int_{-\infty}^{x_{th}} PoD(x) f_X(x) \mathrm{d}x \quad (10)$$

4 QUANTIFYING THE IMPACT OF NDE ON THE DECISION PROCESS

Section 3 shows that several choices are made when modeling the performance of an NDE system. There is no unique description of NDE quality and the quality models depend on threshold values on the measured signal and on the condition, s_{th} and x_{th}. Ultimately, the interpretation and impact of NDE results depend on these choices. For this reason it can be beneficial to scrutinize these choices by a formal analysis.

4.1 *Bayesian analysis*

With Bayesian analysis, the posterior probability distribution of the condition $\mathbf{\Theta}$ is

$$p(\boldsymbol{\theta}|z) \propto \mathscr{L}(\boldsymbol{\theta};z)p(\boldsymbol{\theta}), \quad (11)$$

where $\mathbf{\Theta}$ is either X or Y depending on the setting. z is the measurement, which is either s or i. $\mathscr{L}(\boldsymbol{\theta};z)$ is the likelihood function, i.e., one of the NDE quality models of section 2.

4.2 *Decision analysis and value of information*

In decision analysis, it is assumed that the decision maker selects an action that maximizes the expected utility after obtaining information Z through NDE. Here we consider utility to be negatively proportional to costs. Hence, the optimization problem is written as

$$a_{opt}(z) = \arg \min_{a \in \{a_0, a_1 \dots\}} \mathbf{E}_{\mathbf{\Theta}|z}[C_T(a, \mathbf{\Theta})], \quad (12)$$

where $a_0, a_1 \dots$ are the available actions. The total cost C_T includes the cost of the actions and consequences of failure. $\mathbf{E}_{\mathbf{\Theta}|z}[\cdot]$ is the expectation with respect to the conditional distribution $p(\boldsymbol{\theta}|z)$ from Equation (11).

An optimal action $a_{opt,0}$ can also be obtained for the so-called prior case, where no information is collected. The value of information (VoI) for an NDE system is the difference between the total expected prior cost associated with implementing $a_{opt,0}$ and the total expected cost associated with implementing the maintenance strategy $a_{opt}(z)$. Thus

$$VoI = \mathbf{E}_{\mathbf{\Theta}}[C_T(a_{opt,0}, \mathbf{\Theta})] - \mathbf{E}_Z\left[\mathbf{E}_{\mathbf{\Theta}|z}[C_T(a_{opt}(z), \mathbf{\Theta})]\right] \quad (13)$$

5 CASE STUDY

In the following, we consider a basic decision problem, where one needs to decide on a repair action. We consider an NDE system which measures a crack depth and outputs a continuous signal. We adopt the four different NDE quality models derived from section 3 in turn and assess their impact on the repair decision. We also investigate how recommended NDE model calibrations (e.g., Berens, 2000) perform for this specific case study.

5.1 *Crack detection and repair*

We consider the case of a one-step decision process, shown by the influence diagram in Figure 2. Cracks can be present on the component. The decision-maker can either repair the component or do nothing. Undetected cracks with a depth larger than $x_{cr} = 1$[mm] correspond to a failure. If the component is repaired, there are no more cracks. Under the

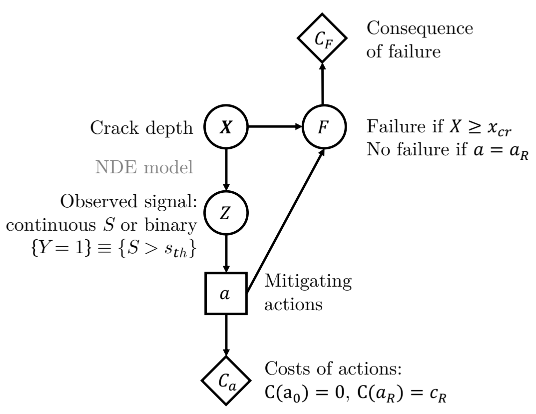

Figure 2. Influence diagram of the one-step decision process.

do-nothing action a_0, at cost €0, the conse-

ICOSSAR 2021
The 13th International Conference on Structural Safety and Reliability (ICOSSAR 2021), June 21-25, 2021, Shanghai, P.R. China
J. Li, Pol D. Spanos, J.B. Chen & Y.B. Peng (Eds)

quence of the crack depth exceeding the critical value $X > x_{cr}$ is $c_F = €1 \cdot 10^4$. If the component is repaired, under action a_R, the only cost incurred is that of the repair, $c_R = €1 \cdot 10^2$. The prior pdf of the true crack depth X is the exponential pdf with mean 0.15[mm].

The crack depth can be measured using NDE. Here, the relationship between the continuous observed signal S and the true crack depth X is modeled such that $\log(S) = \log(X) + \varepsilon$, with $\varepsilon \sim Normal(0, 0.5)$. This model structure is typical for the $\hat{a}$ *vs* a method (Berens, 2000).

The a priori optimal action $a_{opt,0}$, when the crack is not measured, is the one that minimizes the total expected cost. Here,

$$c_F \Pr(X \geq x_{cr}) = 10^4 \cdot 1.3 \cdot 10^{-3} = €12.7 \qquad (14)$$
$$\leq c_R = €100,$$

therefore, the a priori optimal action is to do nothing, i.e., $a_{opt,0} = a_0$.

When we consider the NDE system outcome, the a posteriori optimal action for a measured signal z is a_R if

$$c_R \leq c_F \Pr(X \geq X_{cr} | Z = z), \qquad (15)$$

where $\Pr(X \geq X_{cr}|Z = z)$ is obtained from Equation (11).

5.2 *Optimal actions using the base model*

By solving Eq. (12) for $z = s$, we find that the optimal action is $a_{opt}(s) = a_0$ for $s < 0.74$, and $a_{opt}(s) = a_R$ for $s > 0.74$. The expected total cost computed as $\mathbf{E}_Z \left[\mathbf{E}_{\mathbf{\Theta}|z}[C_T(a_{opt}(z), \mathbf{\Theta})]\right] = €5.1$. The VoI of this NDE system is computed with Equation (13) as $12.7 - 5.1 = €7.6$.

5.3 *Optimizing the PoD curve*

PoD curves are typically given and not specific to the application, and hence are likely suboptimal for a given decision context. For example, a standard calibration of the PoD curve is such that the critical damage (here x_{cr}) is detected with a 90% probability (Berens, 2000). From model one, $PoD(x_{cr}) = 90\%$ is obtained with $s_{th} = 0.53$. It is depicted in Figure 3. For this calibration, the optimal actions are $a_{opt}(I = 0) = a_0$, $a_{opt}(I = 1) = a_R$ and the expected cost is €6.67. The resulting the VoI is $12.7 - 6.7 = €6$, which is below the potentially achievable VoI = €7.6.

For each imposed threshold s_{th} and associated PoD curve, Figure 4 indicates the optimal actions $a_{opt}(I)$ and the expected cost $\mathbf{E}_I \left[\mathbf{E}_{\mathbf{\Theta}|z}[C_T(a_{opt}(I), \mathbf{\Theta})]\right]$. A threshold exists that maximizes the VoI and minimizes the expected cost for this case study. It is the decision threshold 0.74, with the associated expected cost of €5.1, which corresponds to the result obtained with model (1). However, when describing the NDE by the PoD curve, this optimal solution will only be obtained by coincidence. In general, the PoD curve leads to a suboptimal decision.

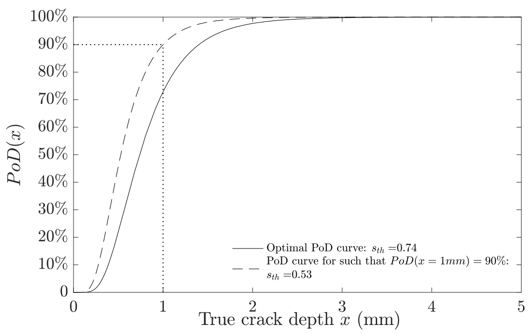

Figure 3. PoD curves for thresholds $s_{th} = 0.53$ and $s_{th} = 0.74$.

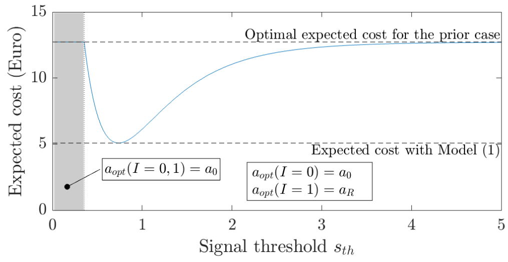

Figure 4. Expected cost using the PoD curve model as a function of the fixed signal threshold s_{th}.

5.4 *Performance of a typical ROC index*

In this decision problem, the analyst is interested in identifying critical cracks such that

$X \geq x_{cr}$. The binary condition Y such that $Y = 1$ corresponds to $X \geq x_{cr}$ is naturally defined. The NDE models (3) and (4) corresponding to this binary condition can be derived. In particular, the ROC curve can be obtained from model (1) above with $x_{th} = x_{cr}$ and $f_X(x)$ is the prior pdf of the crack depth X. The ROC curve is shown in Figure 5.

Indicators have been proposed to qualify the performance of an NDE system with a ROC curve, such as the *Youden index* (Fluss et al., 2005), which is computed as the maximum vertical distance between the ROC curve and the $PoD = PFA$ line starting at $(0,0)$. This index is also typically used to calibrate NDE systems and devise maintenance strategies. We investigate the performance of the NDE system when calibrated on the Youden index. The point corresponding to the Youden index is depicted on the ROC curve of Figure 5, and corresponds to $s_{th} = 0.46$, with $PFA = 0.08$ and $POD = 0.96$. The optimal actions for $I = 0$ and $I = 1$ for this point on the ROC curve are found as a_0 and a_R, respectively. The expected cost is €8.2, with a VoI of $12.7 - 8.2 =$ €4.5.

However, the Youden index is not the best operating point on the ROC curve. To demonstrate it, we plot the optimal expected cost is computed for any pair PFA, PoD in Figure 5. The point on the ROC curve, which minimizes the expected cost, corresponds to $s_{th} = 0.74$, $PFA = 0.02$, $PoD = 0.80$, with the expected cost €5.1. This threshold and the associated actions corresponds here to the optimal threshold found for the base model (1). In a decision problem where the failure consequences depend on the value of the true condition X, the ROC curve model may not perform as well as models (1) or (2).

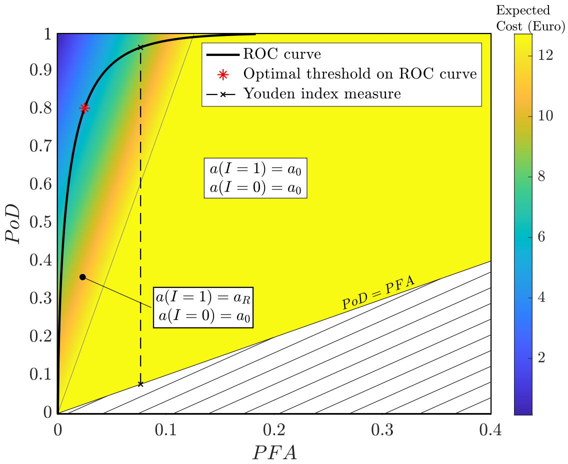

Figure 5. ROC curve for $x_{th} = x_{cr} = 1$[mm] and optimal expected cost as a function of *PoD* and *PFA*. The point on the ROC curve corresponding to the Youden Index is depicted. In this decision context, it does not coincide with the optimal threshold on the ROC curve, also shown. The optimal actions associated with the outcome $I = 0$ and $I = 1$ are indicated for each value of *PoD* and *PFA*. The yellow zone corresponds to a null VoI, i.e., the optimal decision for an NDE system defined by a *PoD* and *PFA* falling within this zone is the same as the prior decision.

6 CONCLUDING REMARKS

In this contribution, we clarify the connection between the different models of NDE quality, such as PoD curves or ROC curves, through a base model in which both the condition and the NDE outcome are modeled continuously. The models with a binary observed signal or binary condition are derived by imposing thresholds on the observed signal and the condition. For models that use a binary condition, one must also impose a distribution on the underlying continuous condition.

In the existing literature, continuous/continuous probabilistic models are often learned in an ad-hoc manner – such as the probability of (correct) sizing (POS) (e.g., da Silva and de Padua, 2012)–, but no application to reliability analysis is documented. For many NDE techniques, the base model (1) remains abstract, as a continuous signal or a continuous condition might not be easily identifiable. In this case, the NDE quality model is chosen among the other three categories. It might not always be possible to reveal a continuous condition or a continuous signal, and only one of the other models (2) to (4) might be identifiable. Still, the base model linking X to S can be considered at an abstract level to ensure correct interpretation of the signal and good experimental design.

With the case study, we demonstrate that calibration of an NDE system in function of the decision settings (e.g., cost model) is beneficial, and that recommended calibration points following indices or standard NDE reliability requirements can result in sub-optimal actions. Bayesian decision analysis provides the means to compute the VoI, which allows for a direct comparison between NDE systems. This analysis gives the opportunity to calibrate an NDE system to suit the decision parameters such as the cost of mitigating actions and the expected consequences of failure.

REFERENCES

Aldrin, J.C., Knopp, J.S., Sabbagh, H.A., 2013. Bayesian methods in probability of detection estimation and model-assisted probability of detection evaluation, in: AIP Conference Proceedings, American Institute of Physics. pp. 1733–1740.

Berens, A.P., 2000. Probability of Detection (POD) Analysis for the Advanced Retirement for Cause (RFC)/Engine Structural Integrity Program (ENSIP) Nondestructive Evaluation (NDE) System-Volume 1: Pod Analysis. Report. University of Dayton Research Institute.

Berens, A.P., Hovey, P.W., 1981. Evaluation of nde reliability characterisation. Air Force Wright Aeronautical Laboratories, AFWAL-TR-81-4160, November 1981 .

Bismut, E., Straub, D., 2021. A unifying review of NDE models towards optimal decision support. Structural Safety. Under review .

Feistkorn, S., Taffe, A., 2011. Die POD-eine Vorgehensweise zum qualitativen Gütenachweis zerstörungsfreier Prüfverfahren im Bauwesen am Beispiel des Impulsradars, in: DGZfP-Jahrestagung 2011, Bremen, Germany.

Fluss, R., Faraggi, D., Reiser, B., 2005. Estimation of the Youden Index and its associated cutoff point. Biometrical Journal: Journal of Mathematical Methods in Biosciences 47, 458–472.

Heasler, P.G., Doctor, S.R., 1996. Piping inspection round robin. Report. U.S. Nuclear Regulatory Commission, Washington, DC (United States).

Hovey, P.W., Berens, A.P., 1988. Statistical evaluation of NDE reliability in the aerospace industry. Springer. pp. 1761–1768.

Lavender, J.D., 1976. Ultrasonic testing of steel castings. Steel Founders' Society of America.

Packman, P.F., Pearson, H.S., Owens, J.S., Marchese, G.B., 1968. The Applicability of a Fracture Mechanics-Nondestructive Testing Design Criterion. Report. Lockheed-Georgia Company Material Sciences Research Laboratory.

Sarkar, P., Meeker, W.Q., Thompson, R.B., Gray, T.A., Junker, W., 1998. Probability of detection modeling for ultrasonic testing, in: D.O. Thompson, D.C. (Ed.), Review of Progress in Quantitative Nondestructive Evaluation, Springer. pp. 2045–2052.

da Silva, R.R., de Padua, G.X., 2012. Nondestructive inspection reliability: state of the art. Nondestructive Testing Methods and New Applications , 1.

Straub, D., 2004. Generic approaches to risk based inspection planning for steel structures. volume 284. vdf Hochschulverlag AG.

Visser, W.P., 2002. POD/POS Curves for Non-Destructive Examination. Report Offshore Technology Report 2000/018. HSE Books, Norwich, United Kingdom.

Youden, W.J., 1950. Index for rating diagnostic tests. Cancer 3, 32–35.