Probabilistic modelling of inspection results for offshore structures

Antoine Rouhan, Franck Schoefs

Probabilistic modeling of inspection results for offshore structures

A. Rouhan[a], F. Schoefs[b]

[a]*Bureau Veritas, Marine division, Research department, 17 bis, place des Reflets, La Défense 2, 92400 Courbevoie, France*
[b]*Laboratoire de Génie Civil de Nantes Saint-Nazaire, 2 rue de la Houssinière, BP 92208, 44322 Nantes cedex 3, France*

Operators of marine structures have to ensure that structural integrity is maintained at a sufficient level during in-service life or in the case of the structure life prolongation. This can be achieved by Inspection, Maintenance and Repair plans (IMR), as rational aid-tools for decisional purposes. Such plans are complex and can be expensive. This leads to their global optimization, particularly regarding inspections. In this context, original results on inspections data in IMR plans are presented. The approach is based on decision and detection theories and include both the probability of false alarm and the probability of detection. It is shown how to introduce these probabilities in a decision scheme. The effect of false alarms and miss detections on the global cost of inspection planning is underlined through a basic example.

Keywords: Existing structures; Probability of detection; False alarm; Cost analysis; IMR plans; In-service inspections; Risk Based Inspections; Non destructive testing

1. Introduction

During the 20th century, new techniques, materials and methods of marine structures construction have given a sense to the concept of durability for an increased number of constructions. Consecutively, design codes and building processes have progressed to ensure a structural life time, making risk more and more explicit. Considering the structural aspect, not the equipment maintenance, these structures have been designed for a predefined period from 20 to 50 years or more, depending on their functions. After these few decades, most of them seems to be in good condition as they do not present large damages. Part of them are considered sensitive due to their impact on human life or environmental preservation and are asked for integrity certifications during this life time. Moreover, after this period most of them are suggested for a life time extension as an economical alternative to destruction. In fact, structures such as bridges, dams, harbors and offshore structures play a major role in these requalification procedures.

Such existing civil and marine engineering structures have to be monitored during their whole in-service lives. The challenge is then (i) to provide strategy for repairs or replacements of damaged components (ii) to inspect the most critical structural parts (iii) to maintain the structure for its operating functions and security requirements. This is the so-called IMR (Inspections, Maintenance, Repairs) plan. For very large structures (or with a significant number of components with potential failure), there is a need to optimize these plans, in terms of costs and performances [1]. The optimal inspection plan would be to inspect at the right place, at the right time, and with the right tool, at lowest cost. It should be done according to a predefined safety level requirement. Recently, there is an increasing need for such plans in marine structures, and particularly for end-of-live structure re-assessment.

In this paper, we will focus on the inspection aspects based on non destructive testing, which is a decision problem from the detection theory point of view. Inspections will be first described from a theoretical point of view in an attempt to carefully define fundamental terms generally used. It is shown that the probability of detection (abbreviated PoD) and the probability of false alarm (abbreviated PFA) are both needed to characterize completely an inspection tool, or inspection results. Then, in the IMR plan context, this set of variables is introduced in an optimization scheme. Effects of poor inspection performances in terms of costs are shown through an example, showing the importance of false alarms. It should be emphasized that false alarms have to be taken into account when introducing inspection results as it may introduce false failure scenario. The reliability of the structure if often evaluated through a decision tree [2]. Hence, a false alarm may lead to the prospection of fault scenarios that actually do not have significant structural importance. Thus, the inspections and repairs planned would be useless, leading to high cost overrun. In this paper, jacket offshore structures and underwater inspections are chosen for illustration.

2. Assessment of global inspection performances

2.1. Global IMR strategy challenge

For risk sensitive marine structures, IMR actions are planned for the whole life of the structure. This means that at the design stage one may predict:

- localization, dates and/or frequencies of inspections and tools used depending on their performances and ability to be used in situ;
- dates and/or frequencies of structural maintenance; and
- strategies of component repairs and/or changes if possible and associated decision criteria.

This leads to an overall cost estimation and operations planning. For in-service structures, the problem is specific and one may consider two cases:

- Numerous damage localizations exist on the structure and one should optimise inspection areas in terms of structural integrity assessment and risk analysis.
- After inspections, unpredicted damages are discovered and the initial plan has to be updated to ensure the safety level requirements. This leads to a general structural safety reassessment.

For in-service conditions, Fig. 1 represents a simplified IMR plan with the connected decision scheme for offshore structures in the case of fatigue behavior which is one of the most important damage for jacket structures: fatigue crack growth in tubular nodes [3,4]. Another one is failure under extreme environmental conditions. Maintenance is not represented, since it is assumed to be done during inspections. This figure illustrates the complexity and the multidisciplinary of the IMR decision context. It leads to consider the informations collected on an in-service structure as the input of successive steps of modeling. To this aim, it is now well established that the probabilistic mechanics approach gives an efficient quantitative means for updating informations from inspections and for measuring the relative changes in safety level compared with a predefined requirement.

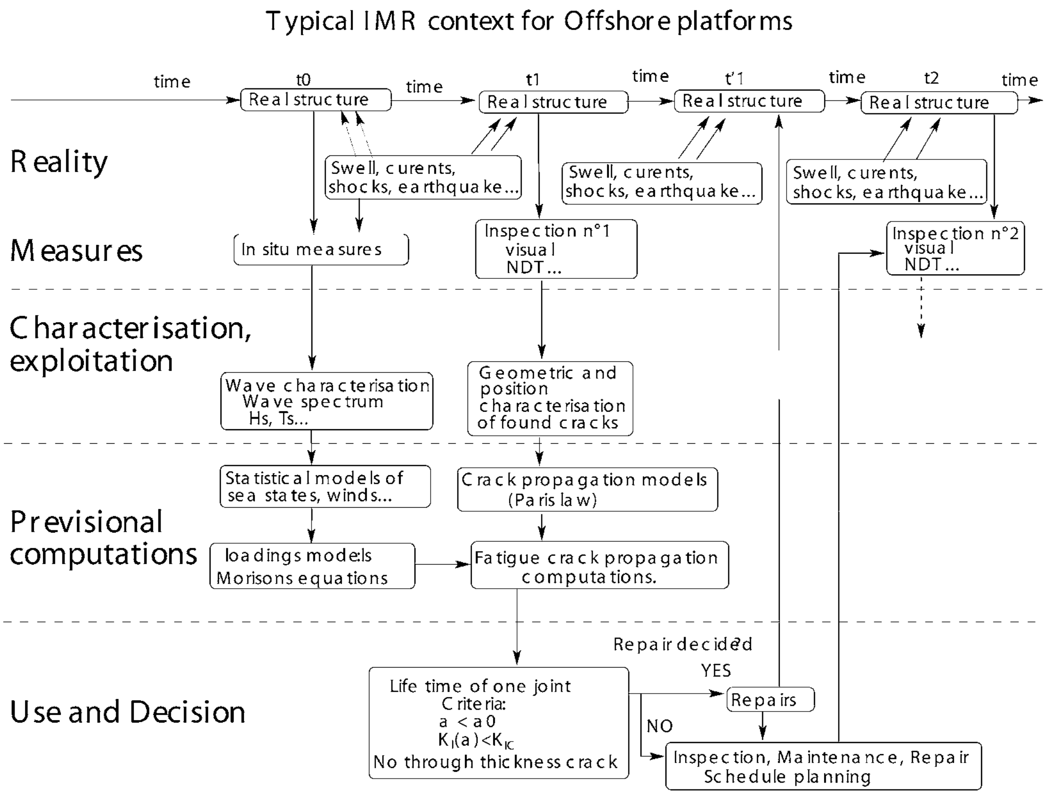

Fig. 1. Illustration of a simplified IMR plan for offshore structures.

2.2. Requirement of NDT tools in IMR plans: need of a probabilistic modeling

Inspection is an essential step in IMR plans, since it's the only way to achieve a partial view of the structural integrity. A complete overview can't be reached due to the size of monitored structures. This underlines the importance of the selection of local areas to be inspected on large structures. An optimal inspection is located where damages should be critical versus a risk criterion and is done with the right inspection tool in the sense of the cost/performance ratio. On existing structures, non-destructive testing (abbreviated NDT) are widely used. Classically, there are two levels of analysing NDT performances: the sizing and the detection capacity. In this paper, the focus will be on the last one (i.e. crack detection) as crack is a very common damage in steel jacket platforms. Moreover cracks can be representative of a significant loss of structural local integrity in the case of large cracks. All NDT tools have limitations and, in complex environment and harsh conditions, their capabilities and abilities to be well operated are different from those given by laboratories and/or factories [5–7], even if a protocol is rigorously followed during inspection. This is the case for underwater inspections of offshore structures where accessibility is limited and conditions of use of the NDT tool are not optimal. This leads to lower performances than expected. In the offshore field, an important work of inter-calibration was made within the ICON project [5,8,9], in order to have an unified overview of several tool performances in realistic in situ conditions. All the data performances where introduced into a single database. The decision-maker has then very powerful informations to decide which best NDT tool to use, relatively to his performances, for a specific application. This allows an optimal choice of different tools in order to use them at their full capabilities. Specifying NDT tools ranking criteria is very difficult in this complex and multi-disciplinary context. It should be based on a detailed analysis of needs and performances.

2.3. NDT performances assessment

One challenge in the IMR plan strategy is to use the whole information existing on NDT performances to optimize their use. Most of the time, inspection results only deal with the probability of detection, which is the probability to detect an existing crack [10]. Let a_d be the minimal crack size, under which it is assumed that no detection is done. Thus, the probability of detection is defined as [11]:

$$\mathrm{PoD}(a) = P(a > a_d) \qquad (1)$$

where a is the crack length. In a probabilistic scheme, a and a_d are stochastic variables, see Refs. [1,11–14]. However, in this paper crack size will not be discussed, as only detection[1] will be focussed on. Fig. 2 shows two different typical PoD curves. The theoretical one is continuous and monotonically increasing, as it is a probability distribution function. The experimental PoD curve is discrete, and not necessarily monotonically increasing. Each point is representative of a crack class range, and the probability of detection in that class is the number of actual detected cracks divided by the total number of existing cracks in that class. As a consequence, it is not necessarily an increasing function. In particular, such a curve is representative of complex tubular joint

[1] Note that, unless expressed, the term *crack detection* will be used for the detection of one crack, with a size included in a given class range. Crack classification into classes is common.

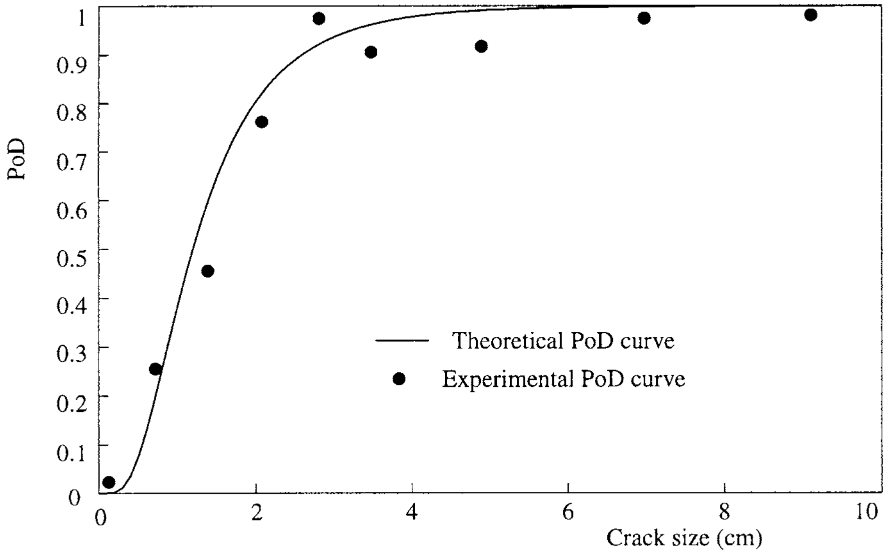

Fig. 2. Typical experimental and theoretical PoD curves.

inspections, with inclined braces on the chord. As some area of such nodes are less easy to access (and because of the lack of samples in the database), the inspection performance decreases [5]: the typology and accessibility of the joint have a great influence on the PoD. This shows that the first PoD model [Eq. (1)] is not satisfactory for inspection data use, particularly for structural safety evaluation. Nevertheless, the use of the PoD alone is not suitable. Another information which is called probability of false alarm (abbreviated PFA) is to be considered. False alarms correspond to detection of non existing cracks. It is induced especially by noise with several possible sources: human, nature of phenomenon to be measured, environmental conditions and so on. It is important to use PFA (see [15,16]), as for underwater technology in offshore structures for example, finding a non existing crack leads to false scenario in the failure tree (changes in the reliability system analysis) and to useless repairs, resulting in a non negligible cost overrun. In harsh in-situ conditions, false detections increase dramatically and non existing large cracks can be detected as demonstrated by the results of the ICON project. In the next section, both PoD and PFA will be introduced in the detection theory context.

2.4. NDT tools ranking from inspection performances

In order to optimize costs and IMR plans, one should rank NDT tools in terms of cost and performances. A first way to achieve, is to plot on the same graph, their receiver operating characteristic (abbreviated ROC) curve. This curve is basically the PoD plotted as function of the PFA [17,18]. From a theoretical point of view, this is a convex, monotonically increasing function, always lying above 45° diagonal of the ROC space, and its first derivative is closely linked to the sensitivity of the receiver, see [17,19]. The diagonal line running from lower left to upper right (curve "PoD = PFA") is the line of no "performance", since in that case the inspection result is the same, no matter what the observation is (see demonstration in Section 3.2).

Looking for the best detection performances, the probability of detection should always take larger values than the probability of false alarm (low noise sensitivity). We have then PoD $\geqslant$ PFA. When reading ROC curves, one must remind that the probability of false alarm depends on the noise and detection threshold only. It does not depend on crack size. The probability of detection is a function of the detection threshold, the crack size, and the noise. Thus, for a given detection threshold, the probability of false alarm is a constant, but the probability of detection is an increasing function of the crack size (see Fig. 2). The ROC curve is a fundamental characteristic of the NDT tool performance. The perfect tool is represented by a ROC curve reduced to a single point whose coordinates in the (PoD, PFA) plane are: (PoD, PFA) = (1,0). Fig. 3 presents four different theoretical ROC curves, corresponding each one to different NDT tool performances. The worst curve is the one with the lowest signal/noise ratio (s/n = 1.0), meaning that some noise can be easily detected, even if nothing is to be detected (no crack presence) and finally leading to a high number of false alarms. At the same time, the PoD is small for low probabilities of false alarms. Overall performances are poor. At the opposite, the best plotted ROC curve is the one with a high signal/noise ratio (s/n = 5.0). Differences with the previous curve are considerable. The probability of detection reaches very quickly values near 1, with small probabilities of false alarms for high values of the PoD. Overall performances are very good. To illustrate the noise ratio and detection threshold effects, four detection cases (here a Gaussian additive signal with a Gaussian noise) are plotted on Fig. 4, and the corresponding points reported on the ROC curves

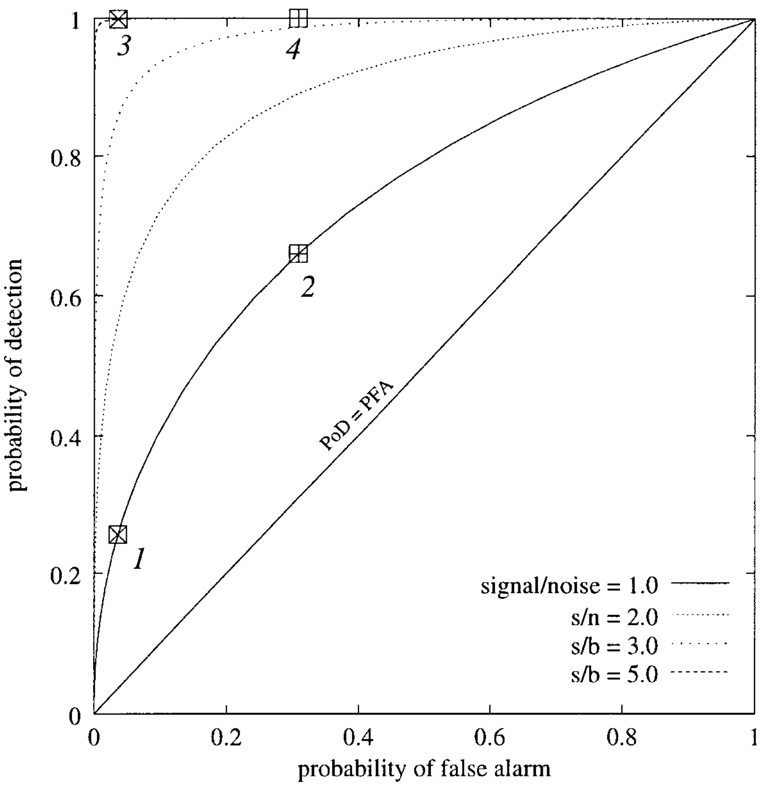

Fig. 3. Example of ROC curves with several signal/noise ratios.

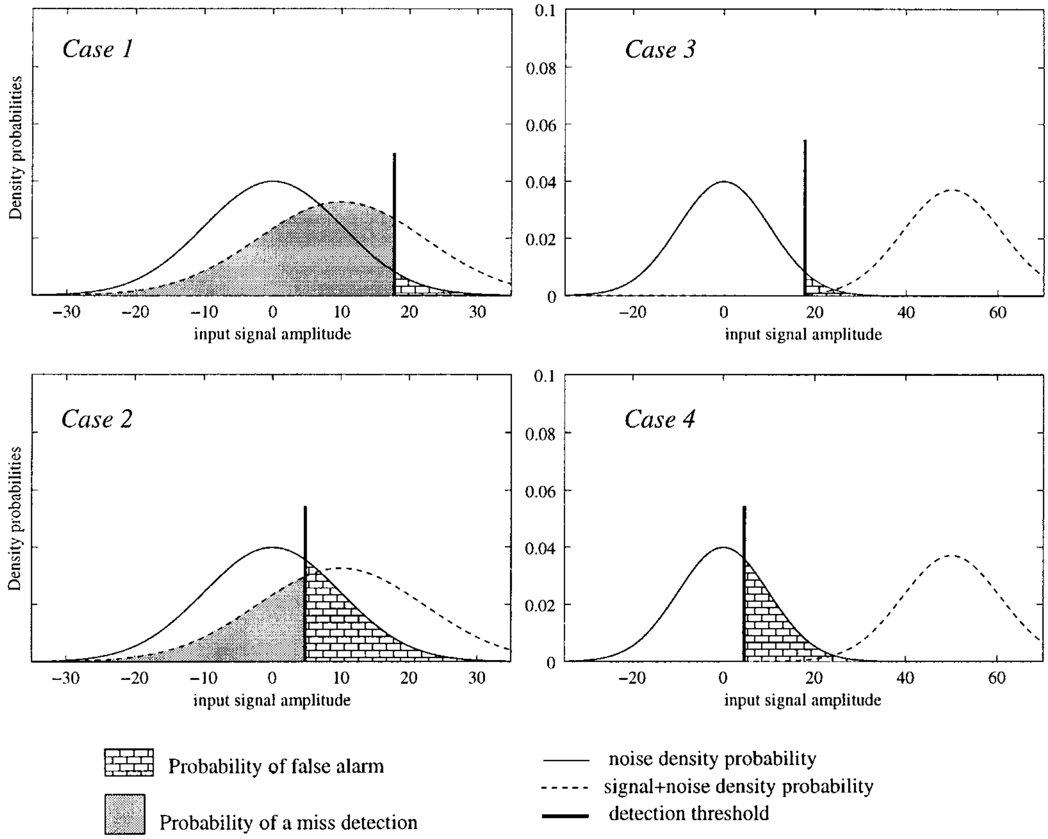

Fig. 4. Detection cases with two signal/noise ratio.

Fig. 3. In these figures, the probability of detection is the probability that the signal "noise+-signal" is greater than the detection threshold and the probability of false alarm is the probability that the signal "noise" is greater than the detection threshold. Cases one and two do have the same low signal/noise ratio but a different detection threshold; they are on the same ROC curve. Cases one and three do have the same detection threshold but a different signal/noise ratio; they have the same probability of false alarm. Cases three and four have the same signal/noise ratio (higher than for cases one and two) but a different detection threshold; they are on the same ROC curve. Cases two and four have the same detection threshold (smaller than for cases one and three) but a different signal/noise ratio; they have the same probability of false alarm. Case one has bad performances, with a very low probability of detection and false alarm. Cases two and four seem to be better with a much higher probability of detection, but are sensitive to noise leading to a high probability of false alarm. Best results are provided by case three, with a low probability of false alarm and a high probability of detection. These cases show clearly that to get good inspections results, a high signal/noise ratio and a well adapted detection threshold are needed.

Considering the crack size and assuming first that a large crack size leads to a high signal/noise ratio, whereas a small crack size gives a small ratio, second that the noise and the threshold are constant, a simple measure of the performance of a NDT tool on a ROC chart is the area under the curve. Its maximum value is 1 and null performance is given by 0.5. Ranking NDT tools can be achieved by the comparison of the area under the ROC curves. Nevertheless, these are theoretical curves, and real ROC curves are not so smooth since it is very hard to obtain regular results with poor data measurements. Otherwise, this is a way to compare two NDT techniques or two operator capacities if the same tool is used.

In the offshore field, the qualification and ability of inspectors to succeed in good detections are essential, and the overall performances of the tool and the diver can be evaluated through ROC curves. Examples of practical results showing the effect of the inspector's experience on the ROC curve shape can be found in [19,20]. The use of this curve as a basis for NDT ranking will be further presented.

3. Implementation of inspection performances in the decision process

3.1. Probabilistic modeling of NDT detection

An inspection is a decision problem: to make an inspection is equivalent to make a decision. To illustrate this, lets consider a typical crack detection problem, see Fig. 5 [19]. Assume we have to detect an existing crack in a body (here a structural offshore tubular node), with a specific NDT tool. After inspection, the NDT result could be: no crack, or presence of crack. In fact this primary result should be interpreted through a decision on the state of the body: cracked or not. The same scheme could be applied if the body is actually not cracked. As for in-service structures the state of the inspected area is not known, it is thus necessary to consider four inspections events:

- E_1: no presence of crack, conditional to no crack detection;
- E_2: no presence of crack, conditional to crack detection;
- E_3: presence of crack, conditional to no crack detection; and
- E_4: presence of crack, conditional to crack detection.

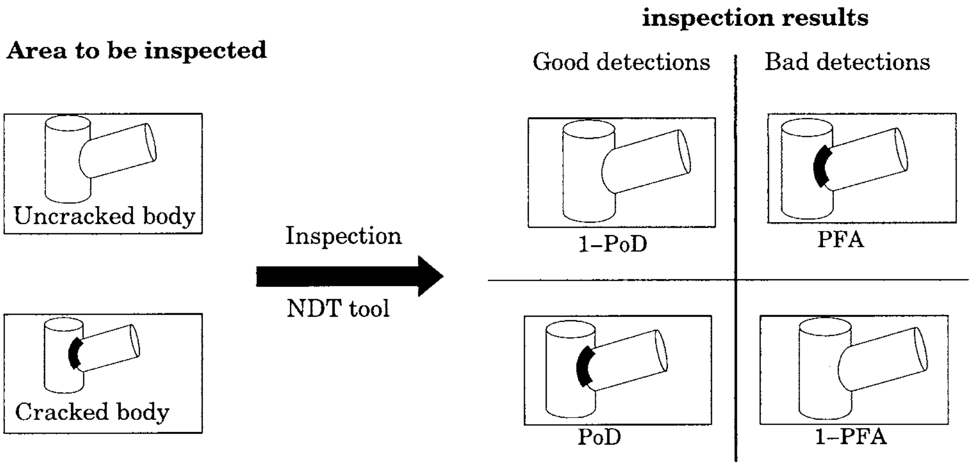

Fig. 5. Illustration of detection theory.

In these events definition, the focus is on presence or absence of crack after an inspection: the aim is finally to know whether or not there is an existing crack. To formalize this, we introduce the decision theory. More details on decision and detection theories could be found in [17,21]. From a probabilistic point of view, we consider the binary random variable 'presence of a crack' X, whose value is 1 if a crack is present, 0 otherwise. We note $d()$, the random inspection decision function, whose value is 1 if a crack is detected (i.e. we decide that one crack is present), 0 otherwise. Thus, the probability of false alarm PFA and the probability of detection PoD could be written, according to Bayes' rule:

$$\mathrm{PoD}(X) = P(d(X) = 1|X = 1) \tag{2}$$

$$\mathrm{PFA}(X) = P(d(X) = 1|X = 0) \tag{3}$$

This gives the right definitions of the PoD and the PFA:

- the PoD is the probability to decide crack presence (crack detection), conditional to an actual existing crack;
- the PFA is the probability to decide crack presence (crack detection), conditional to no actual existing crack.

These definitions are consistent with inspection calibration/inter-calibration aspects of the ICON project [5,22]. Note that an inspection result can be entirely characterized by a set (PoD, PFA).

3.2. Relationship between detection performance and crack events

Expression of the $E_{i(i=1,2,3,4)}$ events introduced in Section 3.1 can be deduced from these previous definitions of the PoD and the PFA, using Bayesian rule. Taking the E_1 event for example, we first have:

$$P(E_1) = P(X = 0|d(X) = 0) = \frac{P(d(X) = 0|X = 0)}{P(d(X) = 0)} P(X = 0) \tag{4}$$

Let's denote γ the probability of presence of crack at the inspected area, then:

$$P(X = 1) = \gamma, P(X = 0) = 1 - \gamma \tag{5}$$

Note that the probability density function of γ, as a function of the size of existing cracks and the inspected area, is related to the probability density function of the natural size of existing cracks and their spatial distribution (see Section 4.2). This underlines the fact that inspection results are conditionals to the inspected area and to the spatial distribution of cracks on the inspected component. The probability of decisions could then be expressed as:

$$P(d(X) = 1) = \mathrm{PFA}(X)(1 - \gamma) + \mathrm{PoD}(X)\gamma \tag{6}$$

$$P(d(X) = 0) = (1 - \mathrm{PFA}(X))(1 - \gamma) + (1\text{-}\mathrm{PoD}(X))\gamma \tag{7}$$

This leads finally to the following probabilities:

$$P(E_1) = P(X = 0|d(X) = 0) = \frac{(1 - \text{PFA}(X))(1 - \gamma)}{(1 - \text{PoD}(X))\gamma + (1 - \text{PFA}(X))(1 - \gamma)} \tag{8}$$

$$P(E_2) = P(X = 0|d(X) = 1) = \frac{\text{PFA}(X)(1 - \gamma)}{\text{PoD}(X)\gamma + \text{PFA}(X)(1 - \gamma)} \tag{9}$$

$$P(E_3) = P(X = 1|d(X) = 0) = \frac{(1 - \text{PoD}(X))\gamma}{(1 - \text{PoD}(X))\gamma + (1 - \text{PFA}(X))(1 - \gamma)} \tag{10}$$

$$P(E_4) = P(X = 1|d(X) = 1) = \frac{\text{PoD}(X)\gamma}{\text{PoD}(X)\gamma + \text{PFA}(X)(1 - \gamma)}. \tag{11}$$

First note that there are two ways to interpret these events:

- If crack detection is considered: the two sets (E_2,E_4) and (E_1,E_3) represent respectively crack and no crack detection.
- If crack existence is considered: the two sets (E_1,E_2) and (E_3,E_4) represent respectively crack absence and presence.

Upon which point of view is to be chosen, one or other of these sets may be used.

Second, some events are complementary. By addition of Eqs. (8) and (10), and Eqs. (9) and (11), one obtain:

$$P(E_1) + P(E_3) = 1 \tag{12}$$

$$P(E_2) + P(E_4) = 1 \tag{13}$$

This means that only one set of two events is sufficient to describe the crack detection scheme. This is due to the fact that (PoD, PFA) is a typical exhaustive set of the detection capacity for the NDT tool. Finally, we introduce the following transformation:

$$\left.\begin{matrix}\gamma \\ \text{PoD} \\ \text{PFA}\end{matrix}\right\} \rightarrow \mathcal{T}\left\{\begin{matrix}\text{1-}\gamma \\ \text{1-PFA} \\ \text{1-PoD}\end{matrix}\right. \tag{14}$$

Lets demonstrate that $\mathcal{T}(P(E_1)) = P(E_4)$ and that $\mathcal{T}(P(E_2)) = P(E_3)$:

$$\mathcal{T}(P(E_1)) = \mathcal{T}\left(\frac{(1 - \text{PFA}(X))(1 - \gamma)}{(1 - \text{PoD}(X))\gamma + (1 - \text{PFA}(X))(1 - \gamma)}\right) \tag{15}$$

$$= \frac{(1 - (1 - \text{PoD}(X)))(1 - (1 - \gamma))}{(1 - (1 - \text{PFA}(X)))(1 - \gamma) + (1 - (1 - \text{PoD}(X)))(1 - (1 - \gamma))} \tag{16}$$

$$= \frac{\text{PoD}(X)\gamma}{\text{PFA}(X)(1-\gamma) + \text{PoD}(X)\gamma} \tag{17}$$

$$= P(E_4) \tag{18}$$

$$\mathcal{T}(P(E_2)) = \mathcal{T}\left(\frac{\text{PFA}(X)(1-\gamma)}{\text{PoD}(X)\gamma + \text{PFA}(X)(1-\gamma)}\right) \tag{19}$$

$$= \frac{(1-\text{PoD}(X))(1-(1-\gamma))}{(1-\text{PoD}(X))(1-\gamma) + (1-\text{PFA}(X))(1-(1-\gamma))} \tag{20}$$

$$= \frac{(1-\text{PoD}(X))\gamma}{(1-\text{PoD}(X))\gamma + (1-\text{PFA}(X))(1-\gamma)} \tag{21}$$

$$= P(E_3) \tag{22}$$

Eqs. (15), (18) and Eqs. (19), (22) underline an interesting mathematical property of $\mathcal{T}$: it is an involution. Thus we have:

$$\mathcal{T} \circ \mathcal{T} = I_{\text{d}}, \text{ and } \mathcal{T} = \mathcal{T}^{-1} \tag{23}$$

$$\left.\begin{matrix} \gamma \\ \text{PoD} \\ \text{PFA} \end{matrix}\right\} \rightarrow \mathcal{T}\left\{\begin{matrix} \text{1-}\gamma \\ \text{1-PFA} \\ \text{1-PoD} \end{matrix}\right. \rightarrow \mathcal{T}\left\{\begin{matrix} 1-(1-\gamma) \\ 1-(1-\text{PoD}) \\ 1-(1-\text{PFA}) \end{matrix}\right. = \left\{\begin{matrix} \gamma \\ \text{PoD} \\ \text{PFA} \end{matrix}\right. \tag{24}$$

The important role of this transformation will be shown later.

This formalism allows to explain why the line PoD = PFA in the ROC plane is the "no performance" curve. When PoD = PFA, which is equivalent both to $P(d(X)=1|X=1) = P(d(X)=1|X=0)$ and $P(d(X)=0|X=1) = P(d(X)=0|X=0)$, Eqs. (6) and (7) give:

$$P(d(X=1)) = \text{PoD} = \text{PFA} \tag{25}$$

$$P(d(X=0)) = 1 - \text{PoD} = 1 - \text{PFA} \tag{26}$$

then the detection result does no more depend on the actual state of the inspected area γ. This explain why the diagonal line in the ROC chart presents the "no performance" curve.

4. Effects of NDT performances on a cost function

4.1. Building the cost function by introducing detection aspects

The optimization of an inspection programme can be achieved by minimizing a cost function [1,2]. The main objective is to reduce costs, for same performance inspections, depending on the required structural informations. This basic cost analysis consists in a cost function assessment $E(C)$, defined by the expected total cost:

$$E(C) = \sum_i C(S_i)P(C(S_i)) \tag{27}$$

$$= \sum_i C(S_i)P(S_i) \tag{28}$$

where: $C(S_i)$ is the cost associated with the ith scenario; and $P(S_i)$ is the probability that the ith scenario occurs.

In the following, $\bar{C}_i$ is defined as a cost overrun resulting from bad decisions, due to bad inspection results. Lets consider the following true scenarios provided by NDT inspection results, using the same definition as the E_i set:

- S_1: no crack presence conditional to no crack detected, the associated cost is C_1 (basically the cost of inspection); and
- S_4: crack presence conditional to crack detected, the associated cost including repair is C_4.

Lets consider the dual scenarios, giving false indications:

- S_2: no crack presence, conditional to one detected crack (event E_2), the cost is $C_2 = C_1 + \bar{C}_4$;
- S_3: crack presence, but missed (event E_3), the cost is $C_3 = C_1 + \bar{C}_1$.

For scenario S_2, and in case of systematic repair decision, the cost is identical as that of scenario 4 ($C_2 = C_4$) and there is no structural failure, due to repairs. However this is not optimal, since the cost overrun $\bar{C}_4$ is high, due to a high repair cost. If no repair is made, the false alarm detection generates a false failure scenario. When considering a reliability component approach (see [23]), such a scenario leads to the modification of the critical failure path of the failure tree. As a result of a false alarm, the cost overrun $\bar{C}_4$ increases [23], as another branch of the failure is explored. In this particular case, we have $C_1 + \bar{C}_4 > C_4$.

For scenario S_3, and depending on the size of the crack and its structural criticality, the cost overrun $\bar{C}_1$ could be very high when missing a large or critical crack, or very low when missing a small or non critical crack.

Now it is possible to evaluate the cost function (27), through inspection results. In case of crack detection (decision of crack presence), using (27) and (13), this function is given by:

$$E(C) = \sum_{2,4} C(S_i)P(S_i) \tag{29}$$

$$= C_2P(E_2) + C_4P(E_4) \tag{30}$$

$$= C_4P(E_4) + \left(C_1 + \bar{C}_4\right)(1 - P(E_4)) \tag{31}$$

In case of no crack detection (decision of crack absence), one obtains using (27) and (12):

$$E(C) = \sum_{1,3} C(S_i)P(S_i) \tag{32}$$

$$= C_1 P(E_1) + C_3 P(E_3) \tag{33}$$

$$= C_1 P(E_1) + \left(C_1 + \bar{C}_1\right)(1 - P(E_1)) \tag{34}$$

Both (31) and (34) show a summation of two terms, one associated with good detection (events E_1, E_2), and the other one with false detection (events E_2, E_3). They are in a certain way complementary and both are function of PoD, PFA and γ. It is attractive to understand how the cost function varies with these parameters.

4.2. NDT tool ranking based on cost function

In this section the effect of the cost overrun in the total expected cost is examined. In Eqs. (31), (34) the expected costs overrun are respectively $(C_1 + \bar{C}_4)(1-P(E_4))$ and $(C_1 + \bar{C}_1)(1-P(E_1))$. By plotting the probabilities of "bad" events $(1-P(E_4)) = P(E_2)$ and $(1-P(E_1)) = P(E_3)$, as function of PoD and PFA for given values of γ, we can show how PoD, PFA and γ influence on the global cost. Figs. 6 and 7 show the evolutions of $P(E_2)$ and $P(E_3)$ in the (PoD,PFA) plane, for three different values of γ:

- $\gamma = 0.1$ which represents a low probability presence of crack. It is typical of large cracks population;
- $\gamma = 0.5$ which represents a mean probability presence of crack, mostly representative of common cracks found during in-service inspections; and
- $\gamma = 0.95$ which is representative of the small cracks population.

All these probabilities can be inferred from natural crack size probability density functions: Cioclov [24], Thurlbeck et al. [25], and Moan et al. [26] assume that the size of pre-existing cracks in welded structures is distributed according to an exponential law. Following this assumption, the probability density function for natural crack sizes a, at point g with coordinates (ζ, ξ) on surface $\mathcal{A}$, is:

$$p_{\mathrm{A}}(\zeta, \xi, a) = \frac{1}{\lambda(\zeta, \xi)} \mathrm{e}^{-\frac{a}{\lambda(\zeta, \xi)}} \tag{35}$$

Considering the class $c_i = \{a > 0 | a_i < a < a_{i+1} = a_i + \delta\}$ with the class range δ, the probability to have a crack in this class at point g is:

$$\gamma(\zeta, \xi, c_i) = \int_{a_i}^{a_{i+1}} p_A(\zeta, \xi, a)\mathrm{d}a \tag{36}$$

Finally, by denoting $\mathcal{A}$ the area of interest that can be inspected by the diver, the probability to have a crack in class c_i on the inspected surface $\mathcal{A}$ is:

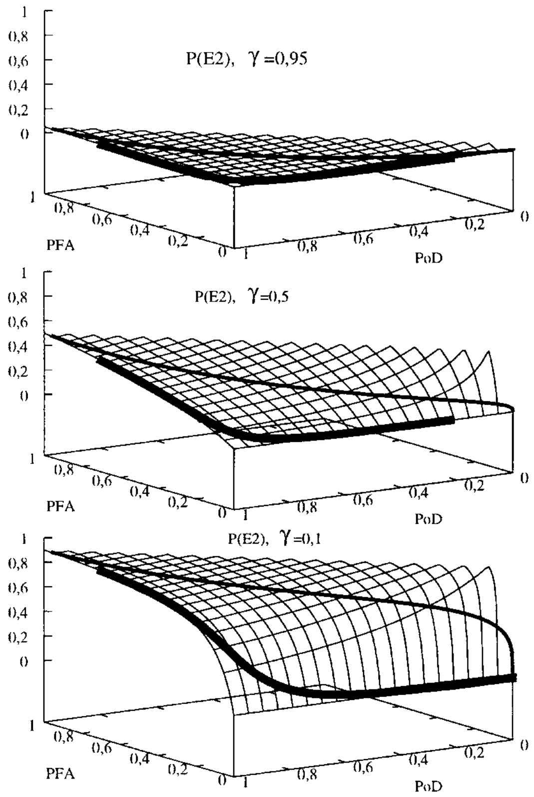

Fig. 6. $P(E_2)$ for $\gamma = 0.95$, 0.5, 0.1.

$$\gamma_i = \gamma(c_i) = \frac{1}{\mathcal{A}} \int_{\mathcal{A}} \gamma(\zeta, \xi, c_i) \mathrm{d}\zeta \mathrm{d}\xi \tag{37}$$

where $\gamma(c_i)$ is the mean of $\gamma(\zeta,\ \xi,\ c_i)$ on the inspected area $\mathcal{A}$. This definition of $\gamma(c_i)$ can be extended to the whole structure or nodes that can be inspected by computing the mean on all the nodes of the structure. By denoting n the number of nodes of the structure, we have:

$$\gamma'(c_i) = \frac{1}{n} \sum_{j=1}^{n} \frac{1}{\mathcal{A}_j} \int_{\mathcal{A}_j} \gamma(\zeta, \xi, c_i) \mathrm{d}\zeta \mathrm{d}\xi \tag{38}$$

This case can be considered when a large number of joints are inspected in one inspection campaign. Fig. 8 shows an example of the surface $\mathcal{A}$ (with geometrical parameters ζ and ξ) that can be used for a tubular T joint. Depending on the hot-spot positions, and from statistical analysis of crack positions and occurrence on the joint, one can propose a $\lambda(\zeta,\ \xi)$ function on this surface $\mathcal{A}$. Values of λ inferred from in-service observations can be found in [26]. In the basic case where λ is constant over the potential inspection area $\mathcal{A}$, $p_{\mathrm{A}}(\zeta, \xi, a)$ is plotted on Fig. 9, with $\lambda = 10$ cm and $\delta = 5$ mm. The $\gamma(c_i)$ curve is plotted against the crack size range Fig. 10. It shows clearly that γ takes high values for small

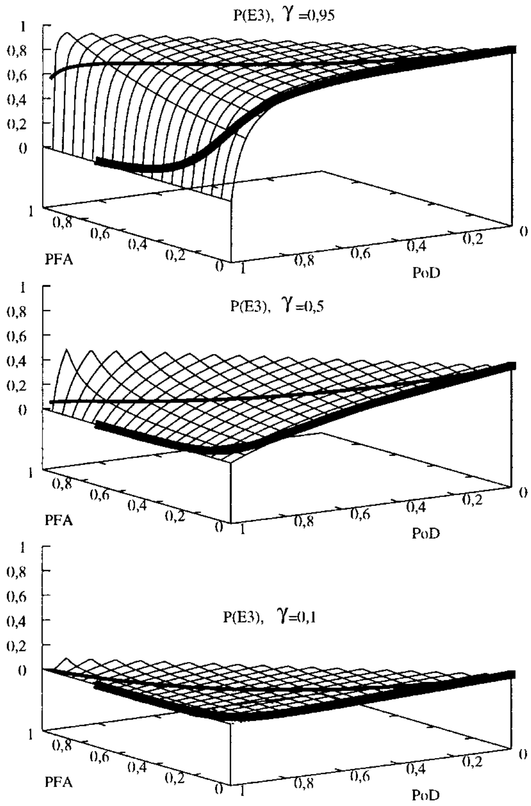

Fig. 7. $P(E_3)$ for $\gamma = 0.95, 0.5, 0.1$.

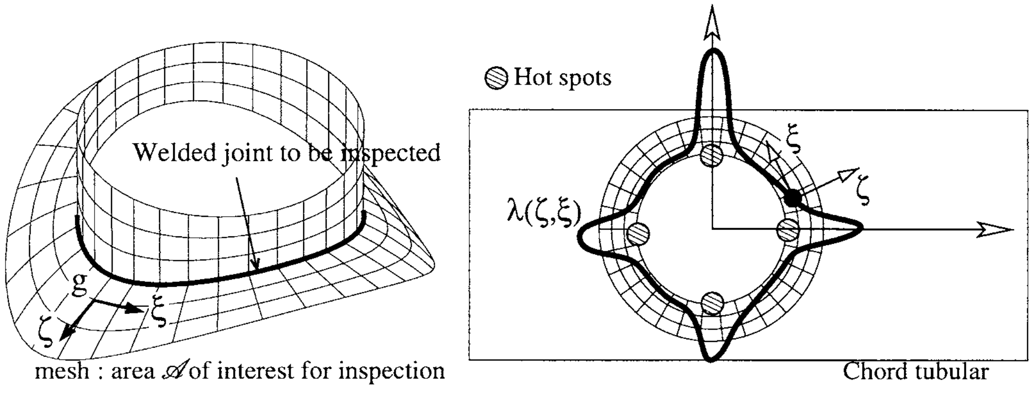

Fig. 8. Example of the area $\mathcal{A}$ of interest that can be inspected, and geometrical parameters in a tubular T joint.

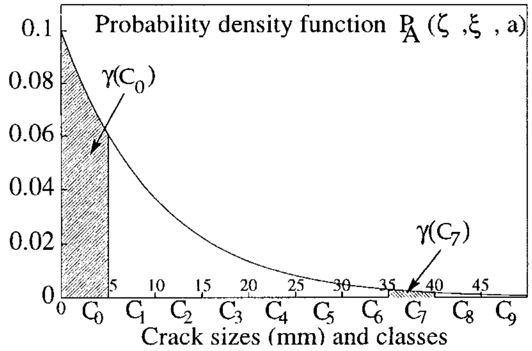

Fig. 9. Probability density function of γ.

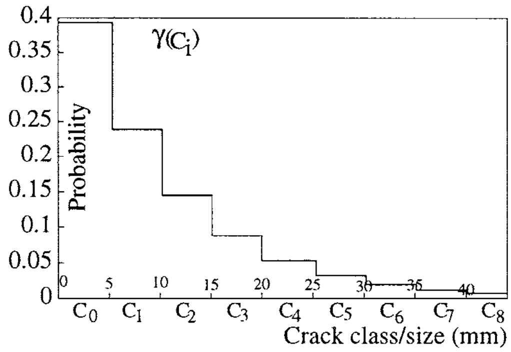

Fig. 10. γ as a function of crack size classes c_i.

crack sizes, and very small values for large cracks. Note that the case $\gamma=0.95$ is not realistic but is kept for clarity.

The interpretation of the surfaces plotted on Figs. 6 and 7 is as follows:

- for $P(E_2)$ which is associated with false alarms, we have (Fig. 6):
  - for $\gamma=0.95$: $P(E_2)\approx 0$ which means that for short cracks, the effect of PFA is not significant, even for small PoD values;
  - for $\gamma=0.5$: the effect of PFA increases, even for PoD values near 1;
  - for $\gamma=0.1$: for large and very large cracks, PFA has a dramatical effect on the overall cost as $P(E_2)$ reaches 1 for wide range of the (PoD, PFA) plan, and is nearly independent of PoD values. As result, the cost overrun has a very high probability to occur. This is why PFA has a significant effect on global IMR plans, for large cracks.

- for $P(E3)$, which is associated with no detection, we have (Fig. 7):
  - for $\gamma=0.95$: concerning small cracks, the effect of the PoD is very high and is almost independent of the PFA. Only in the case of high values of the PoD the effect of false alarms increases. The probability of event E_3 is very high, on a large part of the

(PoD,PFA) plane. Then, the cost overrun can be high in case of small defect propagation without further inspections, low otherwise;

- for $\gamma = 0.5$: for intermediate values of γ, i.e. common cracks, the effect of the PoD becomes less significant, when effect of PFA increases;
- $\gamma = 0.1$: for large cracks, the probability of cost overrun is insignificant.

Note that the effects of the PoD and the PFA are inverted, for inverted probability of crack presence γ: the probability of cost overrun $P(E_2)$ for $\gamma = 0.95$ is high, as the probability of cost overrun $P(E_3)$ for $\gamma = 0.1$, and the probability of cost overrun $P(E_2)$ for $\gamma = 0.1$ is small, as the probability of cost overrun $P(E_3)$ for $\gamma = 0.95$. This is due to the affine transformation $\mathcal{T}$ introduced in Section 3.2. One must consider this transformation as a fundamental one in inspection results, and in IMR plans. It governs the expected cost in an IMR planning, and for sensitive structures where failures leads to serious consequences, as offshore platforms and harbour installations, it is of importance to consider it.

4.3. Illustration

In the following, an illustration of costs overrun depending on the NDT capacities is presented. The repair strategy is based on the following policy:

- no crack detection leads to no action;
- crack detection leads to repair.

The different typical relative costs of failure, inspection and repair used are presented in Table 1. The knowledge of this basic IMR strategy and the associated costs, allow a discussion based on the comparative risk of several tool performances in terms of costs and based on PoD and PFA. The four cases of detection presented Section 2.4 are used. In case of non-detection and according to the defined policy, no action is undertaken: C_1 is thus the cost of inspection and a bad decision (non detection of an existing crack) leads to $\bar{C}_1$ as being the cost of failure. Thus, Eq. (34) becomes:

$$E(C) = C_1 + \bar{C}_1 P(E_3) \tag{39}$$

$$= C_1 + C_3 P(E_3) \tag{40}$$

$$= C_{\text{inspection}} + C_{\text{failure}} P(E_3) \tag{41}$$

Table 1
Cost model

Relative costs	Cost model
C_{failure}	1.0
C_{repair}	0.02
$C_{\text{inspection}}$	0.002

with

$$C_1 = C_{\text{inspection}} \tag{42}$$

$$C_3 = C_{\text{inspection}} + C_{\text{failure}} \tag{43}$$

In case of crack detection, the cost is deterministic because of the undertaken actions and includes the inspection cost, as well as the repair cost and the cost overrun in case of false alarm:

$$E(C) = C_4 P(E_4) + C_2 P(E_2) \tag{44}$$

$$= C_{\text{inspection}} + C_{\text{repair}} \tag{45}$$

with

$$C_1 = C_{\text{inspection}} \tag{46}$$

$$\bar{C}_4 = C_{\text{repair}} \quad \text{(cost overrun)} \tag{47}$$

$$C_2 = C_1 + \bar{C}_4 = C_{\text{inspection}} + C_{\text{repair}} \tag{48}$$

$$C_4 = C_{\text{inspection}} + C_{\text{repair}} \tag{49}$$

The crack size distribution is assumed to be exponential with the following parameters: $\lambda = 10$ cm and $\delta = 5$ cm. Among all the crack size classes c_i where $i \in \mathbb{N}$, the values of γ considered herein are:

$$\begin{aligned} a \in [0; 5] &\Rightarrow \gamma_0 = 0.393469 \\ a \in [5; 10] &\Rightarrow \gamma_1 = 0.238651 \\ a \in [10; 15] &\Rightarrow \gamma_2 = 0.144749 \\ a \in [40; 45] &\Rightarrow \gamma_8 = 0.007207 \end{aligned} \tag{50}$$

where the crack size a is expressed in centimeters. These crack size classes represent both small and large cracks.

In Table 2, the expected cost in case of detection, and the expected cost overrun in case of non detection are reported, as functions of the ROC point in Section 2.4 (see Fig. 3) and γ values. The best technique is the one that minimize costs, both in case of detection and in case of non detection. As the expected cost in case of detection is 0.022 (constant), the cost minimization is based both on the cost overrun $\overline{E(C)}_{\text{d}}$ in case of detection and the expected cost $E(C)_{\text{Nd}}$ in case of non detection: the aim is to select tools with overall good performances.

First, consider one class of crack size, $\gamma_2 = 0.144749$ for example. In that class, it can be seen from Table 2 that the case of detection 3 offers the best compromise in terms of costs. Taking other values of γ leads to the same conclusion. This result is in accordance with the position of point 3 on the ROC plan (see Fig. 3). It should be emphasized that having the best PoD do no let to the best NDT tool: PFA affects global performances. Hence, case 4 is not the best one, in spite

Table 2
Expected cost, depending on inspection performances and γ

Case	ROC point	$E(C)$	$a\epsilon[0;5]$ $\gamma_0 = 0.393469$	$a\epsilon[5;10]$ $\gamma_1 = 0.238651$	$a\epsilon[10;15]$ $\gamma_2 = 0.144749$	$a\epsilon[40;45]$ $\gamma_8 = 0.007207$	$\sum_{i=0}^{+\infty} \gamma_i E(C)$
Case 1	PoD = 0.25611	$E(C)_{\mathrm{Nd}}$	0.33558	0.19676	0.11751	0.00757	**0.206020**
	PFA = 0.03593	$\overline{E(C)}_{\mathrm{d}}$	0.00356	0.00618	0.00906	0.01902	**0.007427**
Case 2	PoD = 0.65896	$E(C)_{\mathrm{Nd}}$	0.24440	0.13590	0.07905	0.00557	**0.146727**
	PFA = 0.30854	$\overline{E(C)}_{\mathrm{d}}$	0.00838	0.01198	0.01469	0.01969	**0.012265**
Case 3	PoD = 0.99852	$E(C)_{\mathrm{Nd}}$	0.00300	0.00248	0.00226	0.00201	**0.002565**
	PFA = 0.03593	$\overline{E(C)}_{\mathrm{d}}$	0.00105	0.00206	0.00351	0.01664	**0.003398**
Case 4	PoD = 0.99999	$E(C)_{\mathrm{Nd}}$	0.00201	0.00201	0.00200	0.00200	**0.002008**
	PFA = 0.30854	$\overline{E(C)}_{\mathrm{d}}$	0.00645	0.00992	0.01292	0.01954	**0.010553**

of his outstanding detection performances (very high PoD, but high PFA too). If the minimization of costs is based upon the expected costs both in the case of detection and non detection, the case 4 is optimal. But it has higher total expected cost overrun, due to false alarms. This shows that the basic policy used herein is not optimal from the false alarms point of view. One should reconsider the decision: detection leads to repair.

Second, consider now a case of crack detection with given tool performances, case 2 Table 2, for example. In case of non detection, the expected costs decreases while the probability of crack presence γ decreases: the more cracks are expected in a given class range, the more missing a crack is not probable. This phenomena is more sensitive for tools which have bad probabilities of detection (cases 1 and 2). This leads to a higher expected cost due to a high probable bad action (do nothing). In case of crack detection, the cost is 0.022 whatever γ. However this is not optimal, because the policy is not too. Considering the cost overrun due to a repair in case of detection, the more cracks are expected in a given class range, the more detecting a crack is probable whatever the tool performance. This leads to a lower expected cost overrun due to high probable good action (repair).

Third, it can be seen from Table 2 that a low probability of crack presence γ and a good non destructive technique lead to a more significant cost overrun due to detection $\overline{E(C)}_{\mathrm{d}}$ than the expected cost due to non detection $E(C)_{\mathrm{Nd}}$. This phenomenon is more sensitive for higher PFA values.

Finally, the last column of Table 2 exhibits in bold the expected costs given a NDT tool for all crack size ranges, in case of detection and non detection. Case 3 appears once again to be minimizing both cost in case of non detection and cost overrun in case of detection. Note that for cases 3 and 4, the cost overrun in case of detection is much higher than the expected cost in case of non detection. This is due to false alarms, and is very sensitive in case 4, because of the high probability of false alarms.

4.4. Global ranking of NDT based on cost functions and tool performances

This cost analysis can serve as a rational basic aid tool for NDT performances ranking. On plots Figs. 6 and 7, two ROC curves are overprinted in the case of two inspection methods, one with very good overall performances [very close to the (1, 0) point, in bold], the other one less efficient. A qualified operator has performance results on the part of the ROC curve which is closest to the (1, 0) point. The estimation of the cost overrun due to bad inspections can then be calculated in terms of cost performances, using previous method. It allows a comparison between two NDT tools, in terms of risk.

5. Conclusions

Integrity assessment and structural safety of existing marine structures should be based as far as possible on rational guidance based on data collection considering costs reduction. For very large structures in harsh environment as offshore platforms, there is a need to optimize inspections planning and to model data both in terms of decision on the actual structural integrity and impact on the global cost. Considering inspection performance modeling, it has been shown that the use of detection and decision theories is very helpful and underline the importance of the set PoD, PFA and γ variables. From a rational aid tool point of view, their definition allows to introduce them as parameters of an explicit cost function. Their structure combined with the ROC curve is analysed in the three-dimensional space (PoD, PFA, γ) underlines the necessity of PFA assessment in view to rank inspection tools both in terms of intrinsic performances and global IMR cost function, as to assess in a reliable manner the structural safety of reassessed structures.

Further works should focus on PFA characterization, include the basic ranking criteria for NDT tool ranking and discuss the cost performances of several repair strategy according to this criteria.

Acknowledgements

The authors would like to acknowledge the scientific and financial support given by IFREMER, Department of Experimental Research and Developments, Materials and structure service, as well as Bureau Veritas (French) for its scientific help and for allowing our participation to the Optimization of Inspection Program, a CEPM R&D Project [27].

References

[1] Goyet J, Paygnard JC, Maroini A, Faber MH. Optimal inspection and repair planning: case studies using IMREL software. In: ASME NY, editor. Proceeding of the 13th International Conference on Offshore Mechanics and Arctic Engineering, vol. 2, 1994.

[2] Faber M, Sørensen JD. Aspects of inspection planning—quality and quantity. In: Melchers RE, Stewart MG, editor. Application of statistics and probability, 1999. p. 739–46.

[3] Dover WD, Dharmavasan S, Brennan FP, Marsh KJ, editors. Fatigue crack growth in offshore structures. EMAS; 1995.

[4] Gandhi P, Ramachandra Murthy DS, Raghava G, Madhava Rao AG. Fatigue crack growth in stiffened steel tubular joints in seawater environment. Engineering Structures 2000;22:1390–401.

[5] Barnouin B, Lemoine L, Dover WD, Rudlin J, Fabbri S, Rebourcet G, Topp D, Kare R, Sangouard D. Under-

water inspection reliability trials for offshore structures. In: ASME NY, editor. Proceeding of the 12th International Conference on Offshore Mechanics and Arctic Engineering, vol. 3B, 1993. p. 883–90.
[6] Hughes G, Bond LJ. Progress towards fatigue crack detection and sizing through partially cleaned or virgin marine deposits. In: ASME NY, editor. Proceeding of the 8th International Conference on Offshore Mechanics and Arctic Engineering, vol. 1, 1989.
[7] Newton K. The development of new techniques for underwater inspections of offshore structures. In: ASME NY, editor. Proceeding of the 9th International Conference on Offshore Mechanics and Arctic Engineering, vol. 3B, 1990.
[8] Rudlin JR, Dover WD. The ICON database. Assisting underwater inspections. In: Offshore technology, vol. 4, 1996.
[9] Rudlin JR. Reliability of inspection for fatigue cracks in offshore structures. In: Colloque lEE, 1996. p. 6/1–6/3.
[10] Berens AP. NDE reliability data analysis. In: Metals handbook, vol. 17. American Society for Metals; 1978. p. 689–701.
[11] Madsen HO, Skjong RK, Talhin AG, Kirkemo F. Probabilistic fatigue crack growth analysis of offshore structures, with reliability updating through inspection. In: Marine Structure Reliability Symposium, 1987.
[12] Guoyang Jiao. Reliability analysis of crack growth with inspection planning, in: ASME NY, editor. Proceeding of the 11th International Conference on Offshore Mechanics and Arctic Engineering, vol. 3, 1992.
[13] Moan T, Johannesen JM, Vàrdal OT. Probabilitic inspection planning of jacket structures. In: Offshore Technology Conference Proceedings, no. OTC paper 10848, Houston, TX, 1999.
[14] Moan T, Song R. Implication of inspection updating on system fatigue reliability of offshore structures. In: ASME editor. Proceeding of the 17th International Conference on Offshore Mechanics and Arctic Engineering, no. 1214, 1998.
[15] Rouhan A, Schoefs F, Labeyrie J. Approche probabiliste du contrôle non destructif des structures pétrolières pour l'évaluation de leur integrite structurelle. Revue Francaise de Mecanique 2000;4:235–41.
[16] Rouhan A, Schoefs F. On the use of inspections results, in IMR plans. In: Mangin MM, editor. DMinUCE—2000. Second International Conference on decision making in urban and civil engineering, vol. 2, 2000, p. 1259–70.
[17] Arques PY. Decisions en traitement du signal. Paris: Editions Masson; 1982.
[18] Rouhan A. Étude sur les probabilités de detection des contrôles non destructifs., Tech. rep., IFREMER, university confidential research report no. 002021316 TMSI/RED/MS, May 2001.
[19] Fücsök F, Muller C, Scharmach M. Human factors: the NDE reliability of routine radiographic film evaluation, In: Proceedings of the 15th World Conference on Non Destructive Testing, vol. 3, 2000.
[20] Nockemann C, Tillack GR, Wessel H, Hobbs C, Konchina V. Performance demonstration in NDT by statistical methods: ROC and POD for ultrasonic and radiographic testing. In: Proceedings of the 6th European Conference on Non Destructive Testing, vol. 44, 1994.
[21] Charabit M. Elements de théorie du signal: les signaux aléatoires. Paris: Editions Ellipses; 1990.
[22] Goyet J. Evaluation des apports de la technique d'inspection FMD, ATA 580 A-OPI. Rapport interne Bureau Veritas, 1996.
[23] Schoefs F, Dubé JF, Goyet J. Semi probabilistic analysis and instability detection on failure trees as rational aid tools for existing structures reassement. In: Mangin, MM, editor. DMinUCE—2000. Second International Conference on decision making in urban and civil engineering, vol. 2, 2000. p. 857–68.
[24] Cioclov DD. A model for fracture risk analysis, Revue roumaine de sciences techniques. Série de mécanique appliquee 1986;6:613–22 [fascicule 31].
[25] Thurlbeck SD, Stacey A, Sharp JV, Nichols NW. Welding fabrication defects in two offshore steel jacket structures. In: ASME editor. Proceeding of the 15th International Conference on offshore mechanics and arctic engineering, vol. 3, 1996, p. 232–46 [Materials Engineering].
[26] Moan T, Vàrdal OT, Hellevig N-C, Skjoldli K. In-service observations of cracks in North-Sea jackets. A study on initial crack depth and POD values. In: ASME NY, editor. Proceedings of the 16th International Conference on Offshore Mechanics and Arctic Engineering, vol. 2, 1997.
[27] Goyet J. Optimisation des programmes d'inspection. Tech. rep., Bureau Veritas, Direction Marine, DR (France), CEPM: M2801/97, final rapport, 2001.