International Journal of Fatigue 86 (2016) 77–87

Contents lists available at ScienceDirect

International Journal of Fatigue

journal homepage: www.elsevier.com/locate/ijfatigue

Advanced ultrasonic “Probability of Detection” curves for designing in-service inspection intervals

Michele Carboni [a],*, Stefano Cantini [b]

[a] Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156 Milano, Italy
[b] Lucchini RS SpA, Via G. Paglia 45, 24065 Lovere, BG, Italy

ARTICLE INFO

Article history:
Received 9 March 2015
Received in revised form 11 July 2015
Accepted 15 July 2015
Available online 21 July 2015

Keywords:
Ultrasonic testing
Probability of detection
Inspection intervals
Railway axles
Boreprobe equipment

ABSTRACT

In order to ensure the safe service of modern railway vehicles, safety critical components are subjected to a dedicated maintenance and inspection plan. Railways axles, in particular, are periodically inspected by the ultrasonic testing method during maintenance interruptions in first level workshops, while the magnetic particles test is carried out in second level workshops. The reliability of such inspections is quantified in terms of “Probability of Detection” curves, which are traditionally related to a specific linear dimension of the defect to be detected. Actually, a fatigue crack is known to change its shape under cyclic loading, so affecting its detectability.

In the present paper, a novel approach to “Probability of Detection” takes into account of this effect, introducing a “Master” probability of detection curve, which is a function of the reflecting area of the crack. A comparative applicative example between standard and such advanced probabilities of detection curves is then presented considering high-speed railway applications, whose hollow axles are made of the medium strength EA4T steel grade.

1. Introduction

Considering mechanical components subjected to fatigue, and assuming during service some surface damages might occur and not be promptly repaired, it is licit to expect crack initiation and consequent propagation. To face this problem, some critical safety fields (such as railways, aerospace, automotive, ...) employ the “Damage Tolerant” design approach as state-of-the-art. Its philosophy consists [1,2] in determining the most opportune in-service inspection interval given the “Probability of Detection” (POD) curve [3–5] of the adopted “Non-Destructive Testing” (NDT) method or, alternatively, in defining the needed NDT specifications given a programmed inspection interval. Structural integrity of safety components during service is then strictly related to the following factors [1,2,6]: (i) the capability of the adopted NDT procedure, i.e. its POD curve; (ii) the crack propagation behavior of the adopted material; (iii) the influence of the geometry of the cracked body on crack driving force; (iv) the reliable knowledge of service loads. An effective damage tolerant approach requires, then, well-defined crack growth lifetime predictions and reliable POD curves.

Focusing here on the capability of NDT, it may be qualitatively defined as “the probability of detecting a crack in a given size group under the specified inspection conditions and procedures” [7]. Even if many similar definitions can be found in the literature, it is well known that it is a statistical matter [7] and that the quantitative and the universally accepted underlying tool is the aforementioned POD curve of the adopted NDT method (a scheme is shown in Fig. 1a). This characteristic statistical aspect of NDT derives from the experimental evidence that repeated inspections of the same flaw size or type do not necessarily result in consistent indications: this is the reason for the “realistic” curve shape shown in Fig. 1a against the “theoretical” expected one. Such a realistic curve is then usually derived by experimental tests on components containing numerous artificial or natural defects. Moreover, traditionally ([3–5]), such probabilities are explicitly expressed and plotted in terms of a characteristic linear dimension of defects (depth, length, diameter, ...). This because the depth is known to be the most relevant parameter for fracture mechanical assessment of superficially cracked bodies [2], but, actually, POD curves are also a function of many other physical and operative factors like the adopted NDT method, material, geometry, defect type and shape, equipment, human and environmental factors. This means that very rarely the POD curve defined for a given inspection procedure can be used for other ones, even if similar. For all these reasons, their definition process is very demanding in terms

---

* Corresponding author. Tel.: +39 02 23998253; fax: +39 02 23998202.
E-mail address: michele.carboni@polimi.it (M. Carboni).

http://dx.doi.org/10.1016/j.ijfatigue.2015.07.018

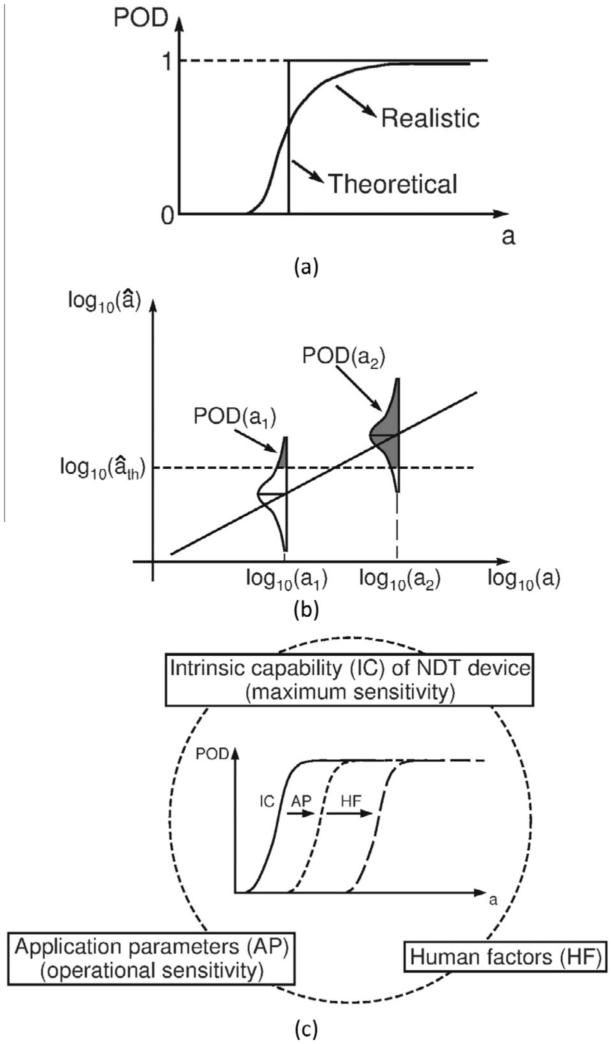

Fig. 1. Derivation of "signal response" POD curves: (a) scheme of a POD curve; (b) decision threshold for "signal response" data; (c) main influences of NDT capability and their effect on POD curves.

of time and costs and, over the last years, experimental responses have begun to be partially substituted and integrated by numerical simulations [8–12].

Two statistical methods are available for analyzing NDT capability data and produce POD curves as functions of the linear flaw size "a". The first one [3–5], also chronologically, is based on "hit/miss" data, where NDT results are only recorded in terms of whether the crack is detected or not. The second one ([3–5]) is based on the presence of more information within the NDT response, typically in terms of peak voltage in eddy currents NDT, the signal amplitude in ultrasonic NDT or the light intensity in fluorescent penetrant NDT. Since, in this case, the NDT signal response is someway correlated to flaw size, this method is named "signal response" or "â vs. a", where "a" is the characteristic linear dimension of the defect and "â" its response to the inspection stimulus. The POD curve of the NDT procedure is, in this case, strictly related to the adopted calibration via the decision threshold $\hat{a}_{th}$ (Fig. 1b). Both methods rely on different probabilistic models to produce POD curves, for more details see [4,5]. Only the signal response approach is considered and described in the present research.

A very important aspect of POD curves is the need, for reliability and design of components, of a statistical characterization of the largest defect that can be missed and not of the smallest one that can be detected. For this reason, POD curves should always be provided along with a suitable lower confidence limit (typically 95%). Moreover, when dealing with real-life POD curves, it is important to distinguish [13] the intrinsic performance of the equipment from its application to different inspecting procedures and from all human factors affecting calibration and inspection operations (Fig. 1c).

Today, these concepts are also applied to railway axles, safety components designed to have an infinite lifetime [14], but showing occasional failures during service [15]. Such failures always occur, at the most stressed regions, as fatigue crack propagations whose initiation can be due to different causes [1]: for example (Fig. 2), wrong handling or maintenance practice, the presence of widespread corrosion [16,17] or the possible damage due to ballast impacts. For this reason, to guarantee adequate reliability and safety during service, NDT is performed during both production, in order to detect internal and surface manufacturing defects, and maintenance, in order to detect in-service surface damages. Specifically, during service, railway axles are periodically inspected by means of ultrasonic testing (UT) at ordinary maintenance service interruptions and by UT and magnetic particles at overhauls. In particular, solid axles are manually inspected by traditional UT probes [18] or, in Italy, by means of a rotating UT probe [19] applied to both ends of the axle and scanning the critical regions (press-fit seats and geometrical transitions). In this case, POD curves are available in the literature [8,20], but more work is still needed to define a generally accepted inspection procedure. Hollow axles are, instead, inspected using a highly automated UT boreprobe roto-translating along the longitudinal bore and scanning the whole external surface, while the availability of POD curves is more meager, an example is given in [21].

The present paper analyses a potential inconsistency regarding the traditional definition and representation of UT POD curves. Even if generally applicable to any UT inspection procedure, such an analysis is carried out considering the specific applicative case of hollow railway axles inspected by the boreprobe. The study is based on an approach recently proposed by the authors [9,22], which is first briefly summarized. Then, a dedicated experimental full-scale fatigue campaign on axles is described: the availability of a large number of natural cracks, propagated under controlled fatigue loads, is the peculiarity of the here-presented research, especially considering this statistical population of natural defects in full-scale axles is currently unique. The derived set of natural defects and the novel approach to POD curves allowed defining those of the boreprobe, in a more flexible way, for its application to the inspection procedure adopted by Lucchini RS during production and maintenance [23]. Finally, some considerations are provided about the effect of the adoption of either traditional or advanced POD curves on the probability of failure of the inspection procedure of hollow axles by the boreprobe.

2. The "Reflecting Area" approach to ultrasonic testing data

2.1. The traditional derivation of probability of detection curves

Considering the signal response approach, a POD(a) function is derived from the correlation of â vs. a data: between the four possible relations required to fulfil the conditions of the POD model [3], an approximate linear relationship will be demonstrated to exist, for the present applicative case, between $\log_{10}(\hat{a})$ and $\log_{10}(a)$:

$$\log_{10}(\hat{a}) = \alpha + \beta \cdot \log_{10}(a) + \gamma \quad (1)$$

where γ is an error term distributed with zero mean and constant standard deviation σ_γ. Actually, Eq. (1) expresses the fact that