A NEW APPROACH FOR THE DEFINITION OF "PROBABILITY OF DETECTION" CURVES

Michele CARBONI[1], Stefano CANTINI[2]
[1] Dept. Mechanical Engineering, Politecnico di Milano, Milano, Italy
[2] Lucchini RS, Lovere (BG), Italy

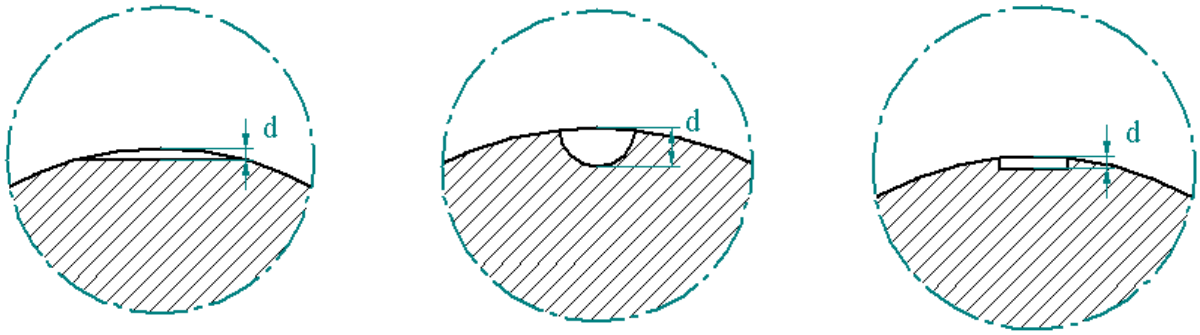
1. INTRODUCTION

Considering components subjected to fatigue loads, in presence of Stress Intensity Factors higher than design assumptions, it is licit to expect crack nucleation and consequent propagation during service. Crack nucleation sites, in the most critical design sections, are expected where manufacturing, or service (also related to aggressive environments) defects are present.
From this point of view, it is possible to employ the "Damage Tolerant" design approach whose philosophy consists ([1],[2]) in determining the most opportune inspection interval given the "Probability of Detection" (POD) curve ([3],[4]) of the adopted Non-Destructive Testing (NDT) technique 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 different factors such as: i) the performance of non destructive testing ("NDT"); ii) the crack propagation behaviour of the adopted material; iii) the reliable knowledge of service loads.
Up till now, the performance of a given NDT technique in detecting cracks and defects has been quantified and summarised by the corresponding POD curve relating the probability to detect the defect to its characteristic linear dimension (diameter, length, depth, etc.). Recent results published by the authors ([5]-[7]) showed that such "linear dimension" approach is not the most effective in order to categorise the response of artificial and natural defects having different dimensions and shapes. In particular, artificial defects obtained by traditional machining (saw-cuts) and EDM (concave and convex shaped) on sample blocks obtained from railway axles and fatigue cracks propagated into full-scale axles were considered and compared. Better results in correlating ultrasonic echoes, for all the shapes, dimensions and artificial or natural defects, could be achieved adopting, as a characteristic parameter, the reflecting area.
The present paper, based onto such results added with more experimental evidences, proposes a novel and more efficient approach to the definition of POD curves in terms of reflecting area of defects instead of their linear dimension. The application and the usefulness of the novel approach will be here exemplified considering the relevant case of hollow railway axles.

2. UT MEASUREMENTS ON ARTIFICIAL DEFECTS

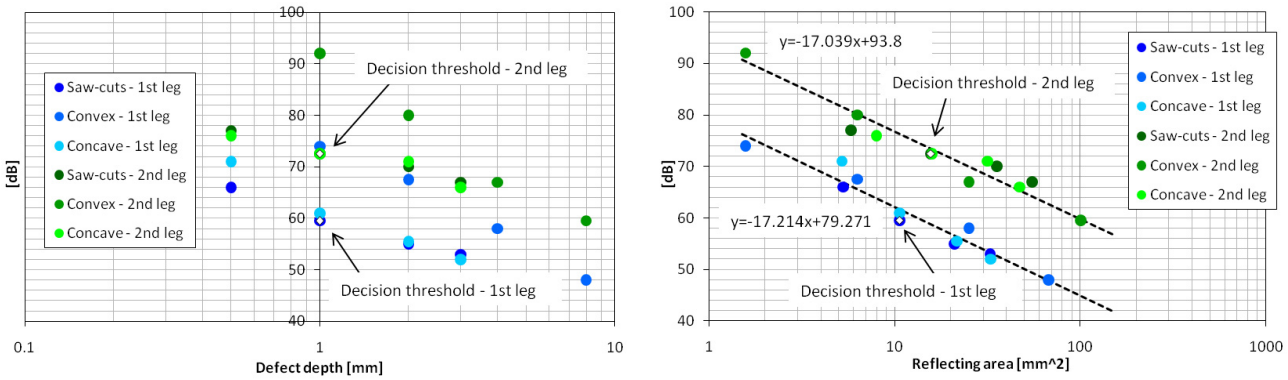
Some artificial defects were realised on the external surface of two different sections of a hollow axle body (D_{ext}=152 mm) made in A4T steel grade (typically used in the production of European railway axles [8]) by means of different technological processes. Twelve defects were introduced: four saw-cuts (Fig. 1a) having depth d=0.5, 1, 2 and 3 mm, four convex electrical discharge machined (EDM) notches (Fig. 1b, representing fatigue cracks typically observed at the body or at the transitions) having radius equal to d=1, 2, 4 and 8 mm and four concave EDM notches (Fig. 1c, representing fretting-fatigue cracks typically observed at the press-fits or surface damages due to impacts, scratches, etc.) having depth equal to d=0.5, 1, 2 and 3 mm.
UT measurements of all these artificial defects were carried out by means of a portable digital RDG500 detector [9] equipped with a 8x9 mm single crystal ATM45-4 probe (45° and 4 MHz) and recording the dB level giving the 80% amplitude on the screen. In all cases, coupling was guaranteed by grease and proper plexiglas wedges. In particular, artificial defects could be inspected adopting both 1st leg (probe coupled in the bore) and 2nd leg (probe coupled on the external surface and reflection of the sound beam on the bore) configurations since the sample blocks were cut longitudinally making the bore available.

Results are summarised in Figure 2a. As it can be seen, there is a strong dependence of UT echoes on the shape of defects. Particularly, saw-cuts and concave EDM notches (geometrically similar for the considered depths) give similar response for both 1st leg and 2nd leg configurations, while the convex EDM notches (at equal depth) give a significantly different response.
As already pointed out in [5] and [6], this behavior suggests that the depth of the defect is not the best parameter in order to entirely characterize its UT response, even if this assumption traditionally comes from the important consideration that the depth is the most dangerous parameter for the reliability and safety of railway axles. It can be more suitable, instead, to consider the area of the defect actually invested by the sound beam (Fig. 2b). In these terms, the effect of the shape of the defect seems to be negligible, since the different curves obtained from different defect shapes (but considering the same time of flight, i.e. the same inspection configuration) collapse onto a single one. It is also worth noting that in a semi-logarithmic plot, such curve is a line. Moreover, even if experimental data are characterized by quite a large scatter band, a difference in the time of flight seems to simply correspond to a rigid translation of the linear behavior, i.e. a parallelism is observed (see the slopes of the fitting lines in Fig. 2b).



(a) (b) (c)

Figure 1 – Artificial defects realised on axles: a) saw-cuts; b) convex EDM notches; c) concave EDM notches.



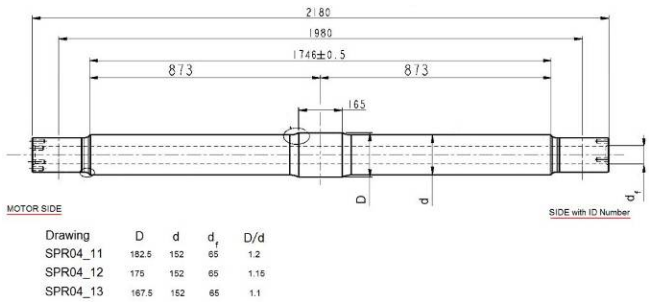
(a) (b)

Figure 2 – Comparison between UT echoes from artificial defects in terms of depth (a) and reflecting area (b).

3. UT MEASUREMENTS ON NATURAL FATIGUE CRACKS

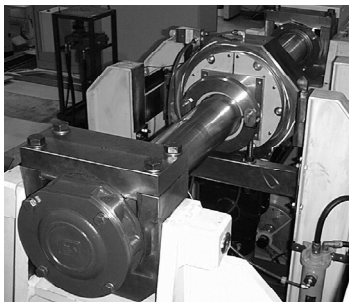
UT echoes from real fatigue cracks were recorded during three different crack propagation tests carried out (in the frame of the WIDEM European Project [10]) on full-scale hollow axles (Fig. 3a, $D_{ext\text{-}body}$=152 mm) made in A4T steel grade. In particular, such tests were carried out by means of the dedicated test bench (Fig. 3b, [11]) available at Politecnico di Milano (Dept. of Mechanical Engineering). Since cracks propagated in the body of the axles starting from some artificial micro-holes and due to rotating bending, their shape is consistent with that of the artificial defects shown in Fig. 1b (i.e. semi-circular or semi-elliptical, Fig. 3c). Extensive UT inspections were then carried

out, using the same equipment already described in the case of artificial defects, in order to characterize and record the response echoes coming from such fatigue cracks growing due to the applied cyclic load. Only inspections characterized by the 2nd leg configuration could be adopted in this experimental campaign.



Drawing	D	d	d_f	D/d
SPR04_11	182.5	152	65	1.2
SPR04_12	175	152	65	1.15
SPR04_13	167.5	152	65	1.1

(a)

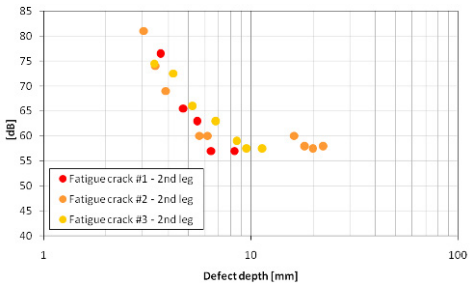


(b)

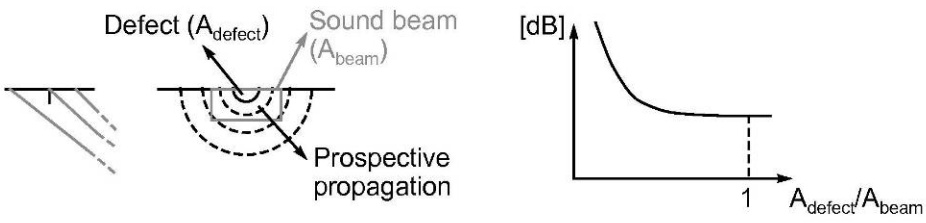
(c)

Figure 3 – UT measurements of natural cracks in hollow axles: a) drawing of the axles; b) dedicated test bench; c) typical fatigue crack.

Such measurements on growing fatigue cracks showed a dB saturation (Fig. 4a). This behaviour has already been evidenced in [12] and is explained in Fig. 4b in terms of the ratio between the area of the defect and the one of the sound beam at the given time of flight. Saturated data will not be considered here because not useful in deriving POD curves.



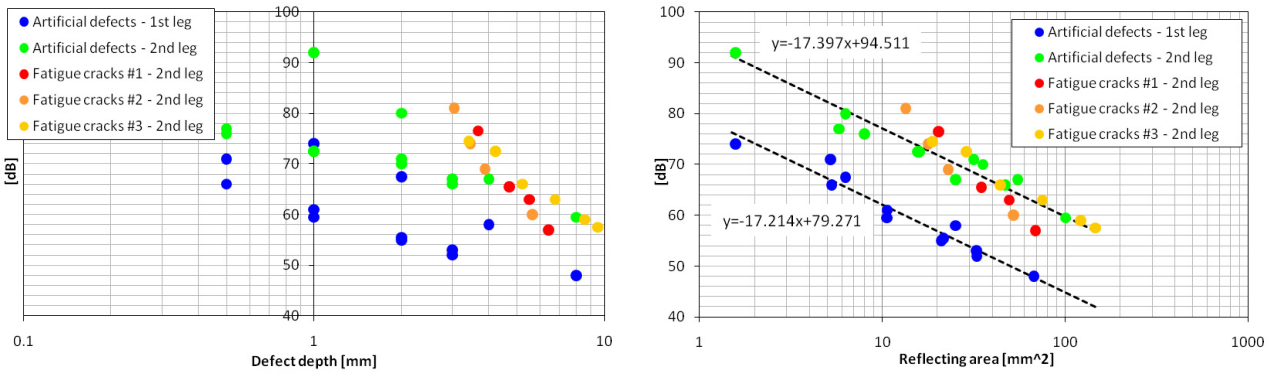
(a)



(b)

Figure 4 – dB saturation of UT measurements carried out on growing cracks.

Figure 5a shows, then, the comparison between the not-saturated fatigue cracks data and the artificial defects in terms of depth. Real fatigue cracks give, as expected, a response similar to convex EDM notches, but require higher energy at lower depths (as often reported in the literature, e.g. [13]). Finally, it is very important to note (Fig. 5b) that also data coming from real fatigue cracks mix up very well with those from artificial defects if plotted in terms of reflecting area.



(a) (b)

Figure 5 – Comparison between UT echoes from artificial defects and real fatigue cracks in terms of depth (a) and reflecting area (b).

4. UT MEASUREMENTS ON NATURAL FRETTING-FATIGUE CRACKS

Eight fretting-fatigue natural cracks (an example is shown in Fig. 6) were nucleated at the press-fits (D_{ext}=175 mm) of some axles, characterized by a low D/d ratio, by means of a test bench (identical to the one described in the previous section) available at Lucchini RS ISO 17025 certified labs. The typical shape of fretting fatigue cracks is consistent with that of the artificial defects shown in Fig. 1c (i.e. shallow). In particular, the surface length of each crack was determined using MPI (maximum 184 mm and minimum 14 mm), while the depth by means of an ACPD instrument (maximum 35.1 mm and minimum 0.6 mm). Such cracks were inspected adopting the already described UT equipment considering only the 2nd leg configuration. Figure 7 shows the comparison between the previous results and the new ones obtained from fretting-fatigue cracks.

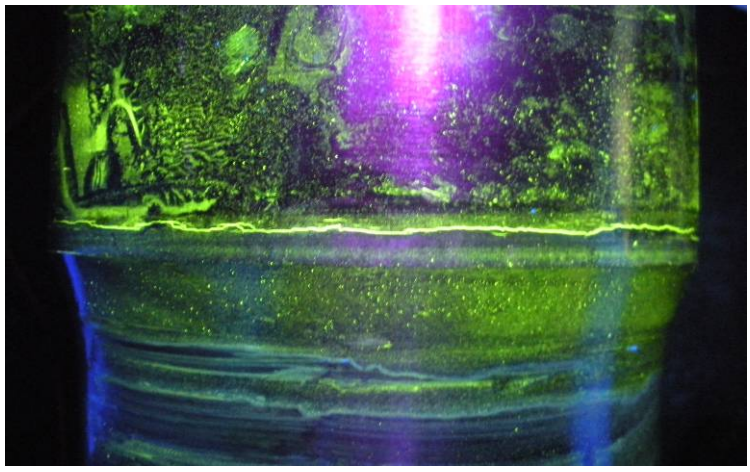
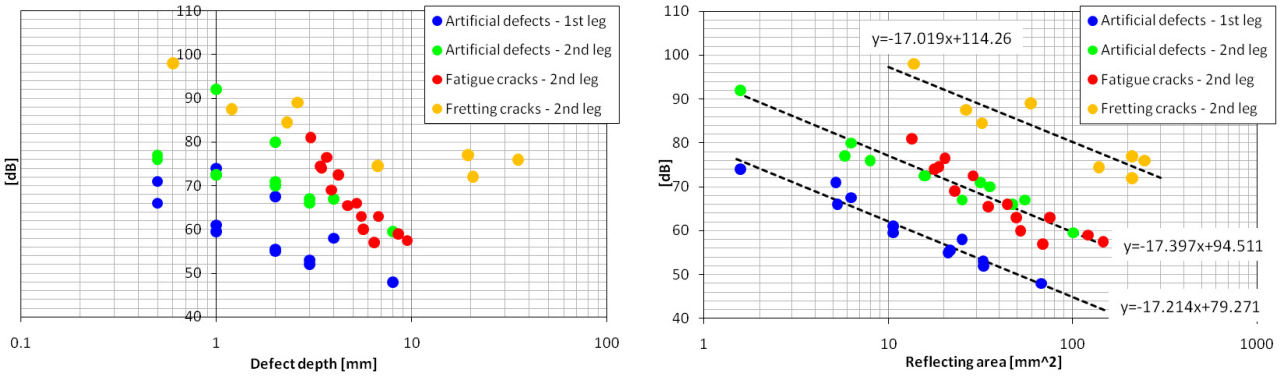


Figure 6 – Fretting-fatigue natural cracks evidenced by means of fluorescent MPI.

As it can be seen, the trend of fretting fatigue defects remains parallel to that of fatigue cracks and artificial defects. Moreover, the difference between fatigue defects in 2nd leg configuration and fretting fatigue defects in the same configuration is about 15-20 dB. Such difference can be ascribed to two different factors: i) the different time of flight between defects located at the body and those located at the press-fit; ii) the different crack morphology, since fretting-fatigue cracks are characterized by higher surface roughness and peculiar complex local paths (zigzagging, branching and kinking, Fig. 6). The amount of dB due to each of these two factors was determined using a

sample block, available at Lucchini RS, presenting the same artificial defects (concave, convex and saw-cuts) both at the body and at the press-fit. It resulted that the difference due to path length is about 6-8 dB, while the rest can be due to the different crack morphology.



(a) (b)

Figure 7 – Comparison between UT echoes from artificial defects and natural cracks in terms of depth (a) and reflecting area (b).

5. DERIVATION OF POD CURVES

Generally, in order to derive POD curves ([3],[4]), it is firstly necessary to define a decision threshold in dB corresponding to the defect to be detected with 50% probability. The POD curve can then be built, for each defect dimension, calculating the probability of a given defect responding with an energy lower than the chosen threshold.

Present UT experimental results in terms of depth do not lie on a single line (Fig. 7a), so they were split in four different groups: semi-circular (convex) defects, shallow (saw-cuts and concave) defects, natural fatigue cracks and natural fretting-fatigue cracks. In these terms, six different POD curves were defined (Fig. 8) considering the application of the already described 8x9 mm probe characterized by 45° shear wave and 4 MHz. Decision thresholds were chosen corresponding to the response of the saw-cut having depth equal to 1 mm, one of the most traditional calibration defect for sample blocks of railway axles. Such decision thresholds resulted to be equal to 59.5 dB and 72.5 dB for the 1st leg and the 2nd leg configurations, respectively (see Fig. 2a). Considering fretting-fatigue cracks, a 90 dB decision threshold was instead estimated, due to the lack of a 1 mm saw-cut defect at the same time of flight of fretting-fatigue cracks, adding 17.5 dB (as explained in the previous section) to the response of the 1 mm saw-cut inspected in 2nd leg configuration.

It is worth noting that defects with different shape, but the same depth, can have completely different POD curves (compare, for example, red and blue curves). This fact suggests that, choosing the depth approach, calibration should be theoretically carried out on shallow defects for detecting fretting-fatigue cracks and surface damages, while on semi-circular defects for detecting fatigue cracks. Moreover, an interesting observation is that the two POD curves for shallow defects (red curves) are very similar for the considered times of flight and the same can be said for the two for semi-circular defects (blue curves): it seems that a suitable calibration of UT systems can then yield similar POD curves for different path lengths, but it remains that shorter times of flight have the advantage to need lower sound energy (in Fig. 2a, data from 1st leg configuration are systematically lower, in terms of dB, than those from 2nd leg configuration). The POD curves for natural cracks (green and orange curves) are comparable to the ones for artificial defects, but their rounded shape is supported by a higher scatter in experimental outcomes (Fig. 7b).

Data reported in Figure 7b can be used to define POD curves in a new perspective based on the reflecting area instead of the traditional defect depth. Particularly, as seen in the previous sections, all data corresponding to a given time of flight lie on a single line: choosing suitable decision thresholds, it is then possible to define a single POD curve for the given time of flight (Fig. 9, detail on the left). The decision thresholds were here chosen as the dB generated by the reflecting areas

corresponding to the saw-cuts having depth equal to 1 mm (see Fig. 2b) or estimated as in the previous section: it is important to understand that these reflecting areas are different because of the divergence of the sound beam with path length. Theoretically, also POD curves in terms of reflecting area should be very similar for different times of flight: in the present case, the difference of the three curves can be ascribed to the significantly different scatter in experimental data (Fig. 7b).

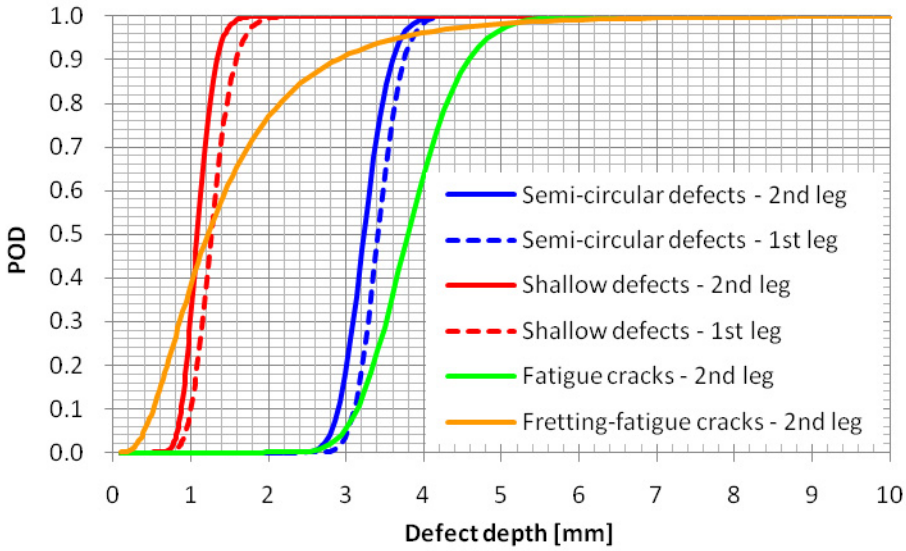


Figure 8 – POD curves in terms of defect depth for a 8x9 mm probe characterized by 45° shear wave and 4 MHz.

The obtained POD curves in terms of reflecting area do not, actually, improve the probability to detect defects in respect to the curves in terms of depth, but they have characteristics of generality and versatility that are not available from the traditional ones. This because POD curves in terms of reflecting area may be transformed into those in terms of depth simply assuming proper shapes and values for the area. Figure 9 shows the procedure for the examples of a shallow defect inspected in 1st leg configuration and a semi-circular defect inspected in 2nd leg configuration. Starting from the POD values in terms of reflecting area, a shape can be associated to the area itself and, consequently, depth values can be derived and used to back-calculate POD curves in terms of depth. As it can be seen from the detail on the right of the figure, the back-calculated curves, obtained by the systematic application of the procedure to a wide range of reflecting area values, correspond almost exactly with the experimental ones. It is important to remark that this transformation regards the abscissa of the plots, while the POD values remain the same. With this procedure, it is also possible to transform POD curves in terms of reflecting area into POD curves in terms of depth of defects for which no experimental data are available: it is then possible to consider the POD curve in terms of reflecting area as a "Master" (or "Parent") POD curve.

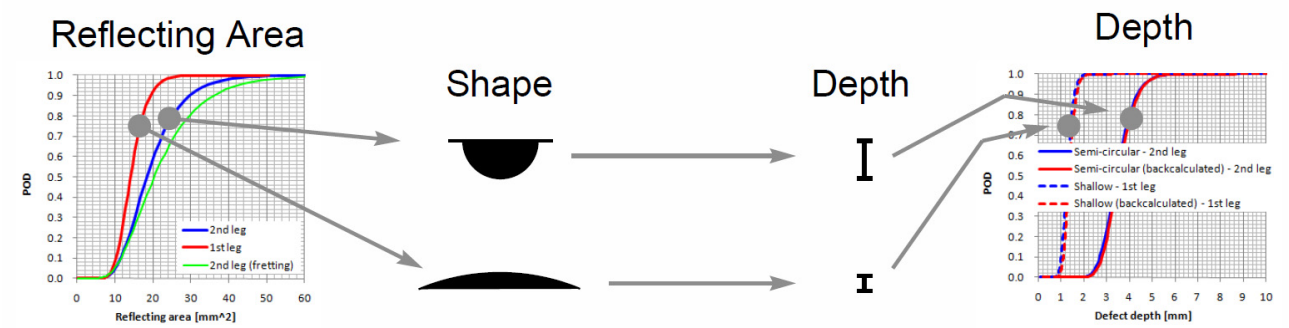


Figure 9 – Versatility of POD curves in terms of reflecting area.

6. CONSIDERATIONS ABOUT THE APPLICATION TO RAILWAY AXLES

The usefulness of the described novel approach in the application to railway axles can be appreciated considering that fatigue cracks during service typically nucleates from surface damages such as impacts, scratches or corrosion pits. All these starters initially present a shallow shape, but growing they tend to become semi-circular. Figure 10 shows the growth prediction, carried out by the dedicated software package AFGrow 4.0011.14 [14], of a shallow surface crack (a/c=0.5) positioned in a hollow axle (D_{ext}=152 mm, D_{int}=65 mm) made of A4T steel and subjected to bending. As it can be seen, the a/c ratio starts to increase immediately, so supporting the changing of shape of the defect during propagation. From this point of view, it can be a problem, for the determination of inspection intervals, to decide the right one between the two POD curves for shallow and semi-circular defects shown in Figure 8, while the POD curves in terms of reflecting area can provide continuous results during the whole transition between shapes.

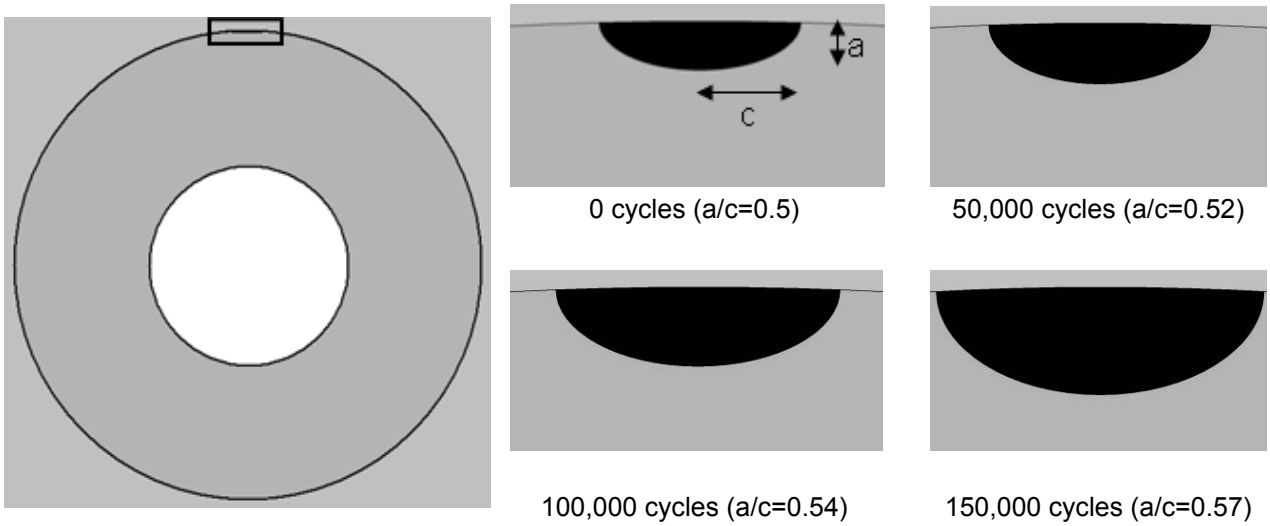
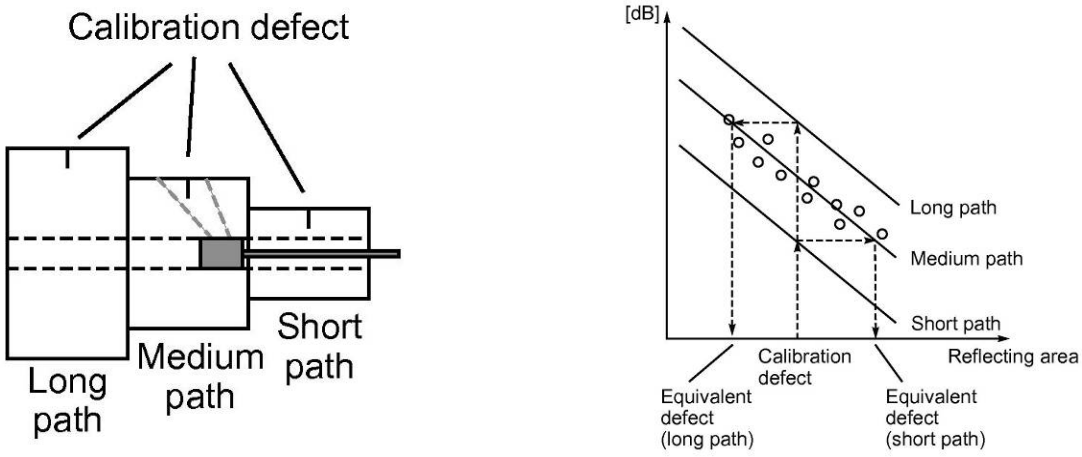


Figure 10 – Shape development of a shallow surface defect in a railway axle.

Another very important aspect consists in the determination of the right decision thresholds, without which POD curves cannot be derived, for the different times of flight of interest. Actually, this aspect is already solved using proper sample blocks where the same defect (typically corresponding to the one whose response has to be adopted as the decision threshold) is positioned at different times of flight (Fig. 11a): the dB levels experimentally obtained for each time of flight correspond, then, to the decision thresholds. A more sophisticated and powerful approach, again based onto the just described sample block, is implemented in some modern UT systems (i.e. [15]) for inspecting railway axles during both production and in-service stages. In particular, the dedicated software of the equipment builds, during the calibration procedure and knowing exactly the geometry of the sample block, DAC curves already adjusted to keep into account for the effect due to the time of flight. This procedure is equivalent (Fig. 11b) to consider only one POD curve (i.e. to keep the same time of flight) and suitably derive the area of an equivalent defect representing the effect of the time of flight (shorter paths give equivalent defects bigger than the calibration one, while, contrarily, longer paths give equivalent defects smaller than the calibration one). Moreover, such UT systems can keep into account for the drawing of the inspected axle, so automatically exclude the response of geometric features that are not useful in the determination of the POD curve.

The algorithm for the definition of maintenance inspection intervals of railway axles is widely known in the literature and has been presented by the authors in different recent papers such as [12], [16] and [17]. What is reported in the present study has a direct influence on the first part (NDT) of the procedure, where it is necessary to determine the dimension of the defect having a given probability to be detected and from which propagation will start. Figure 12 shows the comparison between the traditional (based on POD curves in terms of depth and valid for one time of flight) and

the proposed (based on the Master POD curve compensated for the effect of the time of flight) calculation procedures for inspection intervals. The versatility of the Master POD curve is evident together with its ability to estimate curves for defects for which no experimental data are available. Furthermore, the experimental effort for deriving the Master POD curve is equivalent to that needed for determining just one POD curve in terms of depth. As already said, no advantages are instead achieved in terms of a better probability of detecting defects which is peculiar of the given calibration procedure and inspecting device.



(a) (b)

Figure 11 – Effect of different times of flight on the choice of decision thresholds for the derivation of POD curves.

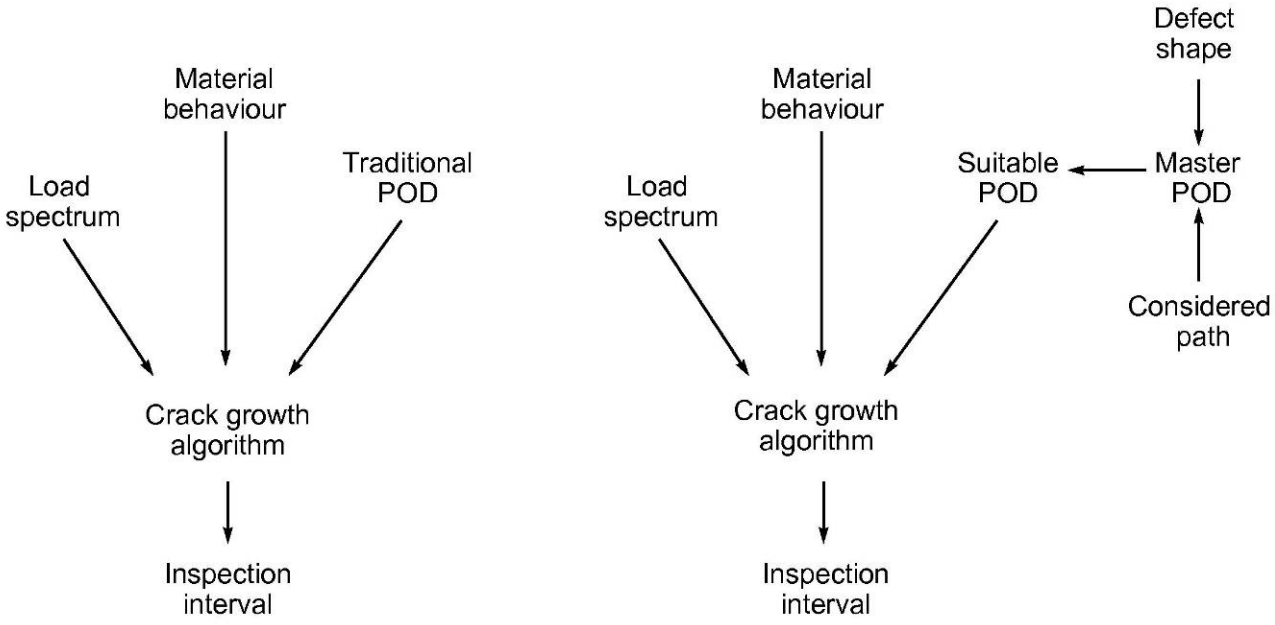


Figure 12 – Effect of POD curve on the determination of inspection intervals.

7. CONCLUDING REMARKS

The present study deepens the matter of the reliability of the in-service UT inspection with particular attention to the special case of railway axles. The conclusions can be so summarised:

- the depth of the defect is not the best parameter in order to entirely characterize its UT response, but it can be more suitable, instead, to consider the area of the defect actually invested by the sound beam;

- defects with different shape, but the same depth, can have completely different POD curves. This fact suggests that, choosing the depth approach, calibration should be carried out on shallow defects for detecting fretting-fatigue cracks and surface damages, while on semi-circular defects for detecting fatigue cracks nucleated in T-transition zones and/or axle body;
- with the new reflecting area approach, the new POD curves are independent from the shape of defects and are also valid for both artificial and natural defects on the body, while more difficulties are to be considered for fretting-fatigue cracks.
- the new POD curves do not, actually, improve the probability to detect defects in respect to the curves in terms of depth, but they may be transformed into those in terms of depth simply assuming proper shapes and values for the area;
- this allows a better estimation of the POD of defects with changing shape during their evolution, like service fatigue cracks which typically nucleate from shallow shape surface damages such as impacts, scratches or corrosion pits and they grow up to become semi-circular.

REFERENCES

[1] U. Zerbst et al.: The development of a damage tolerance concept for railway components and its demonstration for a railway axle, Eng. Fract. Mech., 2003, 72, 209-239.

[2] A. F. Grandt Jr.: Fundamentals of structural integrity, John Wiley & Sons Inc., 2003.

[3] G.A. Georgiou, Probability of Detection (POD) curves: derivation, applications and limitations, Research Report 454, HSE Books, Health and Safety, Executive, UK, 2006.

[4] ASM, ASM handbook – Vol. 17: Non-destructive evaluation and quality control, 1997.

[5] M. Carboni, An analysis of UT echoes coming from fatigue cracks and artificial defects on railway axles, 17th World Congress on NDT, Shangai, China, 2008.

[6] M. Carboni, A new approach to the interpretation of UT echoes coming from artificial notches and fatigue cracks, submitted to NDT&E International, 2009.

[7] S. Cantini, M. Carboni, S. Beretta, NDT reliability and inspection intervals assessment for a safe service of railway axles, Proc. 16th Int. Wheelset Congress, Cape Town, South Africa, 2010.

[8] EN13261: Railway applications – Wheelsets and bogies – Axles – Product requirements, BSI, 2003.

[9] Gilardoni S.p.A., RDG500 detector user's manual, 2006.

[10] WIDEM: Wheelset Integrated Design and Effective Maintenance, website: http://www.widem.org/.

[11] S. Beretta, M. Carboni, M. De Mori, Dynamic Test Bench for Full-Scale tests on Railway Axles - User Manual, Dept. Mechanical Engineering, Quality System, Politecnico di Milano, 2003.

[12] S. Cantini, G. Patelli, S. Beretta, M. Carboni, Assessment of safe life inspection intervals of forged axles/rotors: the influence of in-service NDT reliability (POD curves), Proc. 7th Int. Conf. Eurojoin, Venezia, Italy, 2009.

[13] D. Piotrowski, M. Bode, NDT comparisons of in-service cracks, manufactures cracks and EDM notches, Review of Quantitative Non-destructive Evaluation Vol. 26, ed. by D. O. Thompson and D. E. Chimenti, AIP, 2007.

[14] J.A. Harter, AFGROW users guide and technical manual, AFRL-VA-WP-TR-2006-XXXX, Air Force Research Laboratory, 2006.

[15] C. Rocchi, G. Patelli, C. Gilardoni and co-operators, A new generation UT inspection device for high performance hollow axles, Proc. 15th Int. Wheelset Congress, Prague, Czech Republic, 2007.

[16] S. Cantini, S. Beretta, M. Carboni, POD and inspection intervals of high speed railway axles, Proc. 15th Int. Wheelset Congress, Prague, Czech Republic, 2007.

[17] S. Cantini, A. Ghidini, S. Beretta, M. Carboni, 'Safe life' inspection intervals of railway axles: comparison of crack growth properties of different steel grades, Proc. 14th Int. Wheelset Congress, Orlando, USA, 2004.