Probability of Detection of Ultrasonic In-Service Inspection of Hollow Axles

Michele CARBONI*, Stefano BERETTA*, Stefano CANTINI**,
Cristina GILARDONI***
* Dept. Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156 Milano, Italy; michele.carboni@polimi.it, stefano.beretta@polimi.it
** Lucchini RS SpA, Via G. Paglia 45, 24065 Lovere (BG), Italy; s.cantini@lucchinirs.it
*** Gilardoni SpA, Via A. Gilardoni 1, 23826 Mandello del Lario (LC), Italy; gx@gilardoni.it

Abstract. Hollow axles, widely employed in High Speed and Very High Speed railways applications, are typically inspected by the ultrasonic testing method applied using a suitable boreprobe roto-translating along the longitudinal bore. Different ultrasonic transducers, emitting shear waves at 2-4 MHz and at different refraction angles in steel, are used to inspect the whole external surface of the axle in order to individuate cracks and defects at press-fit seats and at the body, especially along geometrical transitions. The structural integrity of the wheel-set is then based, together with other factors, on the efficient and reliable determination of the Probability of Detection (POD) curve of the boreprobe. For a given probes configuration, the POD curve is function of UT equipment, its calibration and its inspection parameters, therefore the same UT equipment could develop different POD curves as a function of the operating procedure normally defined by wheel-set manufacturers or end users. It is then important to split the absolute performances of the UT equipment by the different performances obtainable with different calibration and operation procedures. In order to derive such curves, a statistically representative sample of natural defects is needed. For this purpose, ten full-scale hollow axles, made of a quenched and tempered alloyed railway steel, were fatigue tested using a dedicated bench in order to develop two natural fatigue cracks in each of them. Different artificial defects were also introduced on the external surface on one of the axles by EDM. A total set of 28 semi-circular defects (17 natural and 11 artificial) was then obtained in a range of depths from 0.4 mm to 12 mm. Such a set allowed the experimental derivation of the POD curve of the boreprobe following the "signal response" approach. A general analysis of different calibration method leads to a comparison between different POD obtainable with the same UT inspecting device. A final part of the research consisted in the optimization of the boreprobe itself, by means of suitable numerical simulations, in terms of frequency, dimension of the transducer, angle of refraction and location of the defect/crack.

1. Introduction

Hollow axles, typically adopted in high-speed railway applications, are inspected, in the last stages of production and during service, using the ultrasonic non-destructive method (UT) and a technique involving a boreprobe roto-translating along the longitudinal bore. Focusing here on in-service inspections, the purpose is to detect any damage, and consequent possible crack propagation up to failure, due to typical phenomena, such as corrosion-fatigue or impact of the ballast, according to a "Damage Tolerant" approach [1]-[2]. From this point of view, the boreprobe considered in this research work (produced by Gilardoni SpA according to the NDT procedures of Lucchini RS SpA [3]) is composed of a series of piezoelectric UT transducers able to inspect the entire outer surface of axles. The structural integrity of axles during service is therefore guaranteed, along with other factors characterizing the Damage Tolerant approach [1]-[2], by the reliable derivation of the "Probability of Detection" (POD) curve [4]-[6] of the boreprobe.

From this point of view, it is also important to estimate the absolute performance of the equipment with respect to its application to different inspection procedures and, especially, to the setting and calibration operations contained in them.

In order to derive such curves, for the considered boreprobe and on the basis of a statistically significant sample, ten full-scale hollow axles made of 30NiCrMoV12 steel grade, according to UNI 6787, were fatigue-stressed in order to generate two natural cracks in each of them and artificial defects were further introduced by EDM on one of them.

The availability of natural cracks, propagated under alternating loads on the test bench, is a peculiarity of the here-presented research: in particular, for natural defects, the reflected energy will be lower than for artificial notches, with consequent risks of underestimating the extent of the damaging phenomenon. This statistical population of natural defects in axles is currently unique and allows to:

- consolidate the POD curves related to the inspection procedures currently required for in-service maintenance and a better estimation of inspection intervals;
- validate new inspection procedures;
- characterize new and/or existing (also for third parties) equipment;
- periodically check the performance of inspection equipment.

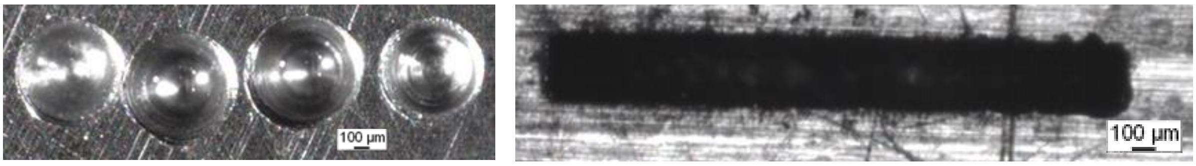
Considering, in this memory, only the first point, the inspection of this set of defects through the considered boreprobe, has allowed the comparison of different POD curves based on different calibration procedures. The final phase of the study then consisted in optimizing the boreprobe by numerical simulations performed using the dedicated software CIVA^{nde} [7].

2. A Set of Fatigue-Cracked Axles

First, it was decided to generate fatigue cracks in the axles through the dedicated test bench available at the Department of Mechanical Engineering - Politecnico di Milano. In particular, cracks were induced at the more dangerous geometrical transitions between the body and the press-fit seat (T-transition, EN13261 [8]), where the stress concentration K_t is approximately 1.2 and the in-service bending moments are at their maximum value.

Ten hollow axles made of 30NiCrMoV12 were, therefore, fatigue-cracked by applying a constant amplitude rotating bending load. In order to obtain two fatigue cracks in every axle, both T-transitions of each axle were prepared by introducing appropriate inclined (5.28°) artificial notches. In particular, the first three axles were prepared by a combination of artificial micro-holes (Pic. 1a) representing possible corrosion pits, while the remaining by semi-circular EDM notches (Pic. 1b) representing possible impacts from ballast or scratches from handling. Finally, at the end of the cracking process, eleven artificial defects, characterized by different morphology (concave and convex), size and position, were introduced by EDM in one of the axle so to enrich the set of defects available for the characterization of the performance of the boreprobe.

At the end of the fatigue tests, 20 defects (characterized by a depth excursion from 0.4 to 12 mm) were available, plus four defects positioned at a different time of flight (on the press-fit seat) and four differently angled ones (on the body).



(a) (b)

Pic. 1. Some example of artificial notches adopted for the cracking of axles.

3. POD Curves for Hollow Axles Made of 30NiCrMoV12 steel

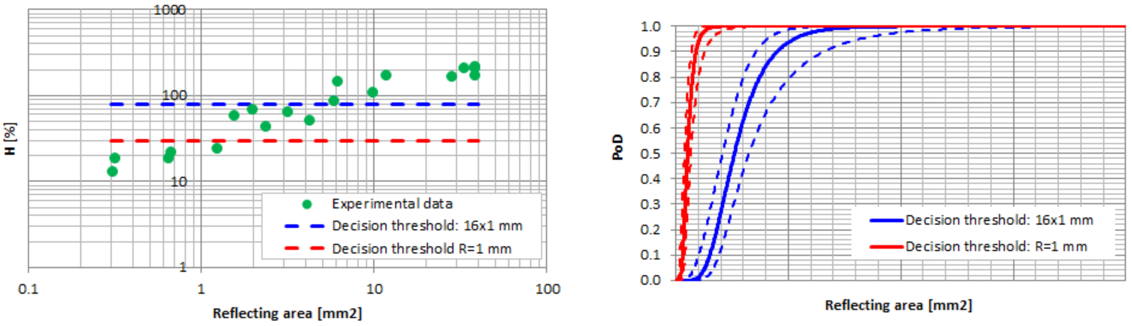
The cracked axles were inspected by the 45° and 38° probes (in both cases oriented along the examination direction or "forward" or along the opposite one or "backward", see Pic. 2a) calibrated using the sample block described in [3] and characterized by 16x1 mm concave defects. The responses from natural cracks were acquired adopting a "signal response" [4]-[6] approach. Furthermore, the boreprobe was introduced, into each axle, from both ends so to inspect any defect from both sides. In particular, since the defects are inclined with respect to the surface, they can be considered, alternately, as "favourable" or "unfavourable" to the inspection (Pic. 2b). Such definitions come from the observation that the relative inclination between probes and defects, in some cases, increases the reflecting area of the defect, in other cases decreases it.

(a) (b)

Pic. 2. Possible approaches of the boreprobe to the inclined defects.

It’s worth noticing that, here, the results will be shown and interpreted considering the innovative approach based on the area of the defect actually invested by the ultrasonic beam [9]. POD curves were then determined by applying the statistical approaches described in [4]-[6], taking into account that, for the analysis, it is simpler and convenient to use data characterized by a growing mathematical trend and then the acquisitions from the boreprobe (in dB) were transformed in echo amplitudes. The reference for this transformation was assumed corresponding to the response of the calibration defect (concave 16x1 mm).

Being the borosonda a complex system formed by different transducers, the data needed for the definition of its POD curve were defined by collecting the best response obtained by the four probes for each inspected defect (Pic. 3a). Picture 3b shows the obtained POD curves and their confidence intervals at 95%. It is important to note that two different decision thresholds (i.e. the response level of the defect which has to be identified at 50%, [4]-[6]) were applied. The first corresponds to the 16x1 mm calibration defect (representing the performance of the boreprobe according to the inspection procedure by Lucchini RS SpA [3]), while the second corresponds to the semi-circular defect having a radius equal to 1 mm (representative of the absolute performance of the boreprobe). Picture 3b shows the obtained POD curves of the boreprobe.

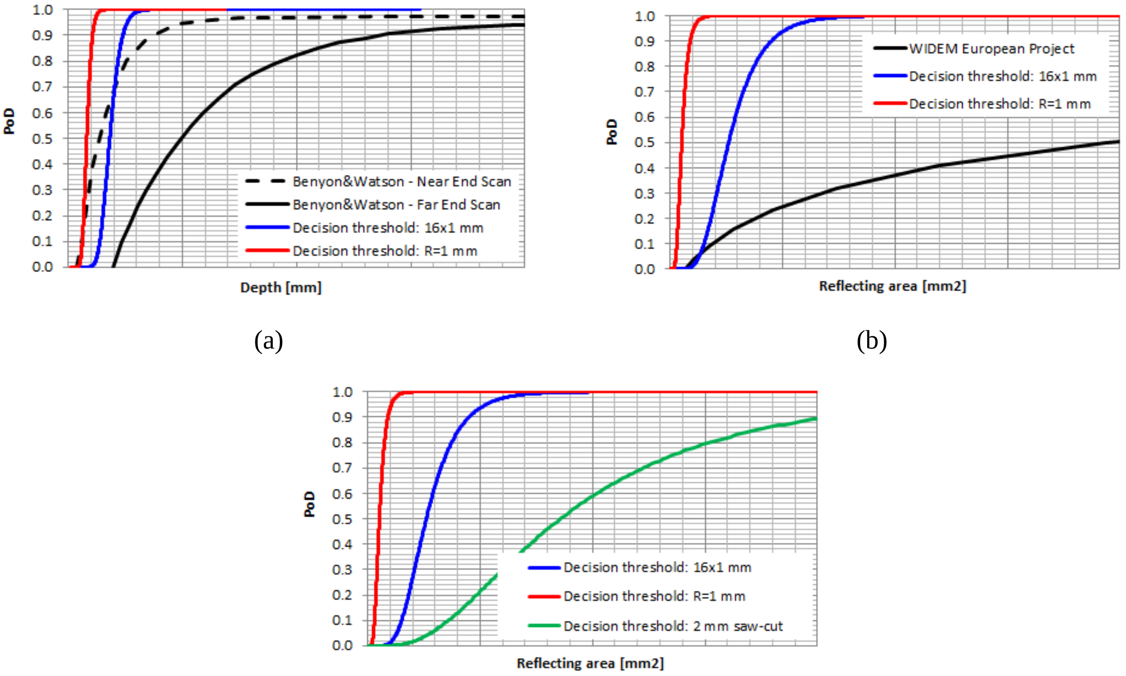


(a) (b)

Pic. 3. POD curves for the boreprobe applied to inclined defects located along the T-transition.

4. Comparison of the Obtained POD Curves to the State-of-the-Art

In order to check the performance of the just-derived POD curves, a comparison with the state-of-the-art (represented by the POD curves by Benyon & Watson [10]) is shown in Picture 4a. It is necessary to add that several details behind the derivation of the experimental curves by Benyon & Watson are not publicly known and that they are only indicative, as they relate to solid axles. It is possible, however, to conclude that the here-obtained POD curves are systematically better than the "far end scan" inspection and comparable with the "near end scan" one.



(a) (b)

(c)

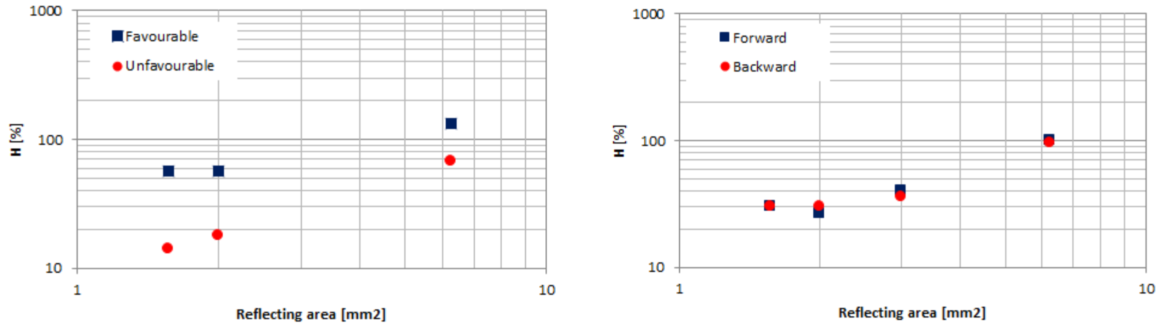
Pic. 4. Comparison of the obtained POD curves to the state-of-the-art.

A second comparison was performed between the just-defined POD curves and that obtained in the frame of the WIDEM European Project [11] for the same boreprobe, but based on a much smaller sample of natural defects. Picture 4b shows this comparison, highlighting how the POD curve obtained in WIDEM seems to be excessively conservative and as the sample size is very important for the derivation of a reliable POD curve.

It was finally carried out a comparison of the here-obtained POD curves to that which would had been obtained adopting, as decision threshold, the typical calibration defect historically used by some European Railways, including DB: the 2 mm deep saw cut. Picture 4c shows this comparison by considering only the mean POD curves for sake of brevity. The results suggest that the calibration defect adopted by some Railways is significantly non-conservative: calibrating the equipment on a 2 mm deep saw cut is equivalent to calibrate it using a semi-elliptical fatigue crack deeper than 4 mm. This is very important, since it has been shown that the 90% of the propagation occurs below the threshold of 4 mm and, beyond a 5 mm depth, the residual life may be shorter than the inspection interval.

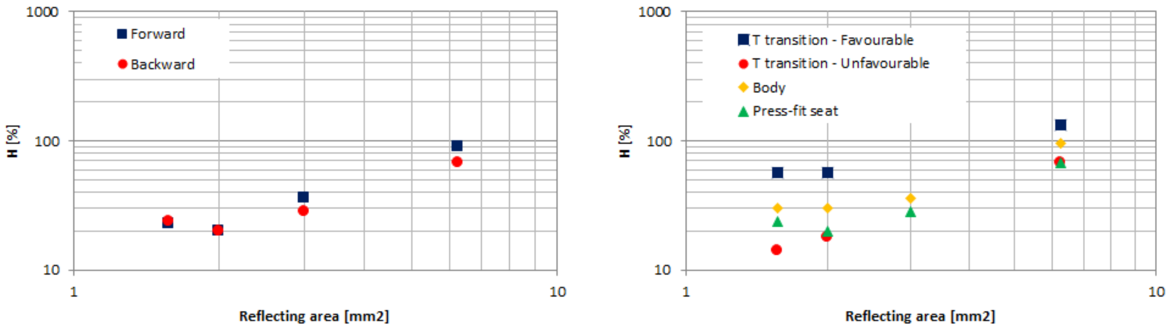
5. Other Exploitations

Some further results were obtained by inspecting the artificial EDM defects. As mentioned in Section 2, they were placed on the central press-fit seat, on the body and inclined along the T-transition. It is interesting to note that the morphology and size of defects were kept the same between these three sections in order to compare the responses of the same defects under different conditions. All artificial defects were, therefore, inspected by the boreprobe using again the probes at 45° and 38° from both ends. Picture 5 shows, as an example, the responses obtained from one of the two ultrasonic probes at 45°. It is important to add that the other probes gave analogous results.



(a) – T-transition

(b) – Body



(c) – Press-fit seat

(d) – Comparison between sections

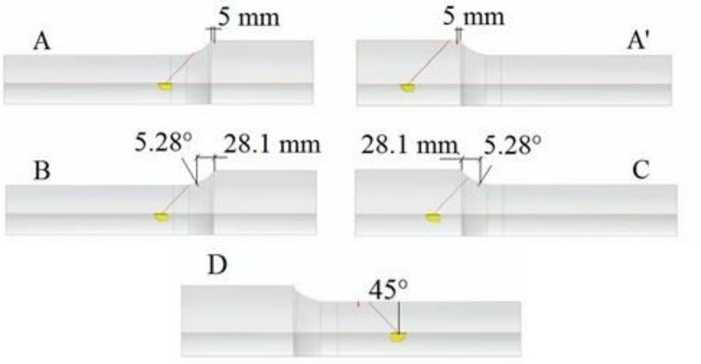
Pic. 5. Inspection of artificial defects by the 45° probe.

Considering the inclined defects along the T-transition (Pic. 5a), the influence of the "favourable" - "unfavourable" configurations, already shown in Picture 2b, is obvious, as approaching the same defects on either side generated significant differences, up to 12 dB, in ultrasonic responses. As expected, however, the not inclined defects (body and press-fit seat, Pic. 5b and 5c), gave ultrasonic responses coincident with respect to the inspection side.

Picture 5d, then, shows a direct comparison between the ultrasonic responses obtained from the three considered sections. The not inclined defects located on the body resulted to respond halfway between the same inclined defects along the T-transition. In particular, it is possible to quantify the response of the ultrasonic defects on the body in about +6 dB with respect to the unfavourable inclined defects and about -6 dB with respect to the favourable inclined ones. It is also possible to observe how the defects on the press-fit seat, characterized by a time of flight greater than the body and T-the transition, respond in a manner coincident with the unfavourable inclined defects along the T-transition, suggesting a similarity of condition between inclination and time-of-flight. These last two configurations were also the most difficult as characterized by the lowest responses.

6. Numerical Optimization of the Boreprobe

The numerical optimization of the boreprobe, defined by the specifications reported in [3], was performed using the dedicated software package CIVA^{nde} 10.0b [7]. The considered scenarios are shown in Picture 6 and are representative of different possible positions of the defect sample (concave 16x1 mm) on the axle.



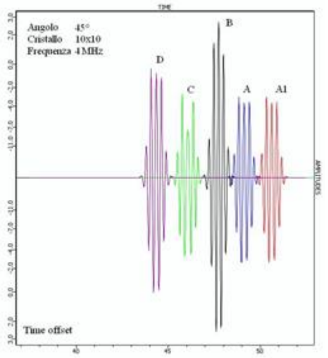
Pic. 6. Scenarios used for the numerical optimization of the boreprobe.

The inspection parameters considered during the optimization were:

- frequency: 2, 4 and 5 MHz;
- refraction angle in steel: 36°, 38°, 42°, 45°, 48° and 50°;
- crystal size: 8x8 mm, 10x10 mm, 12x12 mm and diameter = 11.3 mm (in the latter case, the surface of the circular crystal is equal to that of the square crystal of size 10x10 mm);
- amplitude reference: 0.105, corresponding to the response value of case D in Picture 6.

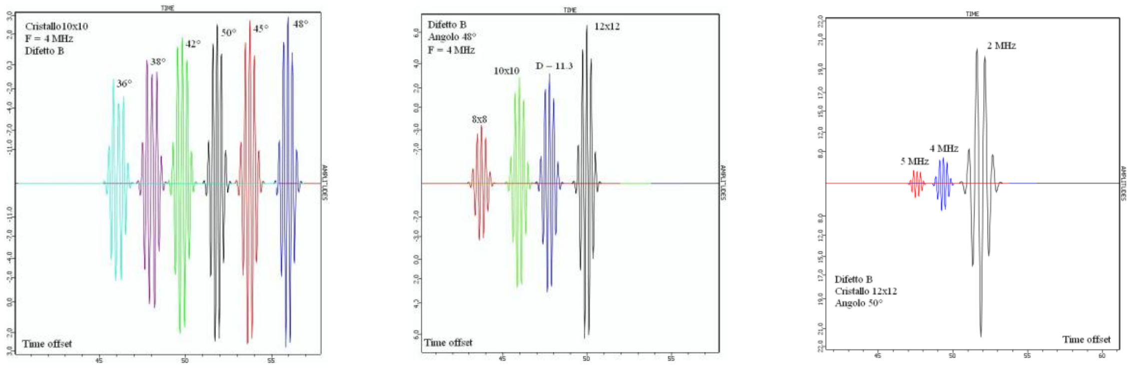
The first series of simulations was carried out in order to verify, by comparison to the experimental results shown in Picture 5, the performance of the software. All possible permutations of the optimization parameters were considered (for a total of 360 simulations) and, as a representative example, Picture 7 shows the case of the currently implemented borosonda (frequency 4 MHz, crystal 10x10 mm and angle of refraction in stele 45°). For a better representation of the obtained results, they were relatively time shifted, so that they are side-by-side rather than overlapping.

The B scenario (T-transition favourably inspected) gave the larger amount of energy. Moreover, as in the experimental evidence, the C scenario (T-transition unfavourably inspected) was the least detectable together with A and A' scenarios, representing the response of the defect in the press-fit seat. Finally, the D scenario (defect in the body) is placed in the middle between B and C. These observations, similarly obtained from all the simulations performed in these analyses, allow assuming the numerical model as properly calibrated based on experimental evidence.



Pic. 7. An example of numerical simulation for validating the model.

Considering, in the following, only the B scenario, the subsequent numerical analyses focused on the optimization of the boreprobe. Picture 8a shows, as an example, the case of frequency 4 MHz, crystal 10x10 mm and a varying angle of refraction. In this particular case, the optimum angle resulted 48°, but generally speaking, it was found to be always 48° (4 MHz) or 50° (at 2 and 5 MHz).



(a) (b) (c)

Pic. 8. Numerical optimization of the boreprobe.

The subsequent analyses (an example is shown in Picture 8b) was targeted to the influence of the size and the shape of the crystal. The best UT response was obtained from the 12x12 mm crystal. It is worth noticing that the circular crystal provided an identical result with respect to the square one of equal area. The last analysis, shown in Picture 8c, compared the influence of the frequency of the ultrasonic signal on the response. It is worth noticing as a decrease of frequency tends to significantly increase the amplitude of the ultrasonic response.

Summarizing, the best-obtained configuration, in terms of amplitude of the UT response, seems to consist in:

- refraction angle in steel equal to 50°
- square crystal 12x12 mm
- signal frequency of 2-2.25 MHz

It is necessary to add that these parameters represent the ideal optimized boreprobe. In the hypothesis of an actual inspection procedure for railway axles, other factors must be considered, forcing the research of a suitable trade-off. In particular:

- the encumbrance due to the typical diameters of the bore-holes prevent, in general, the application of piezoelectric crystals having area 12x12 mm;

- considering the refraction angle, it was possible to note how the numerical responses, in general, do not vary significantly between 45° and 50°;
- as regards the frequency, it has an important influence on the divergence of the ultrasonic beam and, therefore, can affect the extension of the portions of the axle along which the probes are made inactive to not report the indications related to the geometry.

It can be concluded that the performed numerical analysis allowed to identify the optimal configuration for the boreprobe as a tool in itself, but also that the constraints of use and application of the same boreprobe in real inspection procedures suggest a configuration very close to that currently implemented, validating it *de facto*.

7. Concluding Remarks

The most important results obtained from this research work can be described and summarized as follows:

- a set of ten full-scale hollow axles made of 30NiCrMoV12, containing natural fatigue cracks and constituting a statistically significant population currently not available, was prepared for the characterization of the boreprobe;
- the POD curves for Gilardoni's boreprobe were derived both in itself and as applied in Lucchini's inspection procedure;
- the POD curves of the boreprobe were compared to the state-of-the-art obtaining an excellent placement of performance;
- the numerical optimization of the boreprobe was performed using the software CIVA^{nde}: the final configuration was found to be very similar to the current one.

As future developments, the sample of fatigue-cracked axles will be used to characterize other inspection equipment currently on the market, to define new calibration procedures for the European railway operators and to assess reliability and consistency of the today-defined inspection intervals.

References

[1] S. Cantini, S. Beretta (Editors), Structural reliability assessment of railway axles, LRS-Techno Series 4, 2011.

[2] Grandt AF Jr., Fundamentals of structural integrity, John Wiley & Sons Inc., Hoboken, 2003.

[3] G. Patelli et al., Istruzione Tecnica QUA IT 065 Rev.7. 2009, Lucchini RS S.p.A., Lovere (BG), Italy.

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

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

[6] MIL-HDBK-1823A, Nondestructive evaluation system reliability assessment, Department of Defense of the US, 2009.

[7] CEDRAT, CIVA^{nde} 10.0b User's Manual, 2011.

[8] EN13261 – "Railway applications- Wheelsets and bogies – Axles – Product requirements", CEN, 2011.

[9] Carboni M., A critical analysis of ultrasonic echoes coming from natural and artificial flaws and its implications in the derivation of probability of detection curves, Insight 54, 208-216, 2012.

[10] Benyon J.A., Watson A.S., The use of Monte Carlo Analysis to increase inspection intervals, Proc. International Wheelset Congress, Rome, Italy, 2001.

[11] WIDEM EU Project, Wheel-set Integrated Design and Effective Maintenance, Website: www.widem.org.