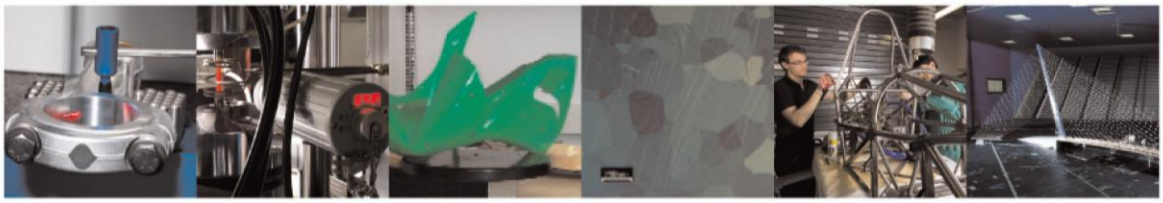
POLITECNICO
MILANO 1863

DIPARTIMENTO DI MECCANICA

mecc



AN ACOUSTIC EMISSION BASED STRUCTURAL HEALTH MONITORING APPROACH TO DAMAGE DEVELOPMENT IN SOLID RAILWAY AXLES

Michele CARBONI, Davide CRIVELLI

This is a post-peer-review, pre-copyedit version of an article published in International Journal of Fatigue. The final authenticated version is available online at:
https://doi.org/10.1016/j.ijfatigue.2020.105753

DIPARTIMENTO DI MECCANICA ■ POLITECNICO DI MILANO
via G. La Masa, 1 ■ 20156 Milano ■ EMAIL (PEC): pecmecc@cert.polimi.it
http://www.mecc.polimi.it
Rev. 0

AN ACOUSTIC EMISSION BASED STRUCTURAL HEALTH MONITORING APPROACH TO DAMAGE DEVELOPMENT IN SOLID RAILWAY AXLES

Michele CARBONI[1*], Davide CRIVELLI[2]

[1]*Politecnico di Milano, Dept. Mechanical Engineering, Via La Masa 1, 20156 Milano, Italy*

[2]*Independent researcher*

Abstract

The in-service safety of railway axles is a very important engineering challenge, as it has a large impact not only from the economic point of view of the railway operator, but it has cascading effects on supply chains, loss of work productivity, and, in the most serious cases, loss of life. It is, therefore, vital that the structural integrity of such components is known, during their lifecycle, with the highest possible accuracy via precise modelling, reliable inspections and, more recently but still at research level, effective condition monitoring.

With a focus on solid freight axles, the research investigates the applicability of Acoustic Emission as a structural health monitoring approach for determining the in-service condition of a full-scale axle. A fatigue crack propagation test is carried out in the lab subjecting the axle to many repetitions of a block load sequence defined from real service measurements. AE data are continuously recorded during the test, whilst crack size is periodically measured by conventional non-destructive techniques.

Eventually, a first-approximation correlation is highlighted between Acoustic Emission data, post-processed by a machine-learning algorithm, and crack propagation ones.

Keywords: structural health monitoring, solid railway axle, acoustic emission, full-scale crack propagation test

[*] Corresponding Author: Tel.: +39-02-23998253, Fax: +39-02-23998202, e-mail: michele.carboni@polimi.it (M. Carboni).

1. INTRODUCTION

In railway vehicles, axles are a fundamental component of the wheelset subsystem and represent safety-critical key elements for railway operations. They are designed against fatigue limit according to relevant standards such as, for example, EN13261 [1] and EN13103-1 [2] in Europe or AAR Section G [3] in North America. On the other hand, in light of a typical design and service life of 30-40 years, the structural integrity of railway axles is prone to service damages [4], like impacts due to the ballast or corrosive aggression, not considered by the cited standardized classical design approach based on the fatigue limit. In particular, such events facilitate the unexpected initiation of fatigue cracks able to cause premature service failures, with important consequences on costs for the infrastructure manager and the railway operator and, in the most serious cases, the onset of accidents and unacceptable fatalities.

To face the abovementioned criticalities, a more effective structural design of railway axles is today commonly looked for complementing the traditional design by more refined approaches based on either fatigue damage assessment or fracture mechanics concepts. Many studies and researches are available in the literature on this topic and, among the most recent ones, a few examples are [5]-[6]. Considering, instead, the in-service structural integrity of railway axles, one of the most successful approaches is Damage Tolerance [7], which is today continuously developed and significantly contributed by, for example, Europe ([4],[8]), China [9] and Japan [10]. By this approach, axles are inspected, at periodical service interruptions suitably planned for maintenance ("inspection intervals"), by well-established [11] non-destructive testing (NDT) techniques, such as visual testing (VT), magnetic particle testing (MT) and ultrasonic testing (UT). Nevertheless, despite the high performance reached by the present application of NDT inspections and by the development of life prediction models, few fatigue-induced failures of axles are still observed. A possible way to further reduce and likely eliminate such in-service failures might be switching the paradigm from the traditional scheduled maintenance to a more affordable on demand maintenance, implementing on-board real time structural health monitoring (SHM) of railway axles. From this point of view, the fact that SHM has shown [12]-[13], especially in the aeronautical and civil fields, the possibility to decrease the costs of preventive damage inspections by 30%, without losing accuracy, should not be underestimated, as well.

To date, very few investigations are available in the literature about SHM applied to railway axles. Even if still at a preliminary stage, the most advanced and promising approaches seem to be acoustic emission [14]-[17], low frequency vibrations (revolution periodicity) [18], high frequency vibrations (changes in natural frequencies) [19] and automated UT [20]. Other methods recently started to be taken into account or just suggested to be applicable to the problem at hand: microwave testing [21], alternating current field measurements [22], induced current focusing potential drop [22], alternated current thermography [23] and laser-air coupled hybrid UT [24]. Each of these approaches has specific pros and cons, but, for the sake of brevity, the reader is referred to the specific literature for details. Nevertheless, the present study focuses on Acoustic Emission (AE) because, up to date, it seems to be the most advanced and because the equipment is less invasive and less bulky. This is a very important and critical advantage considering the configuration of real axles mounted on running trains. Moreover, it is worth adding that, in the railway field, AE is already successfully applied for monitoring damage development in wheels, bearings and welded bogies, suggesting the possibility to take advantage of a synergic approach for the entire rolling stock, and rails.

In the industrial field, SHM by AE is traditionally applied to assess the structural integrity of metallic components, especially pressure vessels and pipelines, under static and fatigue loading [25]. It originated transposing the concepts of seismology to a different (smaller) scale [26] and is based on the observation that, when damage develops in a material (plasticization, fracturing, …), it releases energy as ultrasonic elastic waves (“micro-earthquakes”), which can be fruitfully detected and interpreted. Such waves (the so-called “hits”, Fig. 1) are typically short and transient (“burst events”) and are characterized by a bandwidth in the 100-1000 kHz range, which makes AE quite robust against audible noise and structural vibrations. Historically, the so-called “parametric AE” approach was used in order to decrease the amount of data to be managed: the transient waveforms were discarded after being “fingerprinted” using some well-established features (Fig. 1: amplitude, duration, energy, …) for further processing. Today, with the availability of higher sampling frequencies and the increase of data storage capabilities, the full waveforms can be recorded in order to get the highest flexibility during post-processing and the analysis of the phenomenon, but new issues arise. In particular, a first critical point of AE monitoring is the typical amount

of recorded data, which, even considering short campaigns, can easily sum up to millions of events. Secondly, it is reasonable to expect the recorded data are not just directly related to primary emissions (damage phenomena), but also to different types of secondary emissions (background noise and interference). These two inherent characteristics directly put AE monitoring and analysis into the field of Big Data Analytics [27], i.e. a highly multidimensional problem with little *a priori* knowledge of the underlying data structures, and suggest AE data is the perfect candidate for machine learning classification algorithms.

The present paper investigates the application of SHM by AE to the special case of a deep-rolled solid axle subjected to a variable amplitude full-scale crack propagation test. AE data are continuously recorded during the test, whilst crack size is periodically measured by conventional non-destructive techniques. Eventually, a first-approximation correlation is highlighted between Acoustic Emission data, post-processed by a machine-learning algorithm, and crack propagation ones.

2. EXPERIMENTAL FULL-SCALE PROPAGATION TEST

All the details and results about the full-scale tests carried out, in the frame of the MARAXIL Research Project [28], on deep-rolled axles can be found in [29], while a short summary, useful as a background for the present research, is given in this section.

2.1. Specimen and details on the full-scale crack propagation test

The feasibility of AE as a SHM approach to railway axles was investigated considering a solid full-scale specimen, shown in Figure 2 and named “axle” in the following, designed according to [1] and made of the standardized EA4T medium strength steel grade (quenched and tempered 25CrMo4, [1]). It is worth remarking that the geometry of the specimen does not correspond exactly to that of a real axle, but its central area (press-fit seat, T-transitions and the two portions of the adjacent cylindrical body) is fully representative of the regions of real axles where the wheels are press-fitted and of the adjacent areas [1]. Moreover, a portion of the axle was treated by the deep-rolling procedure typically used, in Europe, for some high-speed applications. The full-scale axle was, then, subjected to a variable amplitude fatigue crack

propagation test. In order to effectively control the initiation site of two independent fatigue cracks, the section of interest of the axle (Fig. 2a and b) was characterized by two artificial notches, manufactured by electro-discharge machining (EDM) in the deep-rolled cylindrical region of the axle, located at 180° one from the other. Such notches had a semi-circular shape with radius R = 4 mm. This particular size was chosen, based on the measured residual stress profile [29], in order to overcome the compressive stress state at the surface and to allow a significant propagation of the cracks during the test.

The full-scale crack propagation test was carried out by means of the "Dynamic Test Bench for Railway Axles" (BDA) available at the labs of the Department of Mechanical Engineering at Politecnico di Milano (Fig. 3a and b). The bench applies, to the tested full-scale specimen, a three-point rotating bending condition by an actuator group (having capacity equal to 250 kNm) and an electrical engine: both fatigue (durability) and crack propagation tests can be carried out applying either constant amplitude or variable amplitude block loads, as described in [30]. The BDA bench is fully compliant to the relevant European standard on the qualification of railway axles [1] and both the fatigue test procedure and the crack propagation results are Quality Certified according to ISO/IEC 17025 [31].

In the present research, the full-scale axle was subjected, by the BDA bench, to a block loading sequence (Fig. 3c) experimentally derived from the typical service spectra of a tilting train on European lines [32] and representing about 57000 km of high-speed service. Eventually [29], the full-scale test lasted for 64 repetitions of the block load sequence, corresponding to an equivalent distance of about 3.5×10^6 km, and was interrupted before the final failure of the specimen.

2.2. Characterization of crack initiation and propagation by non-destructive testing

With the aim to get a feedback on crack initiation and propagation and to evaluate the effectiveness of structural health monitoring by AE, NDT was applied at suitable interruptions of the test, which were planned between the repetitions of the block load sequence. In particular, the size of the cracks originating from the artificial notches was characterized by visual testing (VT) and phased array ultrasonic testing (PAUT).

The rationale behind the application of VT was to size the surface length of cracks. Two different approaches were adopted: digital image acquisition by an optical microscope ("OM" in Figure 2a) and plastic replicas ([33], "PR" in Figure 2a). Eventually, the information provided by the two methods corresponded and, as an example, Figure 4a shows the plastic replica, observed at the microscope, of one of the surface tips of crack A at the end of the test (repetition #64 of the block load sequence). Figure 4b, instead, summarizes the complete trend of crack advance at the surface for crack A, while crack B showed analogous results. As can be seen, no crack initiation could be observed up to block load repetition #15, after which multiple cracks actually initiated and propagated from the artificial notches. Such cracks developed up to a very limited size (about 140 μm) and became non-propagating after about 35 repetitions of the block load sequence. This peculiar behaviour was ascribed to and explained by the influence of deep-rolling residual stresses on crack driving force [29].

On the other hand, the rationale behind the application of PAUT was to size the depth of cracks. To this aim, a Harfang X32 ultrasonic phased array unit, equipped with a linear probe characterized by 32 elements and a central frequency equal to 5 MHz, was adopted. Cracks were inspected using a sectorial scan visualization (S-Scan with refraction angle ranging from 35° to 55°) and sized by means of the ''crack tip diffraction'' method ([34], Fig. 4c). Figure 4d summarizes the trend of crack depth for both crack A and B. Before discussing the results, it is worth noting the significant scatter of PAUT measurements: actually, such a scatter was expected because, even if UT is the only volumetric NDT method applicable (and affordable) to inspect the volume of railway axles, its performance to size defects is one of the most controversial topics. No one UT sizing technique has yet been established which gives a high degree of accuracy or repeatability in all circumstances [35]. Actually, the low accuracy (high uncertainty) of UT sizing and the lack of repeatability of UT measurements (inherent high measurement errors) are very well-known in the NDT field and are confirmed, for the case at hand, by the data shown in Fig. 4d. Nevertheless, as can be seen, both an initiation stage and a propagation one are highlighted again: due to the aforementioned uncertainty of the measurements, the border between these two regions was estimated by the AE results discussed in Section 3.

Considering the initiation stage (Fig. 4d), the experimental average estimations of the depth of the initial notch (nominally equal to 4 mm) resulted to be equal to 4.1 mm (standard deviation equal to 0.26 mm) and to 4.14 mm (standard deviation equal to 0.28 mm) for cracks A and B, respectively. Assuming the nominal 4 mm depth value as a reference (the supplier certified the size of the EDM notches with an uncertainty equal to ±0.1 mm), the mean measurement error by PAUT resulted to be between 0.1 mm and 0.15 mm. Moreover, a 95% confidence analysis of the experimental mean values (not shown in Fig. 4d for the sake of clarity) showed an almost complete superposition of the confidence bands, which also included the reference one, and, consequently, allowed to conclude there is a statistical consensus between them and the reference one.

Regarding the propagation stage, the total crack growth resulted to be at a nearly constant rate and about 1 mm long for both cracks, which propagated up to about 5 mm during the whole fatigue test. Data were fitted by the least squares method trying different mathematical models (linear, logarithmic, exponential and power): the best determination coefficients were achieved by the linear model, which resulted to be characterized by R^2=0.62 and R^2=0.58 for cracks A and B, respectively. Such best values denote a non-optimal, but rather reasonable and acceptable linear fitting. As for the case of the initiation stage, a 95% confidence analysis of the linear fittings was carried out, finding again a statistical consensus between the behaviour of the two cracks.

3. STRUCTURAL HEALTH MONITORING OF RAILWAY AXLES BY ACOUSTIC EMISSION

AE signals were continuously recorded, during the entire full-scale crack propagation test, to monitor damage development. In particular, an AE piezoelectric sensor (Vallen VS150-M with resonance frequency equal to 150 kHz) was applied at the free end of the axle (Fig. 2a) using a custom made mount designed to hold it in position together with a pre-amplifier (Vallen AEP4 34 dB). The AE sensor was coupled to the surface of the axle using bearing grease in order to prevent the formation of air bubbles and to maximize sensitivity: such an assembly was periodically checked to ensure the grease was not deteriorated or moved away during the test. In this bench test, the use of a single AE sensor was forced by the presence of the engine coupling at one end of the axle. In a real application, the use of two sensors (one at each end of the

axle) would allow localizing AE signals along the axle improving the accuracy of defect detection, but this remains an open point and a future development of the present research.

The rotating measuring group was linked to a sliding contact (Michigan Scientific S4) to send the signals to the AE acquisition system (Vallen AMSY-5). The use of a sliding contact, which is simple and cheap, explains the need of the adopted preamplifier: indeed, the signal to noise ratio resulted to be very low because, without preamplifier, the background noise generated by the sliding contact was of the same order of magnitude ([mV]) of the acquired AE signals. Finally, together with AE events, the signal of the applied load over time was acquired, as well, in order to synchronize the outputs.

The AE system was calibrated according to the standard pencil lead break test [36]: signal attenuation along the axle was found to be negligible. After one minute of rotation of the specimen without any load application, the acquisition threshold of AE events was set to 55.8 dB in order to filter noise under the assumption that, in the absence of load, no events related to crack propagation are expected. Moreover, since AE hits related to fatigue damage in homogeneous and isotropic materials have the typical burst morphology, a band pass acquisition filter was set in the 230 - 850 kHz frequency range, with the lower frequency set above the sensor's resonance (150kHz), and the upper frequency set to accommodate a sampling frequency of 5 MHz.

3.1. Analysis of raw events recorded by acoustic emission

During the whole test, the AE set-up recorded 437268 hits. Figures 5a and b compare AE activity, in terms of number of hits and their amplitude, at the beginning of the test (block load repetition #2) and at its end (block load repetition #64). A significant increase of such an activity is evident and can be interpreted as follows.

First, the sources of background noise and interference should be nearly time-independent and keep almost constant during the test: no increase of AE activity should have been observed if crack initiation and propagation were absent. Nevertheless, during the very first repetitions of the block load sequence, like the one shown in Figure 5a, the recorded AE events can be reasonably related to background noise only. In the

case at hand, the events are very few due to the suitably chosen acquisition threshold (55.8 dB) and the applied passband filter (230 - 850 kHz).

A significant and peculiar increase of AE activity started to be observed at block load repetition #4 (Fig. 6a). It can be clearly noticed that there is a strict correlation between the number of hits, their amplitude and the load level: during the load cycles with low load levels, the AE activity is relatively low and almost constant, whereas, during the load cycles with high load levels, it increases instantly and significantly. It has to be also noted that the activity is globally very low during most of the time, with events of moderate amplitude, comparable to the inherent background noise shown in Figure 5a. Since no cracks were detected by NDT up to block load repetition #15, the whole set-up (test bench and measurement chain) did not show any malfunctioning and the test conditions (load levels, temperatures, …) did not justify any structural modifications of the material of the axle, the recorded peaks of AE activity could be ascribed [37] to plasticization and sliding of dislocations in the process zone at the tips of the artificial notches. Figures 6b, c and d confirm such behaviour up to block load repetition #15, with a continuous increase of events in between load transitions, as well. It is worth remarking that the AE activity detected during this (crack initiation) stage seems able to provide a warning about damage development well before the NDT approaches traditionally applied to railway axles. The amplitude doesn't show repeatability across graphs, which is an observation well in line with the Kaiser effect and the nonlinear nature of crack propagation.

Based on the outcomes of NDT inspections, the last (crack propagation) stage is here assumed to start at block load repetition #16, even if it is likely micro-cracks, not detectable by NDT, could have appeared during the previous couple of block load repetitions. Nevertheless, during crack propagation, the size of the crack front increases in length and more point sources of acoustic events are made available (Fig. 7a), suggesting the onset of more sparse AE hits with respect to the shape of the load sequence. This is confirmed by Figure 7b, where the AE activity of block repetitions from #16 to #18 is shown. As can be seen, the AE events correlated to load levels are almost disappeared, while more uniformly distributed hits have taken place. The same can be observed later on during the test (Fig. 7c) and at its end (Fig. 5b). Finally, it is worth noting the ability of SHM by AE to highlight the onset of load interaction effects on crack propagation. Figure 7b, and Figure 7c to a limited extent due to the high number of recorded events, clearly

show that, during the application of the low load levels just following the high ones, AE activity decreases significantly and gradually increases again with the accumulation of low amplitude fatigue cycles. Since a decreased AE activity may be linked to a slower crack growth, this seems to suggest the onset of crack growth retardation due to application of the chosen block load sequence: actually, this corresponds exactly to the results shown in [29] with respect to the crack propagation behaviour observed during the present full-scale test.

3.2. Unsupervised classification of acoustic emission events

As previously stated, AE signals can originate from different sources, not all of which are of interest in a SHM problem. In this research, in order to classify the recorded events, AE parameters (amplitude, rise time, duration, counts and energy, see Fig. 1 for their definition) are used as an input to a Self-Organizing Map (SOM) [38], a particular type of neural network, which tends to group similar inputs together. However, as the number of groups of similar signals is unknown *a priori*, the results are run through a k-means algorithm assuming an increasing number of classes. The classification quality is evaluated using different fitness criteria (Davies-Bouldin, Silhouette, Calinski-Harabasz) which are, then, used to determine the optimal number of classes in the dataset. The full details on the adopted classification algorithm are described in [39], while its full validation is reported in [40].

As the classification algorithm was run on the whole AE dataset recorded during the test, two classes of events were found: Class #1 gathering 87% of the total (380423 hits out of 437268) and Class #2 gathering the remaining 13% (56845 out of 437268 hits). Figure 8a shows, as an example, the classification obtained for block load repetitions #17 and #18. As can be seen, the amplitude of Class #1 is generally low (55-65 dB) and correlated with load levels, whereas the amplitude of Class #2 seems more evenly distributed at all load levels and has, overall, a higher maximum value (up to around 75 dB). In order to provide a physical meaning to the two classes of events, the first step was to check the morphology of the classified waveforms. From this point of view, always considering the exemplificative case of block load repetitions #17 and #18, an example of typical Class #1 waveform is shown in Figure 8b, while Figure 8c shows an

example of typical Class #2 waveform. As can be seen, the first impression is that Class #1 gathers background noise, while Class #2 gathers burst events likely related to damage development.

Another useful and commonly adopted feature for analysing the meaning of AE classes is their cumulated energy, which is a measure of the energy introduced into the system. Accordingly, Figure 8d shows the cumulated energy of the two classes against the applied load for the exemplificative case of block load repetitions #17 and #18. One can see that, despite Class #1 events being far more numerous than Class #2 events, their cumulated energy level is generally lower. This is very informative because, typically, damage development is characterized by a much higher energy content than background noise. In addition, as stated in Section 3.1, noise should remain nearly constant over time and, consequently, show a linear trend in terms of cumulated energy. From this point of view, the trend of Class #1 is actually nearly linear with time, while Class #2 shows, especially during block load repetition #18, sudden significant increments likely related to damage development and crack advance.

It can be concluded that there is a reasonable guarantee that Class #1 is mainly related to background noise, while Class #2 mainly to damage development. Consequently, just the events included in Class #2 will be considered in the following Section.

3.3. Discussion on the performance of classified acoustic emission

Figures 9a shows the trend of the number of hits for each repetition of the block load sequence. Considering the whole test, the discussion presented in Section 3.1 is confirmed: the activity was very low at its beginning, the highest during the initiation stage, but with an uncertain and very dispersed trend, and in-between during the propagation stage with a nearly linear increasing trend, even if quite dispersed again. A clear identification, based on AE activity, of the three stages of damage development (initial noise, crack initiation and crack propagation) might seem a good opportunity for a successful SHM of railway axles, but, actually, the very high level of uncertainty (dispersion of data) makes this approach subtle due to the consequent high probability of false positive and false negative results.

Looking at the trend of the average amplitude of AE events for each repetition of the block load sequence (Fig. 9b), a clearer interpretation is possible. Indeed, in this case, it is easy to identify the different

stages, because the initiation phase is characterized by significantly higher amplitudes than the other two. This observation is aligned to the literature [37]: the initiation stage of cracks usually provides a higher AE activity characterized by a higher amplitude compared to the propagation stage, which is reported to be rather quiet. Consequently, the amplitude of AE events might be a good criterion for an early detection of damage. Nevertheless, the amplitude level during the propagation stage is constant and comparable to the one of the initial noise. This might prevent crack sizing and raises issues in real applications where the noise level is expected to be higher than in the lab.

Figure 9c shows the trend of the average energy of AE events for each repetition of the block load sequence. In this case, it is not possible to identify the three stages because no peculiar behaviours are observed, but the general linear trend show the closest match to crack propagation stages, if compared to the activity and the average amplitude. This means that energy can provide useful information about the presence of damage development, but none on the type or stage of such a damage.

Summarizing, a simple criterion based on AE amplitude might be effective for pointing out the onset of crack initiation: a suitable amplitude threshold should be calibrated on real applications and a warning alarm released when it is reached and overtaken. On the other hand, the most descriptive and robust AE feature for crack propagation, where the key point is a reliable crack sizing, seems to be the energy and a more detailed analysis is required.

4. A FIRST-APPROXIMATION ACOUSTIC EMESSION BASED APPROACH TO CRACK SIZING

AE energy can be represented in many different ways: instantaneous value, total cumulated trend during a given monitoring campaign, block cumulated trend during a fixed time window, etc. All of these possibilities were analysed in order to individuate the most effective one for the study at hand. Eventually, a promising correlation between crack depth and the (logarithmic) AE energy cumulated during each block load repetition, related to the crack propagation stage, could be found (Fig. 10a). It is worth noting that just crack depth was considered as a relevant descriptor of crack size because, as explained in Section 2.2, crack propagation at the surface was strongly influenced (impeded) by the compressive residual stress profile and, consequently, was assumed to be negligible. Moreover, it was not possible to separate the AE events

related to the propagation of each individual crack, but, since they behaved in a very similar way (Fig. 4d), an average crack depth value is here considered as a first approximation.

The empirical relationship between crack depth and the logarithmic AE energy cumulated during each block load repetition was built as follows:

1. AE data from Class #2, previously identified by the classification algorithm, is isolated;
2. the cumulated AE energy related to each block load repetition (E_{block}) is calculated;
3. a linear model is fit to the data:

$$a = \mathrm{C}_0 + \mathrm{C}_1 \cdot \log_{10}(E_{block}) \quad (1)$$

where "a" is the average crack depth. The results obtained by the proposed linear model are shown in Figure 10b. Despite the relatively wide confidence band, which can be explained by the relatively large uncertainty in the crack measurement data, the overall performance of the correlation is reasonable. The R^2 of the linear regression is 0.39, which may appear not enough to provide a traditionally "reliable" model of crack propagation, but is still an indication that such a relationship exists. Both coefficients C_0 (2.671) and C_1 (0.244) have a p-value of less than 0.01 (the tested hypothesis is that the coefficient is equal to 0, i.e. the model is a constant trend): this indicates a strong evidence of the presence of a relationship, with the relatively low R^2 indicating moderate levels of noise/uncertainty.

As a further verification of the claims on classification results given in Section 3.2, the correlation was separately defined for Class #1 data, as well. In this case, the correlation fitted the data poorly, the value of R^2 dropped to 0.24 and the p-value of C_1 rose to 0.03. This information is a further indication that Class #2 signals are indeed related to crack propagation, whilst Class #1 events are linked to sources of background noise.

As a remark, the proposed classifier is trained to the specific axle's data, and, due to its nature, every different combination of axle geometry and material (e.g. hollow versus solid) may affect the signals' propagation path and characteristics. This could mean that a classifier trained on a solid axle will not be directly transferrable on a hollow axle, or on a solid axle made of a significantly different material. However, as the "transfer function" between the source and the sensor will be the same for any category

of signals within the same axle, it is expected that the proposed method would still be applicable, as it will be able to accurately classify different signals. The linear model coefficients may require further analysis before directly extending them to a different geometry.

Finally, it is worth mentioning that the proposed first-approximation empirical (data-driven) correlation is meant to be just the first promising step for the future development of a full crack initiation and propagation (physics-based) tool aimed to build a set up for the diagnostics and the prognostics of in-service railway axles. Accordingly, more dedicated full-scale crack propagation tests are being carried out.

5. CONCLUDING REMARKS

The objective of this study is to highlight the feasibility of using Acoustic Emission for the real-time structural health monitoring of in-service railway axles with initiating and developing damage.

The physical nature of the collected data was clearer after the application of unsupervised classification, excluding effectively signals due to background noise. Consequently, the possibility to clearly identify and separate the initiation and propagation stages of damage development was observed, providing an important advantage when compared with the conventional application of non-destructive testing. Focusing on the propagation stage, an empirical and reasonable relationship, linking suitable AE features to the measured crack size, was found. Further investigations, including two sensors setups and more full-scale tests, can further improve and validate the proposed approach.

In conclusion, the work shows that there is certainly scope for using Acoustic Emission to reduce the frequency and the cost of periodic non-destructive testing inspections.

ACKNOWLEDGMENTS

The full-scale crack propagation test was carried out in the frame of the MARAXIL (‘‘Manufacturing Railway Axles with Improved Lifetime’’) Research Project co-funded by Regione Lombardia (ID 16973, Rif. n° MAN-15). The Authors would like to thank Prof. S. Beretta, Prof. S. Bruni, Prof. M. Guagliano and Dr. P. Rolek for the given support and useful discussions and Mr. G. Cunati for the active help given during the experimental crack propagation test.

REFERENCES

[1] EN 13261. Railway applications – Wheelsets and bogies – Axles – Product requirements. CEN, 2010.

[2] EN 13103-1. Railway applications – Wheelsets and bogies – Part 1: Design method for axles with external journals. CEN, 2017.

[3] AAR Manual of Standards and Recommended Practices. Section G: Wheels and Axles. AAR, 2011.

[4] U. Zerbst, M. Vormwald, C. Andersch, K. Mädler, M. Pfuff. The development of a damage tolerance concept for railway components and its demonstration for a railway axle. Eng. Fract. Mech. 72 (2003) 209-239.

[5] P. Pokorný, T. Vojtek, L. Náhlík, P. Hutař. Crack closure in near-threshold fatigue crack propagation in railway axle steel EA4T. Eng. Fract. Mech. 185 (2017) 2-19.

[6] Z.W. Xu, S.C. Wu, X.S. Wang. Fatigue evaluation for high-speed railway axles with surface scratch. Int. J. Fat. 123 (2019) 79–86.

[7] A.F. Grandt Jr. Fundamentals of Structural Integrity, John Wiley & Sons, 2004.

[8] M. Carboni, S. Beretta. Effect of probability of detection upon the definition of inspection intervals of railway axles. Proc. Instn Mech. Engrs, Part F: J. Rail and Rapid Transit 221 (2004) 409-417.

[9] S.C. Wu, S.Q. Zhang, Z.W. Xu, G.Z. Kang, L.X. Cai. Cyclic plastic strain-based damage tolerance for railway axles in China. Int. J. Fat. 93 (2016) 64–70.

[10] T. Makino, H. Sakai, C. Kozuka, Y. Yamazaki, M. Yamamoto, K. Minoshima. Overview of fatigue damage evaluation rule for railway axles in Japan and fatigue property of railway axle made of medium carbon steel. Int. J. Fat. 132 (2020) 105361.

[11] EN 15313. Railway applications – In-service wheelset operation requirements – In-service and off-vehicle wheelset maintenance. CEN, 2016.

[12] F.K. Chang. Structural Health Monitoring: A Summary Report. Proc. 2nd International Workshop on Structural Health Monitoring, Stanford, US, 1999.

[13] C. Boller, Ways and options for aircraft structural health management, Smart Mater Struct 10 (2001) 432–440.

[14] X.J. Deng, G.J. Xui, S.Q. Liu. Research on Fatigue Crack Detection of Rail Vehicle Axle Based on Acoustic Emission. Proc. 10th International Workshop on Structural Health Monitoring, Stanford, US, 2015.

[15] C.H. Jiang, W. You, L.S. Wang, M. Chu, N. Zhai. Real-time monitoring of axle fracture of railway vehicles by translation invariant wavelet. Proc. International Conference on Machine Learning and Cybernetics. Guangzhou, China, 2005.

[16] C.H. Jiang, S. Pan, C. Zhang, F. Li. Experimental Research on Fault Location for the Axle of Railway Vehicles Based on Acoustic Emission Technique. Int. J. Cont. Autom. 9 (2016) 91-98.

[17] Y. Zhou, L. Lin, D. Wang, M. He, D. He. A new method to classify railway vehicle axle fatigue crack AE signal. Applied Acoustics 131 (2018) 174-185.

[18] P. Rolek, S. Bruni, M. Carboni. Condition monitoring of railway axles based on low frequency vibrations. Int. J. Fatigue 86 (2016) 88–97.

[19] M.J. Gómez, C. Castejón, J.C. García-Prada. New stopping criteria for crack detection during fatigue tests of railway axles. Eng. Fail. Anal. 56 (2015) 530–537.

[20] S.Y. Chong, J.R. Lee, H.J. Shin. A review of health and operation monitoring technologies for trains. Smart Struct. Syst. 6 (2010) 1079–1105.

[21] V. Gorbunov, V. Sutorikhin. Microwave Nondestructive Testing Method. Applied Physics Research 4 (2012) 206-210.

[22] N. Bachschmid, P. Pennacchi, E. Tanzi. Cracked rotors: a survey on static and dynamic behaviour including modelling and diagnosis. Springer Science & Business Media, 2010.

[23] R. Zoughi, S. Kharkovsky. Microwave and millimetre wave sensors for crack detection. Fat. Fract. Engng Mat. Struct. 31 (2008) 695-713.

[24] R. Ngigi, C. Pislaru, A. Ball, F. Gu. Modern techniques for condition monitoring of railway vehicle dynamics. Journal of Physics: Conference Series 364 (2012) 1-12.

[25] M. Huang, L. Jiang, P.K. Liaw, C.R. Brooks, R. Seeley, D.L. Klarstrom. Using acoustic emission in fatigue and fracture materials research. JOM 50 (1998) 1–12.

[26] C. Grosse, M. Ohtsu. Acoustic Emission Testing, Springer-Verlag, Berlin Heidelberg, 2008.

[27] A. Deshpande, M. Kumar, Artificial Intelligence for Big Data, Packt Publishing, 2018.

[28] MARAXIL project. http://maraxil.mecc.polimi.it/, 2013. (accessed on July 2019).

[29] D. Regazzi, S. Beretta, M. Carboni. An investigation about the influence of deep rolling on fatigue crack growth in railway axles made of a medium strength steel. Eng. Fract. Mech. 131 (2014) 587–601.

[30] S. Beretta, M. Carboni. Variable amplitude fatigue crack growth in a mild steel for railway axles: experiments and predictive models. Eng. Fract. Mech. 78 (2011) 848-862.

[31] ISO/IEC 17025. General requirements for the competence of testing and calibration laboratories, ISO, 2017.

[32] S. Beretta, M. Carboni, S. Cervello. Design review of a freight railway axle: fatigue damage versus damage tolerance. Mat-wiss U Werkstofftech 42 (2011) 1099-1104.

[33] M.H. Swain. Monitoring small crack growth by the replication method, ASTM STP 1149, 1992.

[34] J. Krautkrämer, H. Krautkrämer. Ultrasonic testing of materials, 4th edition, Springer-Verlag, 1990.

[35] J.C. Drury. Ultrasonic flaw detection for technicians, 3rd edition, Silverwing Limited, 2004.

[36] ASTM E976. Standard Guide for Determining the Reproducibility of Acoustic Emission Sensor Response, ASTM International, 2010.

[37] R.K. Miller, E.v.K. Hill, P.O. Moore. Nondestructive Testing Handbook, Third Edition: Volume 6, Acoustic Emission Testing (AE). ASNT Publications, 2005.

[38] S. Haykin. Neural Networks and Learning Machines. Pearson International Edition, 2009.

[39] D. Crivelli, M. Guagliano, A. Monici. Development of an artificial neural network processing technique for the analysis of damage evolution in pultruded composites with acoustic emission. Composites: Part B 56, (2014) 948–959.

[40] D. Crivelli. Structural health monitoring with acoustic emission and neural networks. PhD Thesis, Politecnico di Milano, Milano (Italy), 2014. Open access at https://www.politesi.polimi.it/handle/10589/89521.

List of Figures

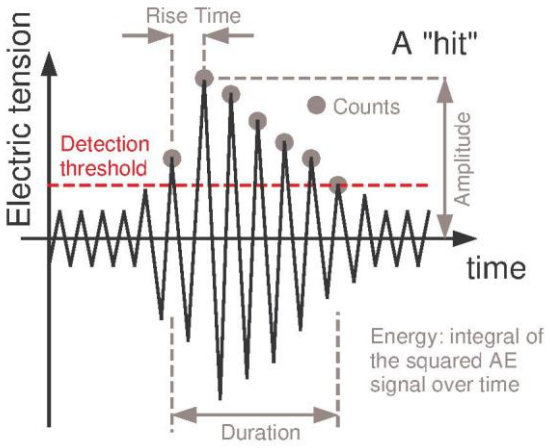

Rise Time
A "hit"
Electric tension
Counts
Detection
threshold
Amplitude
time
Energy: integral of
the squared AE
signal over time
Duration

Fig. 1

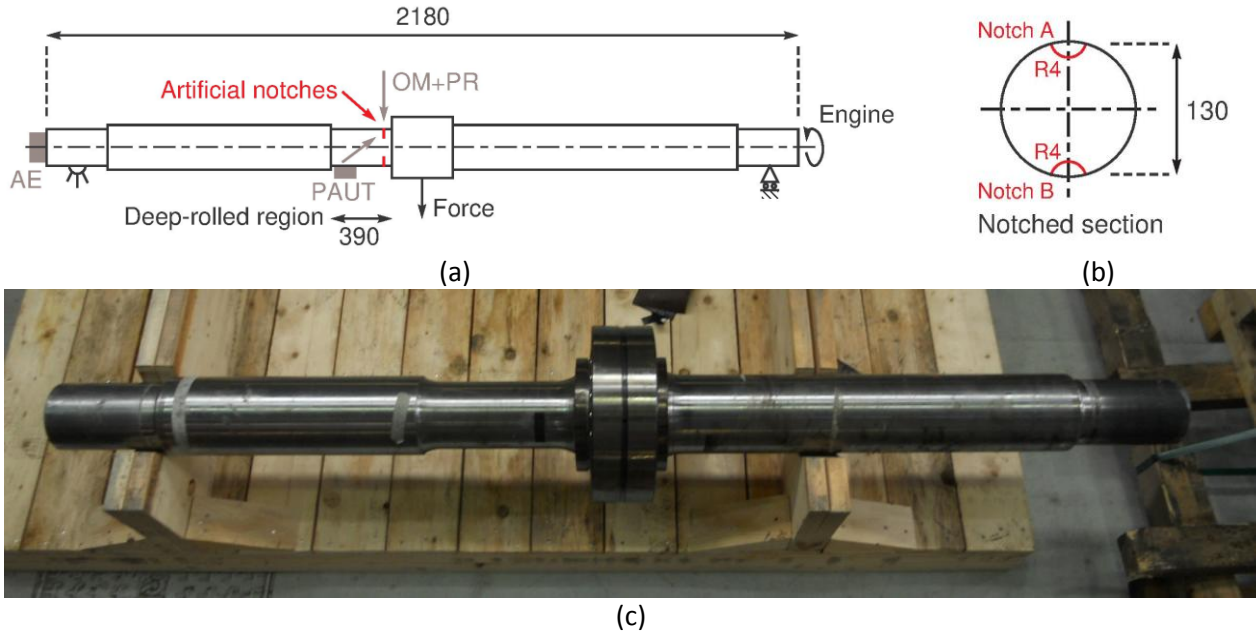

2180
Artificial notches
OM+PR
Engine
AE
PAUT
Deep-rolled region
390
Force
(a)
Notch A
R4
130
R4
Notch B
Notched section
(b)
(c)

Fig. 2

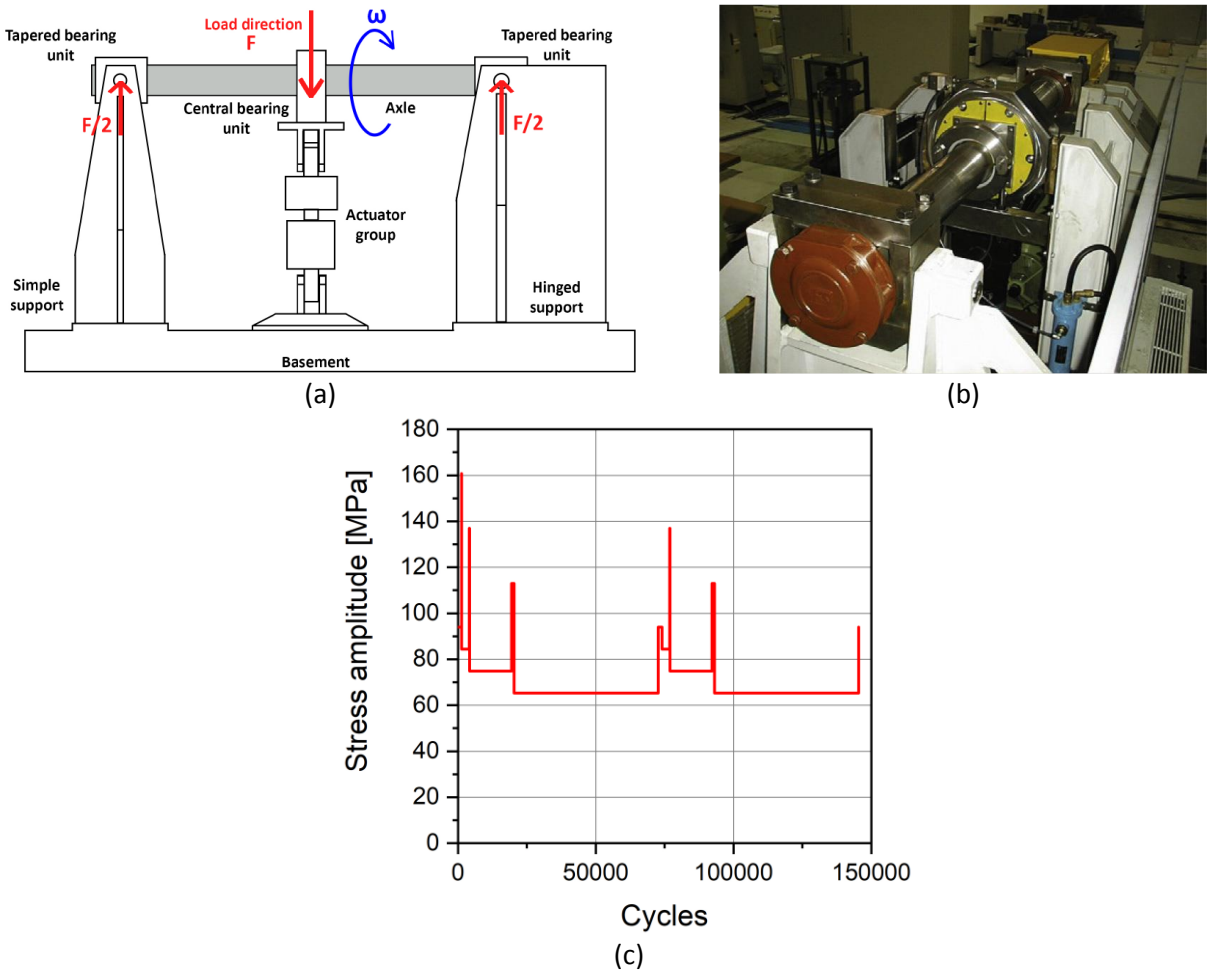
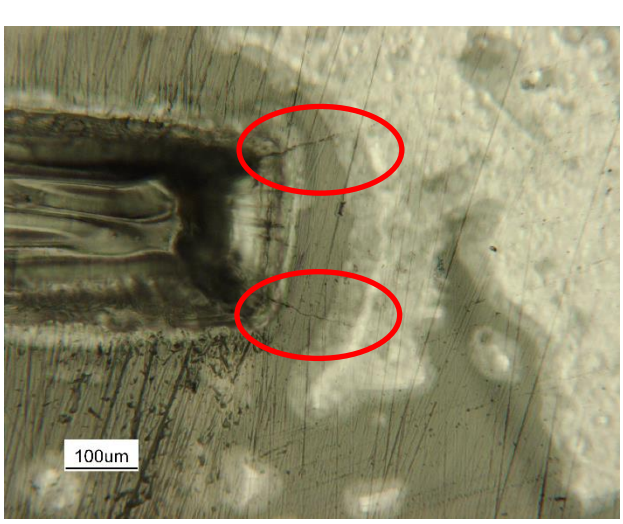
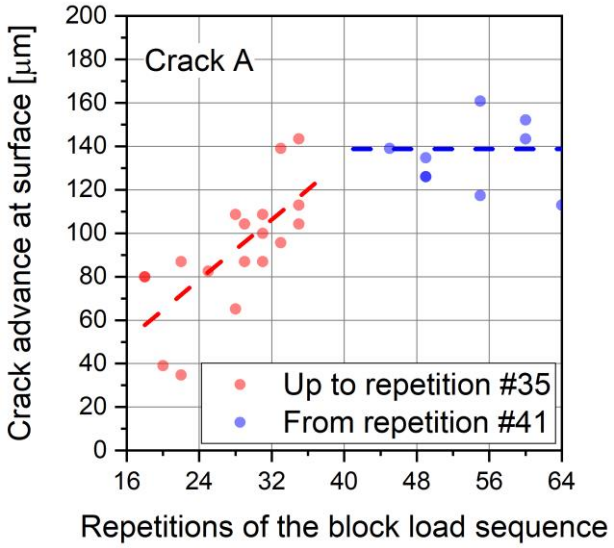

Tapered bearing unit
Load direction
F
ω
Tapered bearing unit
F/2
Central bearing unit
Axle
F/2
Actuator group
Simple support
Hinged support
Basement
(a)
(b)
Stress amplitude [MPa]
180
160
140
120
100
80
60
40
20
0
0
50000
100000
150000
Cycles
(c)

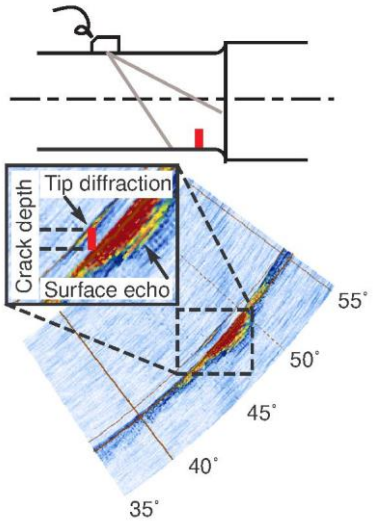
Fig. 3


100um

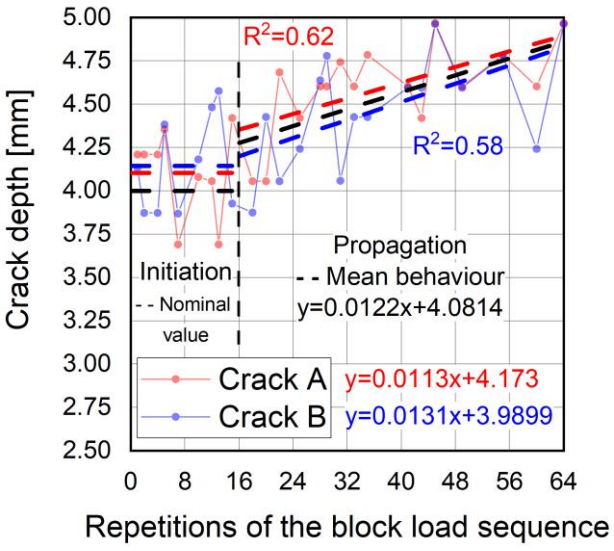
(a)


Crack A
Crack advance at surface [μm]
Repetitions of the block load sequence
Up to repetition #35
From repetition #41

(b)


Crack depth
Tip diffraction
Surface echo
55°
50°
45°
40°
35°

(c)


R²=0.62
R²=0.58
Crack depth [mm]
Initiation
- - Nominal value
Propagation
- - Mean behaviour
y=0.0122x+4.0814
Crack A y=0.0113x+4.173
Crack B y=0.0131x+3.9899
Repetitions of the block load sequence

(d)

Fig. 4

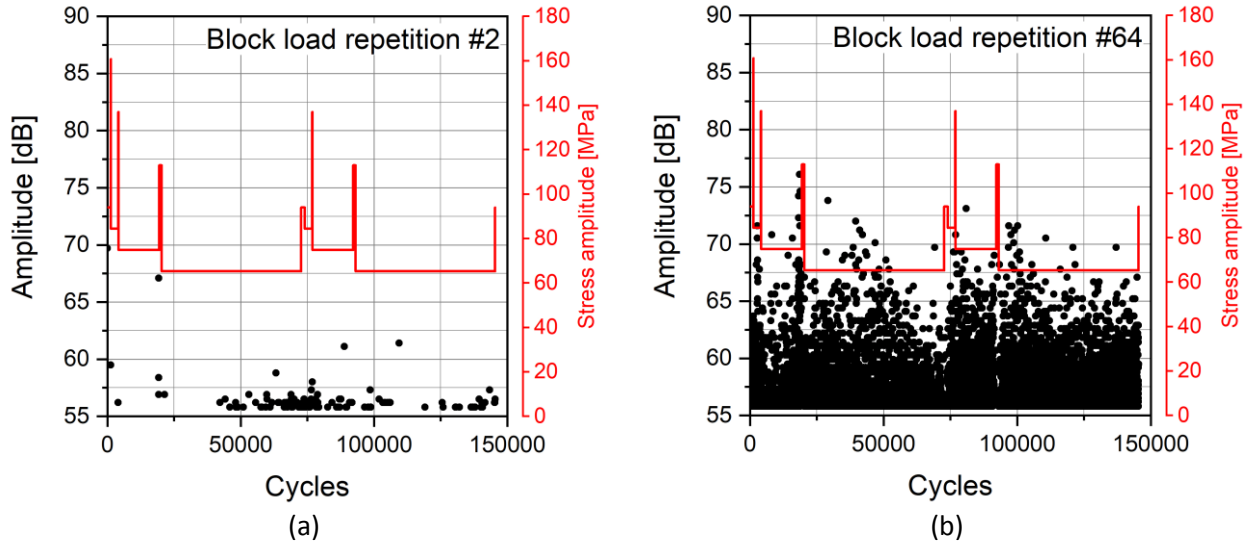

Block load repetition #2
Amplitude [dB]
Stress amplitude [MPa]
Cycles
(a)
Block load repetition #64
Amplitude [dB]
Stress amplitude [MPa]
Cycles
(b)

Fig. 5

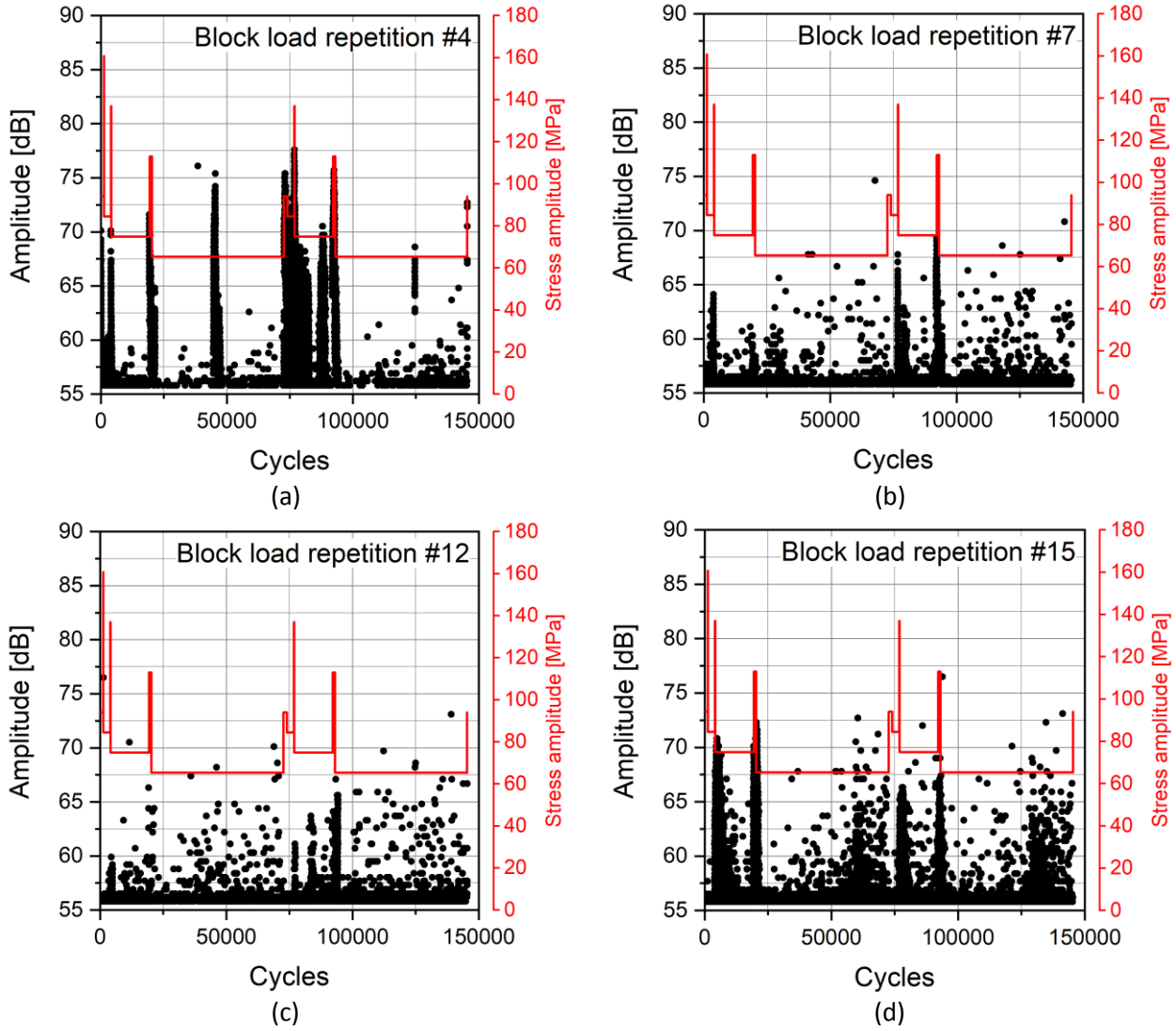
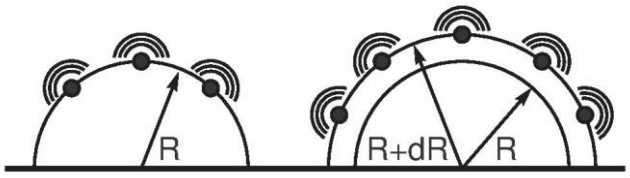
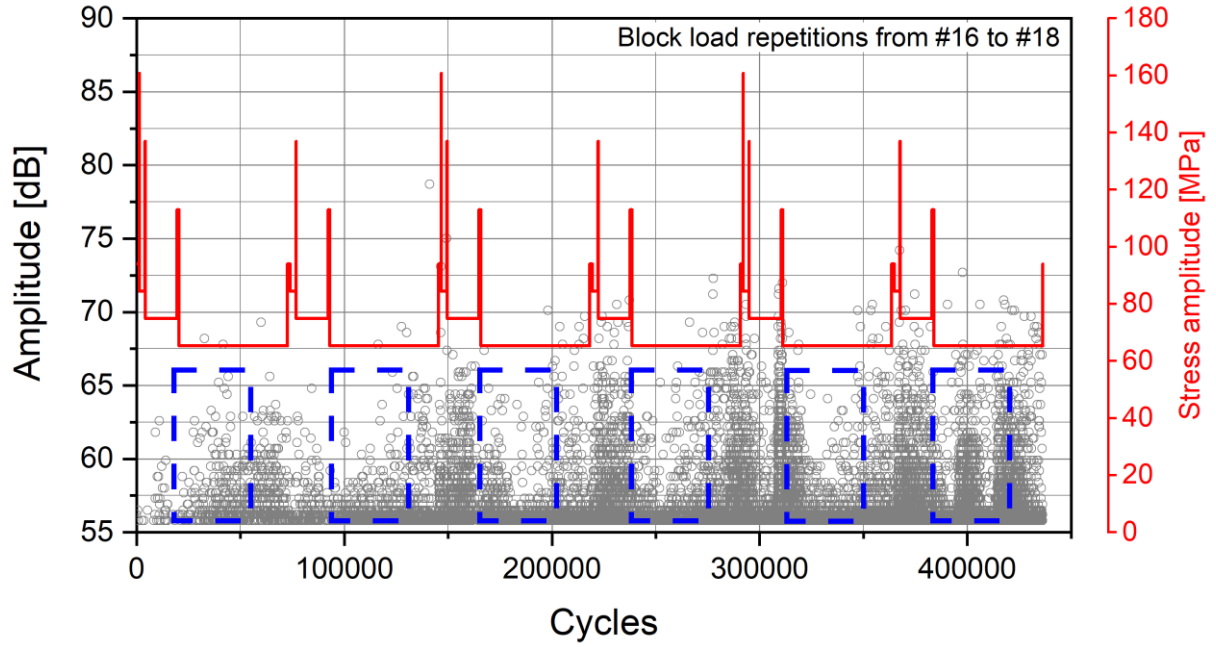

Block load repetition #4
Block load repetition #7
Block load repetition #12
Block load repetition #15
Amplitude [dB]
Stress amplitude [MPa]
Cycles
(a)
(b)
(c)
(d)

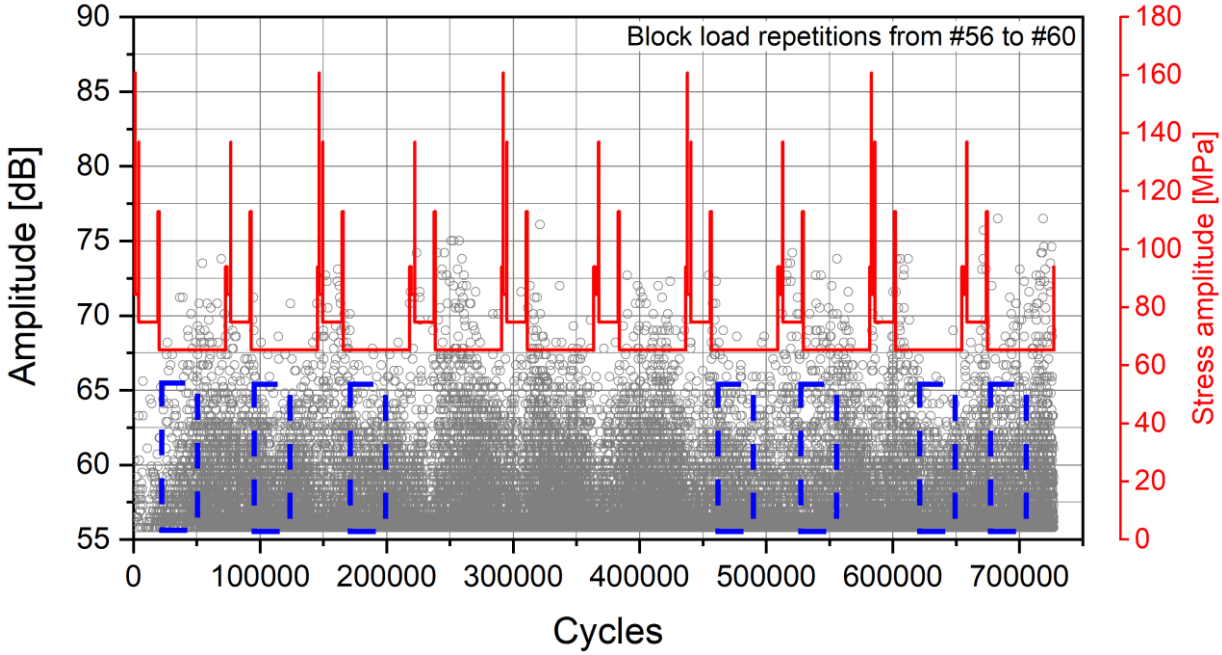
Fig. 6


R
R+dR
R

(a)


Block load repetitions from #16 to #18
Amplitude [dB]
Stress amplitude [MPa]
Cycles
90
85
80
75
70
65
60
55
180
160
140
120
100
80
60
40
20
0
0
100000
200000
300000
400000

(b)


Block load repetitions from #56 to #60
Amplitude [dB]
Stress amplitude [MPa]
Cycles
90
85
80
75
70
65
60
55
180
160
140
120
100
80
60
40
20
0
0
100000
200000
300000
400000
500000
600000
700000

(c)

Fig. 7

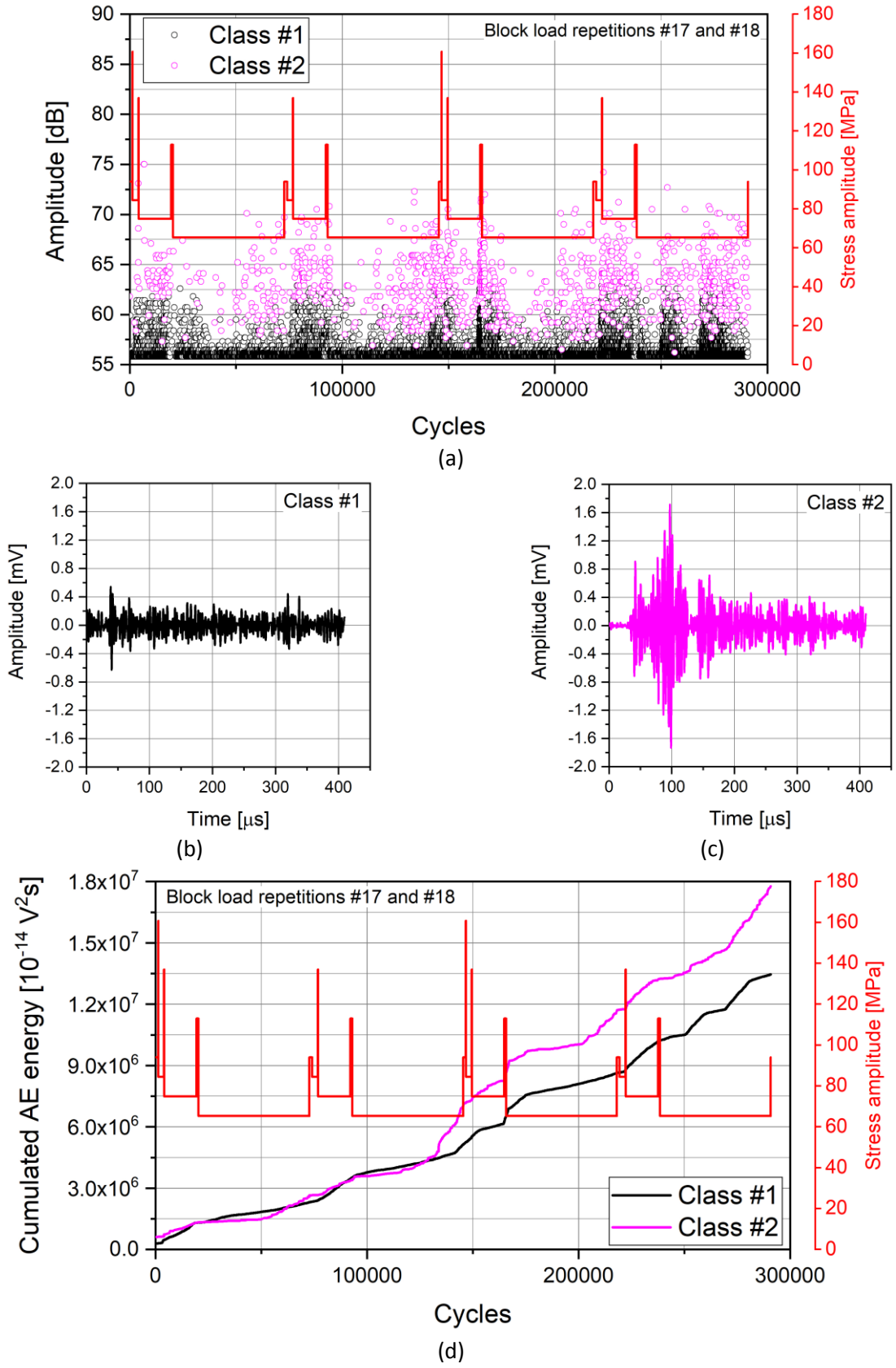
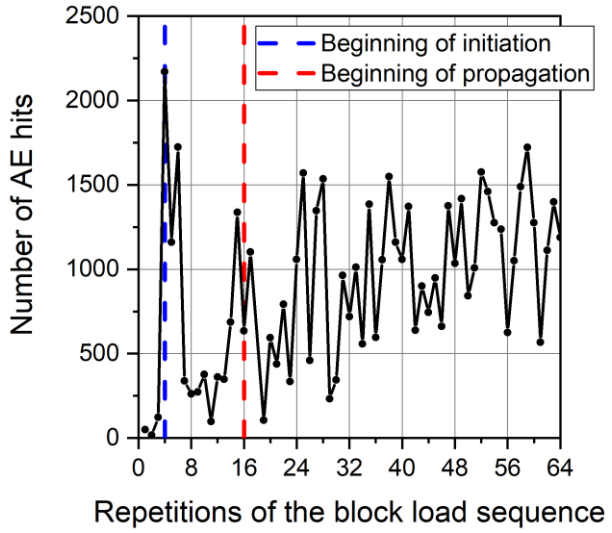
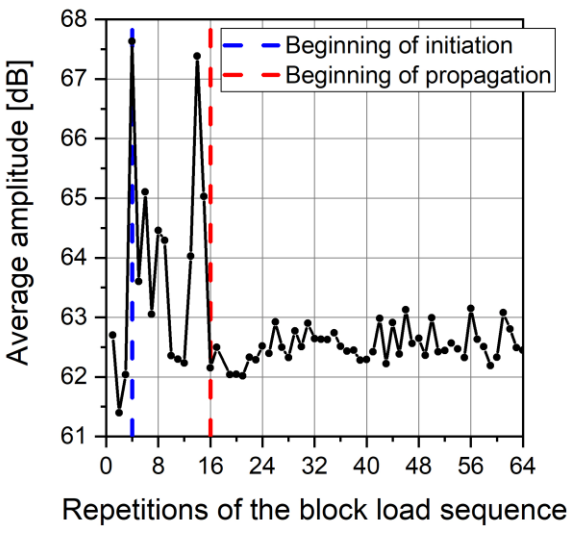

Class #1
Class #2
Block load repetitions #17 and #18
Amplitude [dB]
Stress amplitude [MPa]
Cycles
(a)
Class #1
Amplitude [mV]
Time [μs]
(b)
Class #2
Amplitude [mV]
Time [μs]
(c)
Block load repetitions #17 and #18
Cumulated AE energy [10^-14 V^2s]
Stress amplitude [MPa]
Class #1
Class #2
Cycles
(d)

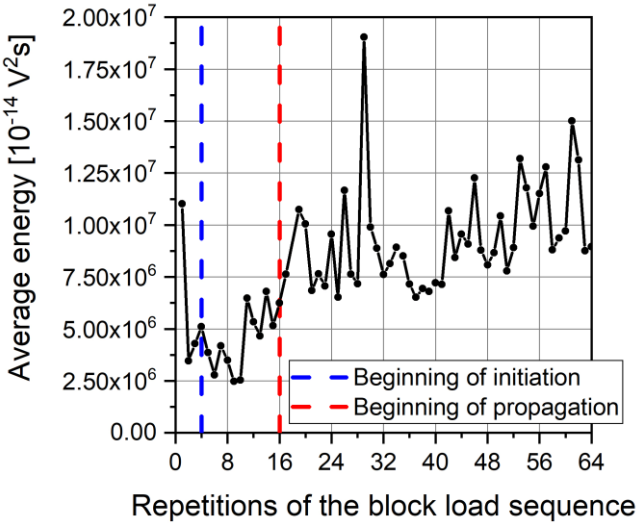
Fig. 8


Number of AE hits
2500
2000
1500
1000
500
0
0 8 16 24 32 40 48 56 64
Repetitions of the block load sequence
Beginning of initiation
Beginning of propagation

(a)

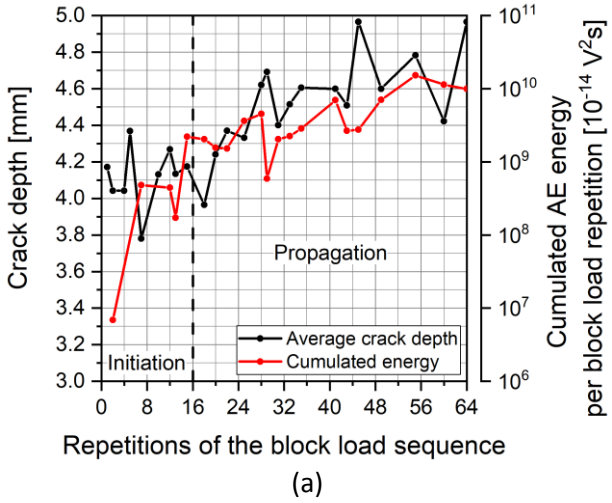

Average amplitude [dB]
68
67
66
65
64
63
62
61
0 8 16 24 32 40 48 56 64
Repetitions of the block load sequence
Beginning of initiation
Beginning of propagation

(b)


Average energy [10^{-14} V^{2}s]
2.00x10^7
1.75x10^7
1.50x10^7
1.25x10^7
1.00x10^7
7.50x10^6
5.00x10^6
2.50x10^6
0.00
0 8 16 24 32 40 48 56 64
Repetitions of the block load sequence
Beginning of initiation
Beginning of propagation

(c)

Fig. 9


Crack depth [mm]
Cumulated AE energy per block load repetition [10^-14 V^2s]
Repetitions of the block load sequence
Initiation
Propagation
Average crack depth
Cumulated energy
(a)

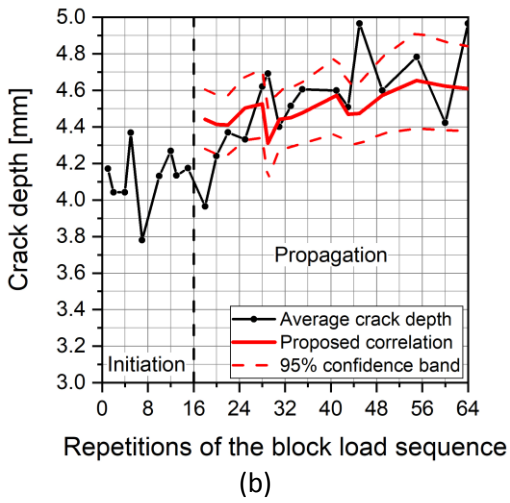

Crack depth [mm]
Repetitions of the block load sequence
Initiation
Propagation
Average crack depth
Proposed correlation
95% confidence band
(b)

Fig. 10