Numerical modelling of rolling contact fatigue damage initiation from non-metallic inclusions in bearing steel

Ravi, Gopalakrishnan; De Waele, Wim; Nikolic, Ksenija; Petrov, Roumen; Hertelé, Stijn

DOI
10.1016/j.triboint.2023.108290

Publication date
2023

Document Version
Final published version

Published in
Tribology International

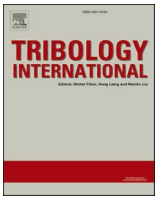
Citation (APA)
Ravi, G., De Waele, W., Nikolic, K., Petrov, R., & Hertelé, S. (2023). Numerical modelling of rolling contact fatigue damage initiation from non-metallic inclusions in bearing steel. *Tribology International*, *180*, Article 108290. https://doi.org/10.1016/j.triboint.2023.108290

Tribology International 180 (2023) 108290

Contents lists available at ScienceDirect

Tribology International

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



Numerical modelling of rolling contact fatigue damage initiation from non-metallic inclusions in bearing steel

Gopalakrishnan Ravi [a,b,*], Wim De Waele [a], Ksenija Nikolic [b,c], Roumen Petrov [c,d], Stijn Hertelé [a]

[a] Dept. EMSME, Soete Laboratory, Ghent University, Belgium
[b] SIM vzw, Technologiepark 48, 9052 Zwijnaarde, Belgium
[c] Dept. EMSME, Materials Science and Technology, Ghent University, Belgium
[d] Dept. MSE, Delft University of Technology, Delft, the Netherlands

ARTICLE INFO

Keywords:
Rolling contact fatigue
Crack initiation
Finite element simulation
Non-metallic inclusion
Debonding

ABSTRACT

Bearing failure is a cause of concern in a variety of machinery such as turbines, transmissions, drills, engines, etc. It is often associated with rolling contact fatigue (RCF) triggered from damage initiation at non-metallic inclusions (NMI's). Experimental evidence shows that damage initiation lifetime is highly sensitive to the NMI characteristics and its bonding with the steel matrix. This study numerically investigates the role of NMI features and its bonding with the steel matrix on damage initiation lifetime. NMI characteristics modelled in this study are derived from an experimental investigation of a failed bearing. Simulation results highlight a near to instantaneous debonding at the matrix-inclusion interface followed by accelerated crack initiation. The critical depth for damage initiation shifts towards the surface with the increase in friction coefficient between roller and raceway. The simulations also reveal that larger inclusions show earlier damage initiation, indicating a size effect. The damage hotspots from the simulation results were compared with experimental findings and a hypothesis for crack initiation from a NMI is put forward.

1. Introduction

Roller element bearings are critical engineering components that permit rotary motion whilst transmitting loads. They often fail due to a process called rolling contact fatigue where the contact parts are subjected to cyclic rolling or sliding contact. This can lead to either surface or subsurface initiated failure. When the bearing contacts are polished, lubricated and the bearing is installed properly, subsurface initiated damage tends to become dominant [1–3]. It has been experimentally verified that damage develops at the vicinity of subsurface defects, namely non-metallic inclusions (NMI's) which occur as an inevitable by-product of the bearing steel manufacturing process [4–6]. The most common NMI's are alumina (Al_2O_3), manganese sulphide (MnS) and other mixed type inclusions [5]. Although the advancements in steel-making processes led to cleaner steels [7], their presence remains unavoidable [6]. Rolling contact fatigue (RCF) driven subsurface damage estimation is a key challenge for the bearing community.

Despite numerous experimental investigations on the role of NMI's in bearing fatigue failure [4,5,8,9], only few researchers have attempted to capture the same from a numerical perspective [2,10–16]. Numerical modelling has its challenges in replicating the underlying physics behind RCF damage initiation from NMI's. It requires assumptions on failure mechanisms, inclusion modelling complexities such as accurate representation of their physical features. Nonetheless, acknowledging its approximate nature, numerical modelling of non-metallic inclusions in bearing steels offers a systematic way to investigate how the complex stress states around inclusions (and the influence of inclusion characteristics on these) link to fatigue failure initiation.

Historically, there have been attempts to model the RCF crack initiation problem either numerically or analytically based on contact stress or strain analysis [1]. Preliminary work was established by Jiang and Sehitoglu [17] and Ringsberg [18]. They used finite element (FE) analyses and strain-life based multiaxial fatigue crack initiation models accompanied by critical plane analysis concepts to predict a procedure for crack initiation under RCF loading. Liu et al. [19] extended this unified multiaxial fatigue damage model to railway wheel contact fatigue and anisotropic materials. Early works from Melander et al. [20] described the propensity of debonded inclusions having larger effect on early crack initiation via crack-tip displacements and total energy release rates. The majority of numerical models were built by applying

* Corresponding author at: Dept. EMSME, Soete Laboratory, Ghent University, Belgium.
E-mail address: gopalakrishnan.ravi@ugent.be (G. Ravi).

https://doi.org/10.1016/j.triboint.2023.108290
Received 11 December 2022; Received in revised form 6 January 2023; Accepted 22 January 2023
Available online 25 January 2023

continuum mechanics, cyclic contact loading and a multi-axial critical plane criterion for fatigue damage initiation [12,13]. Their main drawback is that they assumed the material to be homogenous and initially defect free without any imperfections such as non-metallic inclusions, which play an important role in the mechanics of the steel.

In the last decade, some researchers have attempted to address this issue by incorporating the role of material imperfections in their models. Recent numerical works that model RCF subsurface damage initiation from defects are discussed in the following. Sadeghi and Moghaddam et al. [2] used the finite element method along with a microstructure based 2D damage model to investigate the effects of inclusion size, depth and stiffness on rolling contact fatigue lifetime. Stiffer inclusions and inclusions present at the critical depth of maximum shear stress reversal were found more detrimental to bearing lifetime. However, their model was limited to fully bonded circular inclusions only. Another micro-mechanical study by Cerullo et al. [10] used a 2D plane strain FE model and continuum based linear elastic mechanics to study the fatigue damage initiation from NMI's using the Dang Van multiaxial fatigue criterion. Their study showed that maximum damage always occurred in the matrix for the cases examined, more specifically at the matrix-inclusion interface. Their study was limited to alumina (Al_2O_3) and titanium nitride (TiN) inclusions. Both studies did not consider the effects of other important aspects such as other types of inclusions (MnS), inclusion shape effects (e.g. MnS being potentially highly elongated), inclusion orientation and the effect of debonding between the inclusion and matrix on RCF initiation lifetime.

Recently, Mahdavi et al. [14] used a 3D micro-mechanical model to study the micro-scale cyclic plasticity at the vicinity of isolated subsurface inclusions such as Al_2O_3 and MnS to understand its impact on fatigue crack initiation. They studied the role of macro-residual stresses and the effect of steel matrix yield limit on the accumulated plastic strain for hard and soft inclusions. However, inclusion debonding and size effects were excluded. Work by Al-Tameemi et al. [11] tried to address missing links by considering the role of elongated inclusions such as MnS and introducing debonding at the inclusion-matrix interface to provide an insight into the stress state that could give rise to fatigue crack initiation. However, they did not go forward to establish the link between stress state around the inclusion to RCF lifetime estimation. Also, it did not consider other inclusion types such as alumina nor inclusion size effects. Schäfer et al. [15] thoroughly investigated the fatigue crack initiation from NMI's in steel including the effect of debonding via introducing different possible inclusion–matrix interface configurations by means of micromechanical simulations. However, this was done for a uniaxial loading case whereas RCF modelling involves non-proportional multiaxial stress states.

Role of inclusion-matrix debonding on early crack initiation is a growing area of interest in RCF [5,11,21]. This research complements the recent experimental findings by Vegter et al. [5]. They applied a tensile load to a bearing steel sample with a bonded inclusion, during which a gap was observed in-situ between the matrix-inclusion interface. This supported their hypothesis of crack initiation from an NMI. Another characterization study by Bruce et al. [22] confirmed this finding. Starting from an undamaged inclusion, separation between the interfaces and internal cracking occurs as an initial damage during the loading process.

This work contributes to the state-of-the-art by extending a computational methodology developed by the authors [23,24] to predict the RCF damage initiation lifetime triggered by the presence of NMI's. The proposed framework addresses abovementioned research gaps, i.e., the role of (a) inclusion characteristics such as size, type, shape, orientation, and (b) debonding on RCF initiation lifetime.

In addition to this, the role of inclusion depth below the surface and the influence of surface traction between roller and raceway on mechanical damage development is also investigated. Critical lifetime reduction factors are studied and a hypothesis of near to instantaneous debonding at the matrix-inclusion interface (henceforth simply referred to as "interface") leading to early crack initiation is proposed.

2. Material characterization study for a failed bearing

2.1. Characterization of NMI features

To investigate the damage evolution around inclusions, it is important to provide a realistic definition of inclusion features as input to the model. A material characterization study of a failed bearing acts as a starting point to this end. It was removed from a gearbox following the formation of an axial crack. Its loading history is unknown. The studied bearing is through-hardened and its chemical composition corresponds with a typical AISI 52100 steel grade [25], as shown in Table 1. After the final thermal treatment, which includes oil quenching and low-temperature tempering (~160 °C), the microstructure consists of spherical secondary cementite type carbides formed during spheroidization annealing, which are embedded in a matrix of tempered martensite, retained austenite, and one or more of the η, ε and θ phases, as presented in Fig. 1 [26].

For this study, the samples were cut along circumferential and radial planes, and further prepared for metallographic observation by a series of grinding and polishing steps. In the first step, the samples for further metallographic preparation were cut from the zone of interest of the bearing with a Struers type 50A15 cut-off wheel on a Discotom-5 cutting machine. The cutting was done with a very low cutting speed and under constant water cooling in order to prevent any heating of the sample that may cause phase transformation or microstructural changes. In the second step, the sample was embedded in a conductive polymer and prepared for microstructural observation following the standard metallographic sequence of grinding and polishing steps. These steps contain automatic mechanical grinding with SiC grinding papers P180, 220, 400, 600, 800, and 1200, followed by polishing with diamond paste with particle size of 3 μm and 1 μm and a final polishing step with 35 nm colloidal silica. During each grinding and polishing step the parameters (speed, pressure, cooling, lubrication) were strictly controlled in order to avoid the appearance of any artifacts (plastic deformation, debonding of the inclusions, etc). Finally, the samples were etched with nital 2% (2 ml nitric acid and 98 ml ethanol) to inspect the microstructure for the presence of damage. SEM images were acquired on an FEI Quanta 450-FEG-SEM (FEI Company, Hillsboro, OR, USA) operating at 15 kV, with a final beam aperture of 50 μm and spot size 5 that corresponds to a probe current of 2.3 nA in a secondary electron (SE) imaging mode. The chemical composition of the inclusions was determined using the energy dispersive X-ray spectroscopy (EDS).

The morphologies of the detected NMI's were thoroughly studied and their features such as type, size, shape, orientation with respect to the raceway and their debonding with the matrix were documented down to a depth of 1 mm below the raceway surface as shown in Fig. 2, where ORD refers to the over rolling direction of the roller. This depth range coincides with the zone of high rolling contact stresses. All of 209 examined inclusions were observed using scanning electron microscopy (SEM) and characterized as bonded or debonded in accordance to their coherence with the bulk material.

The majority (77%) of the inclusions were identified as pure manganese sulphide (MnS). The remaining 23% of inclusions were of a mixed type (MnS + Alumina (Al_2O_3)). Henceforward, MnS will be considered as the inclusion type of interest for the numerical investigation. This is in line with other experimental findings that support the detrimental role of MnS in RCF [8,9]. The inclusion database gathered from this study will be investigated to understand the inclusion characteristics. Based on the inclusion database, a parametric study using the numerical model will be constructed to reveal trends in the damage initiation lifetime driven by rolling contact fatigue.

Table 1
Chemical composition of through hardened bearing: AISI 52100.

(wt%)	C	Mn	P	S	Si	Ni	Cr	Cu	Mo
	0.95–1.05	0.25–0.45	0.025	0.015	0.15–0.35	0.25	1.35–1.6	0.3	0.1

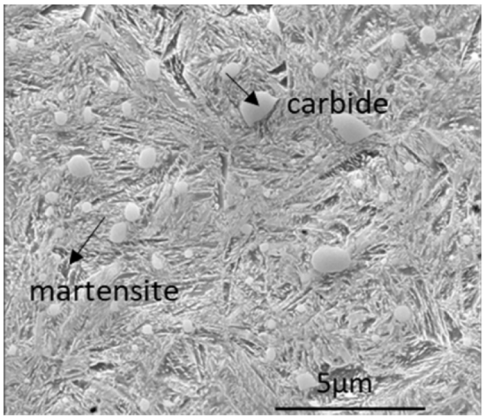


Fig. 1. Typical microstructure of the investigated AISI 52100 bearing steel consisting of tempered martensite and carbides.

2.2. Statistical results from inclusion database

Inclusions data collected from the characterization study were subjected to statistical correlation tests to identify the relationship between different variables. This was done using Pearson analysis, and their respective p value and correlation coefficient R for each combination are shown in Fig. 3. The results from the analysis show very low correlation between the variables. The following conclusions can be made:

- The majority (85%) of the inclusions had an aspect ratio of less than 3.
- The majority (94%) of the inclusions showed a major axis length of less than 10 µm.
- Inclusions were randomly oriented with their major axis between 0° and + /- 45° to the ORD.
- The majority (80%) of all inclusions observed across the depth were either partially or fully debonded.

3. Numerical modelling of damage initiation from NMI's

This section outlines the numerical modelling strategy implemented to estimate the RCF driven damage initiation and subsequent evolution from a non-metallic inclusion.

The overall approach is divided into two stages. The first is the mechanical RCF model that captures the stress field around an NMI and the corresponding damage state via a crack initiation framework based on continuum mechanics, cyclic contact loading and a multi-axial critical plane fatigue damage model [23,24]. The second stage links the damage state calculated in stage 1 to a crack initiation lifetime. The approach is summarized in Fig. 4.

3.1. Stage 1: RCF driven damage initiation model

The stage 1 numerical modelling approach is briefly explained hereunder; the reader is referred to [23,24] for an in-depth description of the model design and implementation. The adopted finite element analysis considers two levels: a global model and a submodel. The global model mimics the passage of a roller element on a bearing raceway by means of the theoretical (Hertzian) contact pressure it invokes (see Eq. (1)).

$$p(x) = p_{\max}\sqrt{1 - \frac{(x - x_c)^2}{b^2}} \tag{1}$$

Where x refers to the coordinate position where the pressure is evaluated, x_c denotes the center of the Hertzian pressure distribution, b represents the half contact width, and $p_{\max}$ is the magnitude of pressure at the center of the Hertzian pressure distribution. The rolling contact is mimicked by changing x_c linearly over time, such that the pressure distribution is displaced accordingly. Shear surface traction $t_x(x)$ in the

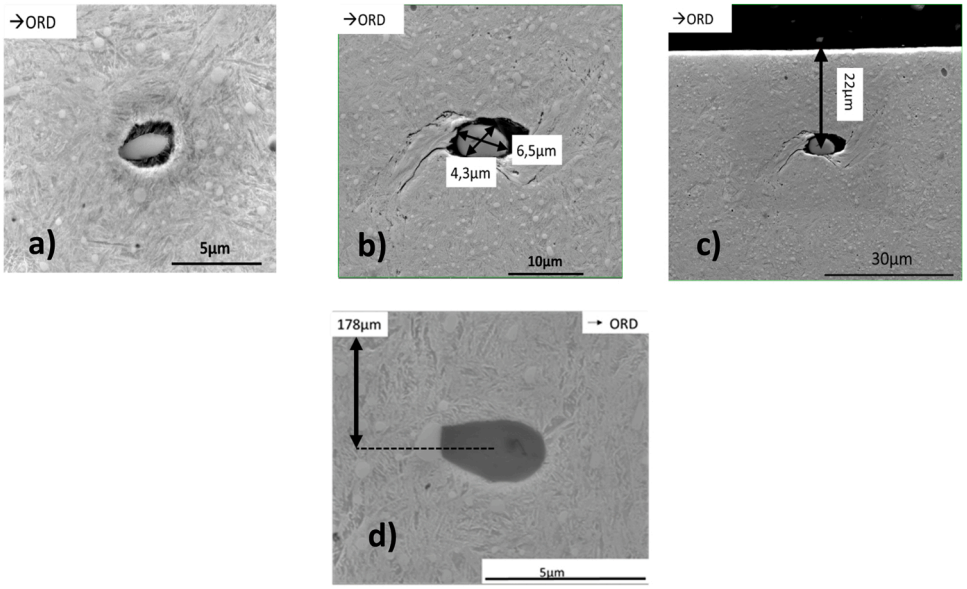


Fig. 2. (a) Example of non-metallic inclusion debonded from the matrix, (b) determination of size, and (c) depth below the surface. (d) Fully bonded inclusion, located 178 µm from the raceway surface.

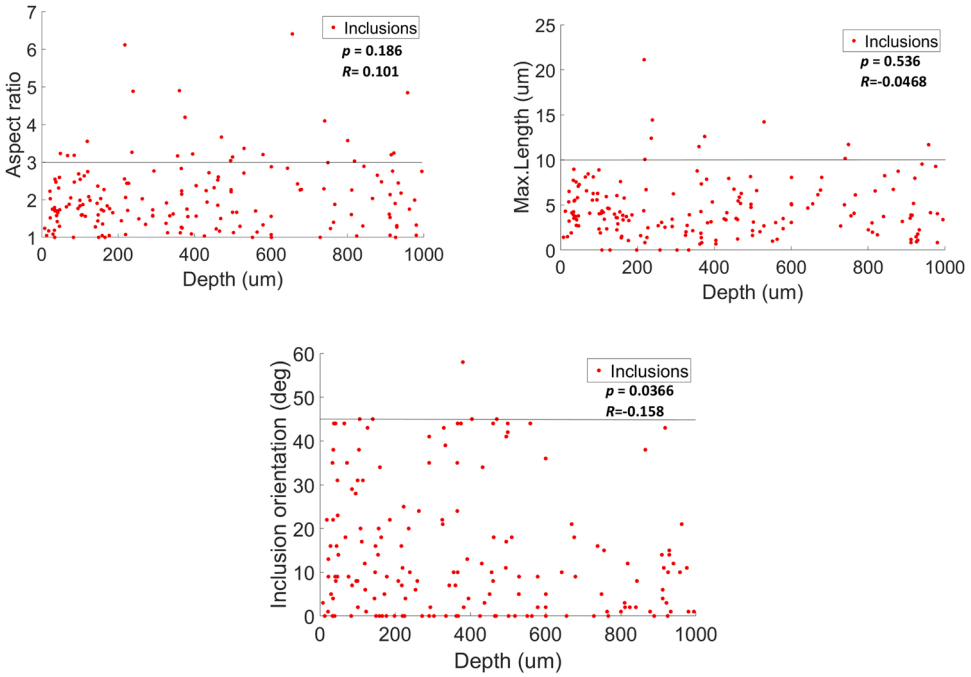


Fig. 3. Results from material characterization study - Inclusion database.

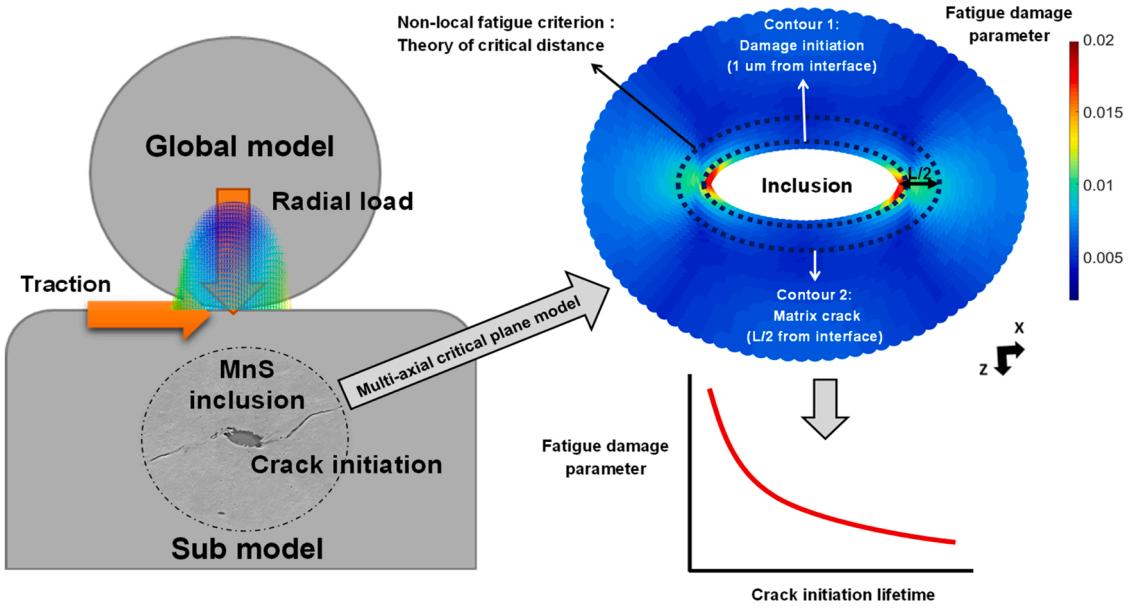


Fig. 4. Numerical RCF driven damage initiation lifetime model (Representation is not to scale).

rolling direction is taken proportional to the normal pressure according to the friction coefficient μ (Eq. (2)).

$$t_x(x) = \mu|p(x)| \tag{2}$$

The depth is defined in normalized terms against half-contact width b as per a Hertzian line contact [12] (See Eqs.(3) – (6)). Here, F refers to normal contact force, R and E^* denotes the equivalent radius and Young's modulus that includes the parameters E_1, R_1 and υ_1 and E_2, R_2 and υ_2 which represent the Young's moduli, curvature radii and Poisson's ratios of both contacting cylinders. In this study, the specifications of a cylindrical bearing of type NU 232 ECML [27] were used to calculate R. 210 GPa and 0.3 were taken for both E and ν values, corresponding to a steel-to-steel contact.

$$b = 2\sqrt{\frac{FR}{\pi E^*}} \tag{3}$$

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} \tag{4}$$

$$\frac{1}{E^*} = \frac{1-\upsilon_1^2}{E_1} + \frac{1-\upsilon_2^2}{E_2} \tag{5}$$

$$p_{\max} = \sqrt{\frac{FE^*}{\pi R^*}} \tag{6}$$

Using the time dependent stress tensors of this global model as boundary conditions, an NMI is subsequently introduced at the level of a

submodel to have a realistic estimation of the time dependent stress state around the inclusion.

In this study, both fully bonded and debonded inclusions were studied. Debonding is introduced at the inclusion-matrix interface by allowing separation between the contact surfaces. This is implemented by defining a frictionless hard contact without penetration at the interface [11,15]. This is a simplification of reality, as there will be some amount of interlocking or chemical forces present between the interfaces. As these are difficult to quantify and subject of ongoing research [5], the adopted contact definition is deemed to be a sound starting point to understand the influence of interface separation on rolling contact fatigue damage. For a partially bonded inclusion, the lower half of the interface was modelled as fully connected by means of tie constraints and the other half as debonded.

The crack initiation from an NMI in RCF is hypothesized to occur due to cyclic shear strain accumulation at the interface. This was clearly documented in the experimental studies of Grabulov et al. [4] and Vegter et al. [28]. The overrolling generates local interactions of normal and shear stresses at the interface, which over time lead to damage accumulation and eventually to crack initiation at the interface. In order to capture this damage mechanism, the theory of critical distances (TCD) is coupled with a shear based Fatemi and Socie (F-S) damage model which uses normal stress and shear strain range as its primary and secondary channel to calculate the material damage. This critical plane model takes cyclic, multiaxial, non-proportional loading behaviour into account, which makes it suitable to RCF based applications. Recall that the F-S damage model has been adopted in [13,29,30] for similar purposes.

The stresses extracted just outside the interface (see contour 1 in Fig. 4) are used to determine the damage initiation from an NMI. This was taken as 1 μm from interface as there will be stress singularities leading to higher damage values if extracted right at the interface. Whereas, the stresses extracted at a certain distance from the interface represent the transition into a matrix crack (see contour 2 in Fig. 4). A point based TCD (theory of critical distance) method is used in this work, similar to a recent RCF railway application study [31]. The stresses are extracted at a distance of $L/2$ from the inclusion, where L refers to a material characteristic length that is calculated as follows:

$$L = \frac{1}{\pi}\left(\frac{\Delta K_{th}}{\Delta\sigma_0}\right)^2 \tag{7}$$

With $\Delta K_{th}$the threshold range of the stress intensity factor and $\Delta\sigma_0$ the plain fatigue limit. The material parameters to calculate this TCD length for the same material are taken as follows [32]: ΔK_{th}= 5 MPa $m^{1/2}$ and $\Delta\sigma_0$ = 850 MPa for a subsurface crack. Using these values, the highest F-S damage reached at a critical distance of 5 μm around the circumference of the interface is used to predict potential size effects on fatigue crack initiation and also this captures the progression from debonding to crack formation in the matrix (Fig. 4).

3.2. Stage 2: Linking damage to initiation lifetime model

In this subsection, a simple and comprehensive procedure to estimate crack initiation lifetime for subsurface NMI's is presented.

Commonly in RCF applications, the material damage is linked to initiation lifetime by means of the Basquin-Coffin-Manson equation [13] (see Eq. (8)).

$$\left(1 + k.\frac{\sigma_{n,max}}{\sigma_y}\right).\frac{\Delta\gamma_{\max}}{2} = \frac{\tau_f'}{G}(2N_f)^{b_\gamma} + \gamma_f'(2N_f)^{c_\gamma} \tag{8}$$

In this equation, k is a material constant arbitrarily taken as unity, σ_y represents the yield strength, $\Delta\gamma_{\max}$is the maximum shear strain range on the critical plane and $\sigma_{n,\max}$ represents the maximum stress perpendicular to this material plane. N_f is the number of cycles to failure for the given F-S material damage. G is the shear modulus, τ_f' is the shear fatigue strength coefficient, γ_f' is the shear fatigue ductility coefficient, b_γ and c_γ are the shear fatigue strength exponent and shear fatigue ductility exponent respectively. The left side of Eq. (8) represents the F-S fatigue damage. This is calculated via a critical plane search algorithm developed for subsurface inclusions, as documented in our previous studies [23,24].

One challenge of this approach is that it demands several material parameters which can only be obtained from a time consuming and elaborate experimental multi-axial fatigue test program. As a result fatigue curves for multi-axial loading are not as widely available as compared to uniaxial cases. Researchers have tried to address this issue by using material data available for common bearing steel or using approximations from uniaxial strain-life properties [13,33]. In this work, an alternative approach has been applied to come up with an approximate fatigue life estimation using simple and commonly available material properties.

The authors propose to link the F-S damage to a Roessle-Fatemi hardness method [34], which estimates fatigue initiation lifetime purely based on material hardness (Eq. (9)). This approach is a variation of the Basquin-Coffin-Manson equation based on commonly available material properties. Brinell hardness is used to calibrate the cyclic material parameters needed for the right hand side of Eq. (9). As these material coefficients were calculated via hardness measurements for the failed bearing used in this study, this method offers a pragmatic way to indirectly couple the effect of inelastic behavior on initiation lifetime [13].

$$\left(1 + k.\frac{\sigma_{n,\max}}{\sigma_y}\right).\frac{\Delta\gamma_{\max}}{2} = \left[A(2N_f)^{-0.09} + B(2N_f)^{-0.56}\right]\left[1 + kC(2N_f)^{-0.09}\right] \tag{9}$$

The feasibility of this approach has been evaluated for steels with Brinell hardness (HB) between 150 and 700 kgf/mm^2 [34,35]. The parameters A, B, and C can be written in terms of HB as follows (Eqs. (10–12)):

$$A = \frac{5.53(HB) + 293}{200000} \tag{10}$$

$$B = \frac{0.48(HB)^2 - 731(HB) + 286500}{200000} \tag{11}$$

$$C = \frac{1}{0.0022(HB) + 0.382} \tag{12}$$

For this study, a Vickers indentation was performed on a sample of the failed bearing traversing across the depth from raceway. This was carried out with a load of 20 kgf and dwell time of 10 s using a DURAMIN-40 AC3 tester. As the tested sample was taken from a through hardened bearing, the profile showed a relatively constant hardness across the depth. The mean hardness across the region of interest (1.5 cm) was found to be 701 HV, which can be translated into an equivalent Brinell hardness of 656 HB (kgf/mm^2) as per ISO 18265 [36]. This hardness value has been input to Eq. (9) to calculate damage initiation lifetime from F-S damage.

3.3. Model optimization: Time and spatial discretization

The rolling pass was discretized into 21 time frames to obtain a satisfactory balance between iteration time and numerical accuracy. The rolling pass length was taken to be from *-2b* to *2b*, where *b* refers to half contact width. This provides adequate stress histories with good accuracy as shown in previous work [23] and is also consistent with the inferences from [2,37].

Subsurface stresses scale linearly with applied contact pressures due to the linear elastic nature of the adopted model. Hence, the reference contact pressure p_{max} is fixed to 1 GPa and the obtained stress distributions are scaled with contact pressure.

Fig. 5 represents the mesh design used in this study to model different inclusion features. A biased mesh design inspired from Cerullo et al. [10] was found useful, where more interest is shown to the regions closer to the interface and the mesh coarsens towards the boundaries. A CPE4R plane strain quadrilateral type element was used in this study. Element size was taken to be 0.2 μm in the matrix region close to the interface to clearly capture the stress gradients.

After mesh design, the damage reached in the matrix was chosen as a representative output to check the mesh convergence. Referring to recent literature that involves modelling NMI's [2,10,15,38], it can be unanimously seen that damage reached in the matrix can be a tangible output parameter. In this study, as discussed earlier we use a non-local fatigue criterion based on TCD (Eq. 7) to calculate damage in the matrix. The obtained F-S damage values were used as a representative output in a mesh convergence study to find a satisfactory balance between mesh density and computational time. Fig. 5b shows the mesh convergence analysis for a fully debonded interface. This was done for an inclusion length of 15 μm, aspect ratio 5, depth of 0.5 b and for a contact pressure of 1 GPa. From the figure, it can be understood that the output converges when the number of elements crosses 10k with a reasonable computational time. This was chosen as an optimum for in this study.

4. Qualitative comparison to literature

In this section, the damage results predicted from the numerical model are benchmarked against similar studies reported in literature. As different researches use different methods and damage models, a direct quantitative comparison cannot be made. Hence the assessment is done qualitatively at two levels: stress raising factor and damage prediction.

Fig. 6 shows the fatigue damage prediction using our model, for a fully bonded hard inclusion (Al_2O_3), a fully bonded soft inclusion (MnS), and a fully debonded hard inclusion (Al_2O_3). The simulations were performed using a reference contact pressure of 1 GPa. Inclusions were circular, having a diameter of 10 μm. The inclusions mechanical properties were: $\nu = 0.25$, $E_{\text{(alumina)}} = 380$ GPa, and $E_{\text{(MnS)}} = 108$ GPa. The inclusion was positioned at a depth of 0.5 b below the contact surface with $b = 250$ μm. This depth corresponds to the region of the maximum subsurface shear stress reversal under rolling contact.

The damage trends predicted from this study (Fig. 6) are qualitatively in line with the trends in stress raising factor reported by Lai et al. [16] (Fig. 7). In both studies, a fully bonded hard inclusion introduces the least damage, a fully bonded soft inclusion generates relatively higher damage than a hard inclusion, and the damage increases further with the introduction of debonding between the inclusion-matrix interface.

Further, Fig. 8a represents the debonded regions observed around the interface for an alumina inclusion [4]. Fig. 8b shows the values of the F-S fatigue damage parameter predicted around the interface for a bonded circular Al_2O_3 inclusion. Fig. 6a & c show the damage evolution around a bonded and debonded alumina inclusion predicted in this study. This prediction is in line with the experimental finding from Fig. 8a confirming that maximum damage accumulation at the 45 deg. intercepts on the interface dominates the debonding for an alumina inclusion. Similar damage predictions were observed by Lai et al. [16] and Mahdavi et al. [29] in their simulations for alumina.

5. Design of virtual experiments

Representing an NMI involves the definition of multiple characteristics. This study focuses on effects of inclusion size, shape, orientation, depth, and debonding of its interfaces. The inclusion type is fixed to MnS based on the experimental evidence from the failed bearing discussed above. A full-factorial parametric study devised to capture these effects is shown in Table 2.

6. Results and discussion

This section highlights how the abovementioned variables affect the bearing damage initiation lifetime. By doing so, we try to interpret and understand how debonding and other inclusion characteristics can be an important influencing factor on bearing damage initiation.

6.1. Debonding increases fatigue damage

Fig. 9 summarizes the fatigue damage contours around a selection of different inclusion morphologies and for different bonding ratios. The results shown were obtained for simulations with a contact pressure of 1 GPa and zero friction. The inclusion length was equal to 10 μm, with varying aspect ratios (AR) of 1, 3 and 5. Orientation angles of 0 and 45 degrees with respect to the ORD were considered and all inclusions were positioned at a depth equal to 0.5 b. The contour plots confirm the influence of inclusion-matrix debonding on the F-S fatigue damage value near the interface. It is clear that both partially and fully debonded interfaces experience significantly higher damage than a perfectly bonded inclusion. This observation will be quantified below.

A perfectly bonded inclusion already acts as a damage raiser. Once debonding is introduced between the inclusion and the surrounding steel matrix, the separated interfaces behave as a crack which alters the stress field around them. As soon as debonding comes in to picture, the damage becomes localized to the debonded edges. The consequence of this increase in damage can be noticed by comparing the damage around perfectly bonded and debonded inclusions. A significant increase in fatigue damage is present which eventually affects the damage initiation lifetime. This can be observed for all cases and orientations varying from circular to elongated inclusions. Also, for a fixed inclusion length, the sharper the inclusion, the higher the damage at the interface.

This is in line with our experimental findings which indicate that the inclusion tips act as crack initiation hot spots for elongated inclusions and their side faces initiate cracks for globular type inclusions as

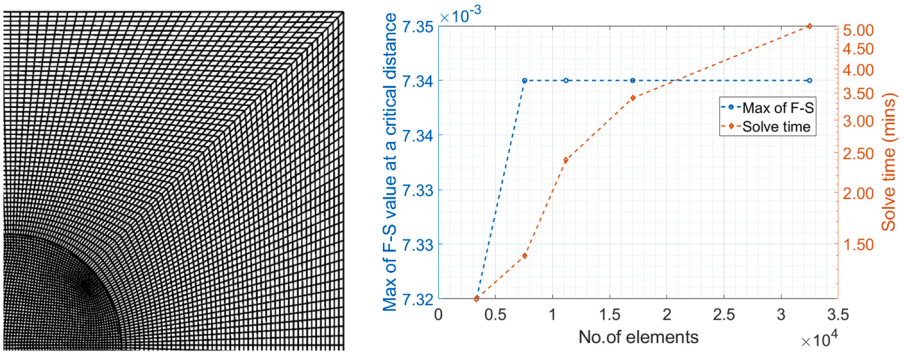


Fig. 5. (a) Mesh design, and (b) mesh convergence.

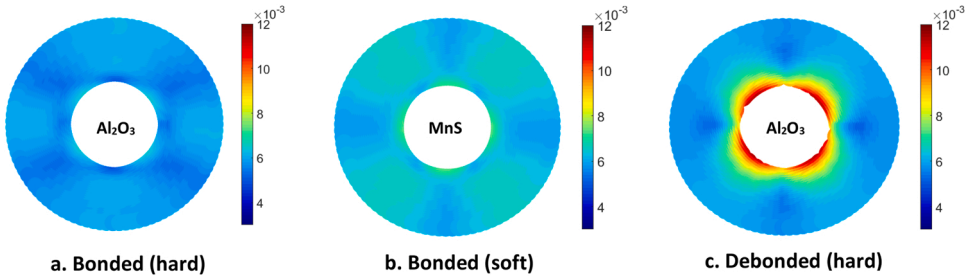


Fig. 6. Distribution of F-S fatigue damage around different types of imperfections and for different bonding conditions.

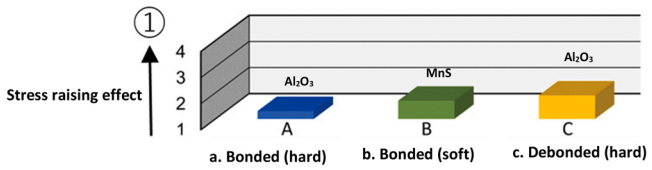


Fig. 7. Literature results - Stress raising effect for different types of NMI's and bonding conditions [16].

predicted by the simulations. Fig. 10a shows the damage initiation from the side faces for a globular inclusion and Fig. 10b shows how damage initiates from the crack tips for an elongated inclusion.

Fig. 11 shows the damage extracted at the contour 1 (see Fig. 4) for all inclusions numerically investigated in this work and classified based on different bonding ratios. This figure includes damage results from 405 simulations (i.e. 135 simulations each for fully bonded, partially and fully debonded cases), obtained for the case of zero friction and other inclusion parameters shown in Table 2.

The debonded inclusions have a higher median damage, hence cracks will generally initiate earlier than for perfectly bonded inclusions. Although not profound, slightly higher median damage is predicted for partially debonded inclusions as compared to fully debonded ones. This is visible in the case of AR 3 in Fig. 9 where slightly higher damage is predicted at the tips of a partially debonded inclusion compared to a fully debonded case. Contrary, the damage saturates to similar values in AR 5 for either case due to the sharp tips and thus high stress concentration. The depth range was limited from 0.25 *b* to 1 *b*, below which no significant RCF damage was predicted due to the reduced stresses at deeper levels.

The underlying mechanism behind the debonding phenomenon are explored in the next section.

6.2. Near to instantaneous debonding occurs around the interface

Our experimental study as well as literature show that inclusions tend to act as crack initiators [2,4,8]. This investigation shines light on how the separation of the interface between the inclusion and matrix takes place early and moreover, how it can accelerate the crack initiation process. Fig. 12 gives an idea on how a crack initiates from a NMI. It displays the damage development around a fully bonded inclusion where it can be seen that highest damage is reached at top and bottom poles of the inclusion. Consequently this damage build up quickly leads to increased and localized damage around the tips as shown for a debonded inclusion. By extracting the damage at the interface (contour 1) for a fully bonded inclusion and linking to lifetime, it is noticed that in a time span of around 10^4 – 10^5 cycles, damage builds up indicating damage initiation. In the context of high speed stage wind turbine gearboxes, this lifetime is a matter of hours for a roller element bearings present inside them. This means a near to instantaneous debonding occurs around the interface. Once this interfacial gap comes into the picture, this acts as an initial crack and accelerates the crack initiation process by raising the stress concentration in the matrix.

Table 2
Factors affecting damage initiation lifetime - simulation matrix.

Contact pressure / Half contact width *b*	1 GPa/ (250 μm)					
Inclusion size - maximum length (μm)	5		10		15	
Aspect ratio	1		3		5	
Depth (μm)	0.25 *b*	0.5 *b*	0.75 *b*	1*b*	2 *b*	
Orientation (deg.) w. r.t ORD	0		45		-45	
Debonding ratio	0% (bonded)		50% (partially bonded)		100% (debonded)	
Friction coefficient at surface for calculation of traction	0		0.15		0.3	

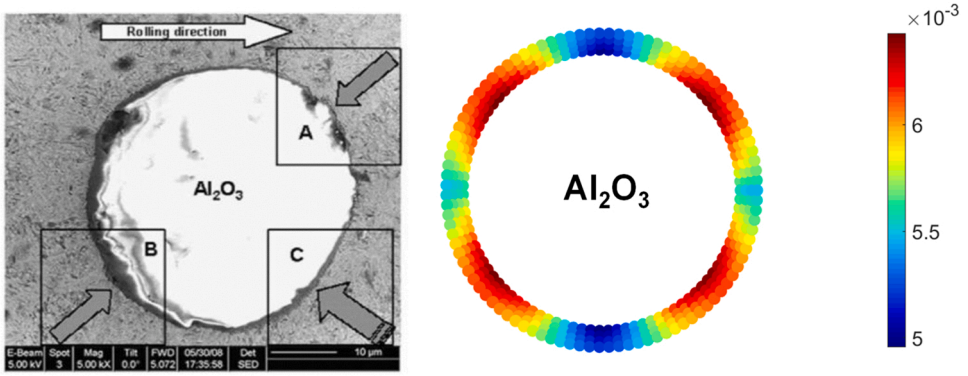


Fig. 8. Debonding and deformation regions at inclusion-matrix interface for alumina (a) Literature results - as observed by Grabulov [4] (b) Simulation results.

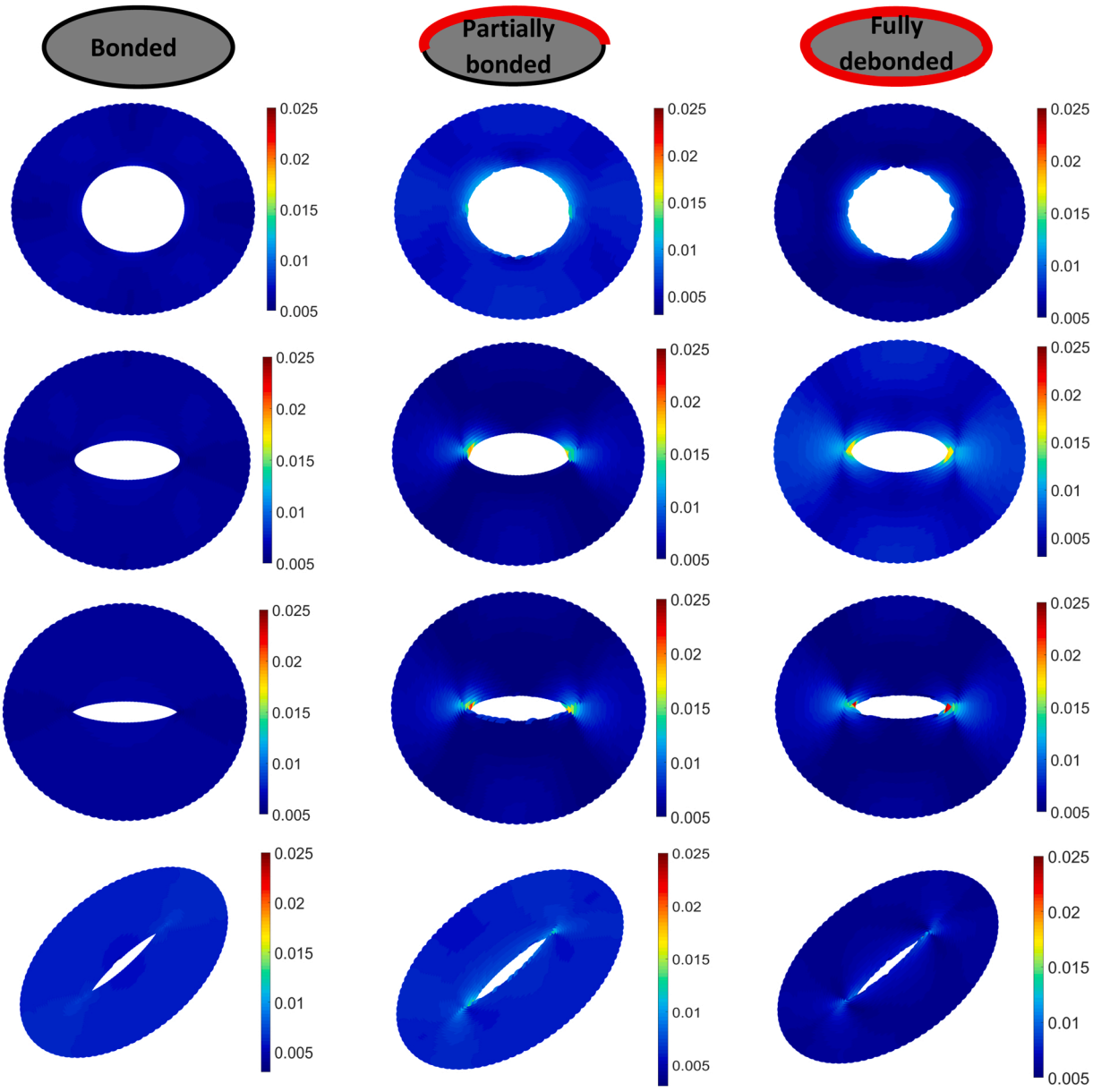


Fig. 9. Damage evolution around various NMIs for different bonding ratios with the surrounding matrix.

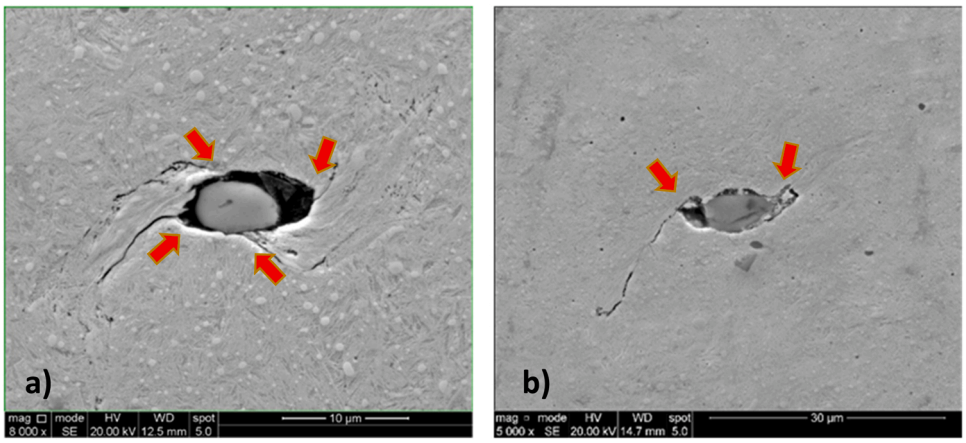


Fig. 10. Crack initiation from a MnS inclusion: (a) globular and (b) elongated.

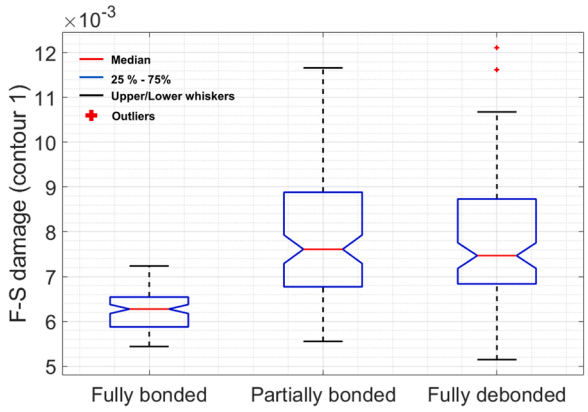


Fig. 11. Effect of debonding on F-S fatigue damage initiation.

From our simulation results, this section puts forward a crack initiation hypothesis. Starting with an inclusion which is bonded to the matrix, once the loading begins, damage starts to build up around the inclusion. As the global rolling contact stress field generates a normal (crack opening) stress and shear stress around the interface, this damage induces a local contact between the steel matrix and inclusion. The phenomenon over time causes a separation between the inclusion-matrix interface and creates a gap between them. The beginning of this gap at the inclusion-matrix interface can be the initiating point of the crack. A schematic representation of the mechanism is shown in Fig. 13.

The same has been experimentally observed by recent studies from Vegter et al. and Bruce et al. [5,22], where starting with a bonded inclusion, separation between the interfaces was observed during the early stages of the loading process. Hence, this hypothesis based on numerical evidence provides essential insight on how NMI's can be crack initiators. This initiated crack can induce microstructural phase transformation in later stages, such as formation of white etching matter by crack rubbing mechanism [28,39], which is beyond the scope of this paper.

This investigation suggests that debonding happens early during the loading process and subsequently leads to early damage initiation. Hence, from now on, only debonded inclusions will be considered for further analysis in this study.

6.3. Role of inclusion depth and surface traction

Fig. 14 demonstrates the damage initiation lifetime (contour 1) predicted at various subsurface depths for different characteristics of a fully debonded inclusion (bearing in mind that debonding is expected to occur very early on, as inferred from previous section). As an example, results are shown for a debonded elongated inclusion (AR = 3), for varying lengths and increasing friction coefficient between roller and raceway. The lifetime profile clearly follows the subsurface stress distribution across the depth. The F-S damage calculated from subsurface stresses and strains start to increase around 0.25 *b*, it reaches a peak at 0.5 *b* after which it gradually starts to drop from 0.75 *b* and eventually fades away after a depth of 1 *b*. Below 1 *b,* the damage becomes insignificant, hence the inclusions present below 1 *b* become less important. This trend is reflected in terms of initiation lifetime predicted across the depth, as expected the least lifetime is predicted at 0.5 *b* and 0.75 *b*, out of which the depth of 0.5 *b* becomes critical as this where the maximum shear stress reversal occurs.

Similarly, Fig. 15 (taken at contour 2) shows the lifetime needed to form a matrix crack for different friction coefficients. It can be seen that damage initiation occurs at an order of magnitude quicker than matrix crack formation. The role of surface traction on damage initiation was documented in our previous study [23]. Adding to that, this study shows how much it affects the damage initiation lifetime across the depth. As can be seen from Fig. 15, the friction coefficient has a high impact on the depth range closer to the surface i.e. from 0 to 0.5 *b*. Another important observation which supports the previous studies is that [9,23], increase in surface traction shifts the subsurface stresses closer to the raceway i.e. the critical depth moves from 0.5 *b* towards 0.25 *b* for the simulated cases. This trend can be clearly seen in 0.3 friction coefficient case, as we increase the friction coefficient from 0.15 to 0.3, inclusions close to the surface experience high stresses and fail early than the ones at 0.5 *b*.

From this analysis, it can be concluded that for a general scenario of normal operating conditions, the earliest crack initiation happens at a

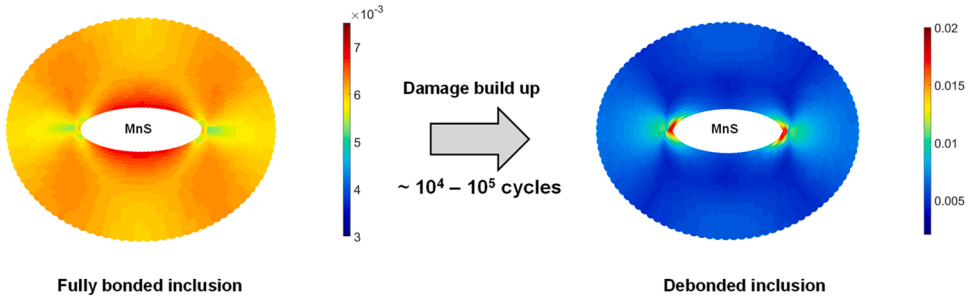


Fig. 12. Crack formation from a NMI via debonding.

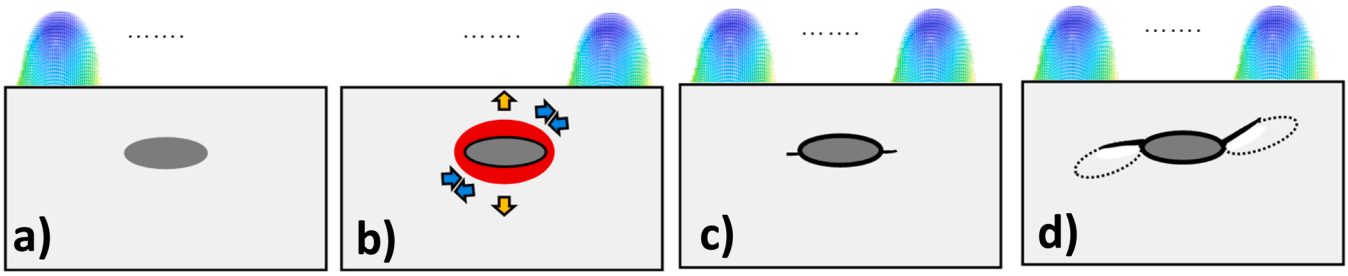


Fig. 13. Schematic representation of crack initiation from an NMI a. Bonded inclusion b. Local interaction due to normal and shear stresses creates debonded interfaces c. Debonding acts an initial crack d. Crack formation, can lead to butterfly and WEC formation in later stages.

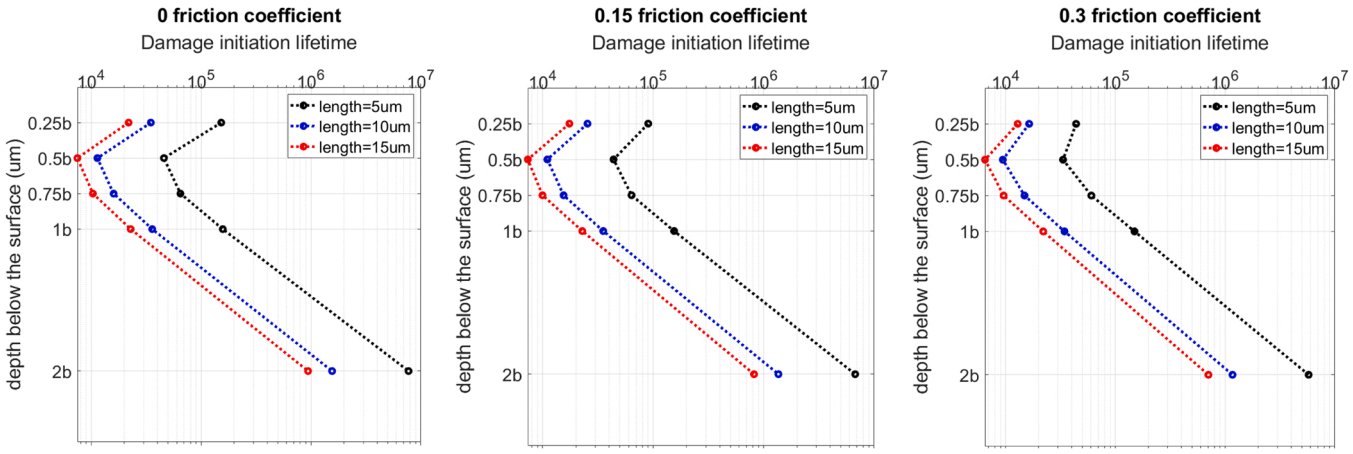


Fig. 14. Role of critical depth and surface traction on damage initiation lifetime.

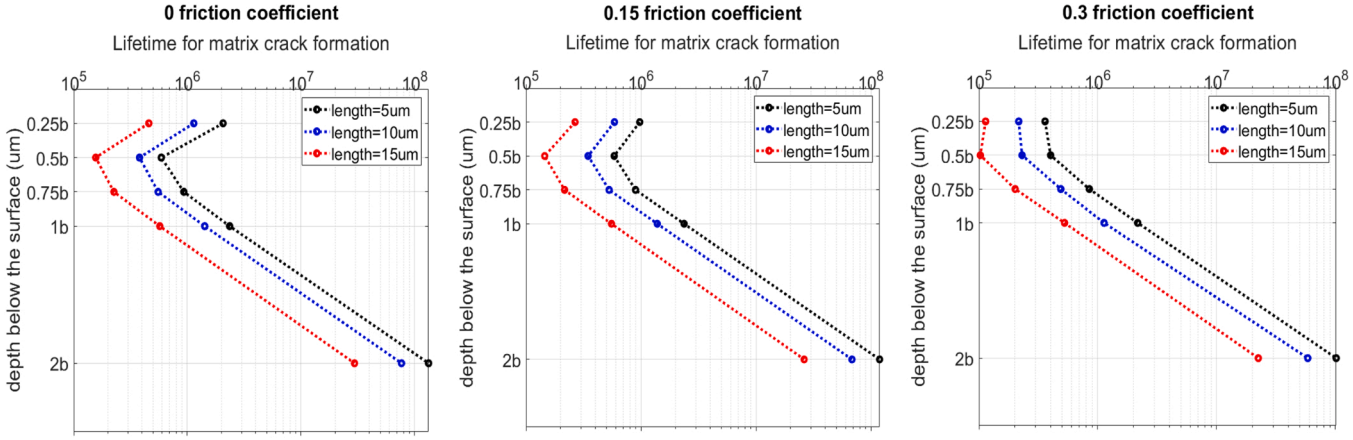


Fig. 15. Role of critical depth and surface traction on matrix crack formation lifetime.

critical depth of 0.5 *b* to 0.75 *b*. Whereas, when the operating conditions generate high traction, the critical depth shifts towards the surface around 0.25 *b*.

6.4. Role of inclusion characteristics

Fig. 16 summarises the effect of different inclusion characteristics such as its length, aspect ratio and orientation against its corresponding initiation lifetime (contour 1) of all debonded inclusions positioned at the depth of 0.5 *b* and 0.75 *b*. Therefore, as mentioned above, for a zero friction coefficient and other inclusion parameters in Table 2, 405 simulations were made. The results from these simulations were then filtered to only debonded cases present at critical depths to plot Fig. 16.

Although median life prediction for 5 and 10 μm inclusion sizes were more or less the same, the mean line shows that increasing inclusion length tends to reduce the crack initiation lifetime. Keeping in mind that these results show the lifetime range for a fixed length that includes varying AR's and angle effects as well. This is in line with the experimental trends from literature [6]. As far as aspect ratio goes, simulations predict AR 3 fails earlier i.e. the inclusion shapes that are in transition from round to elongated experiences higher damage at their tips. For AR 1, the box plot notches extend beyond the 75th percentile. This happens due to small sample size which subsequently means there is uncertainty around the true median value. Finally, coming to inclusion orientation, we can see that the damage profile is symmetrical for + /- 45 deg. owing to the linear elastic nature of the simulations. In this scenario, inclined inclusions (+/- 45 deg.) experience higher damage and initiate earlier compared to equally shaped and sized, horizontally oriented inclusions.

Fig. 17 demonstrates the matrix crack lifetime (contour 2) for different inclusion characteristics positioned at the depth of 0.5 *b* and 0.75 *b*. For inclusion length and AR, the trends remain the same i.e. least lifetime prediction for longer and sharper inclusion. Whereas, opposite trend was observed for inclusion orientation, simulations predict flat inclusions develop matrix crack earlier than inclined ones. This is because, for inclined inclusions, the damage gradients are high close to their tips and they fade away from the interface. For the studied cases, it can be concluded that an elongated and debonded MnS inclusion positioned parallel to the raceway and located at a critical depth of 0.5 *b* to 0.75 *b* is the most detrimental.

7. Conclusions

This study analyses the effects of different inclusion characteristics, degree of debonding, depth and surface traction on RCF damage initiation and crack formation lifetime. As far as methodology goes, based on the location of damage extraction around the interface, a distinction was made between damage initiation and its subsequent progression into a matrix crack. To accomplish this, a non-local fatigue criterion, i.e. the theory of critical distance (TCD), was coupled with a multi-axial critical plane fatigue damage model. The Roessle-Fatemi equation is linked with fatigue damage to calculate damage initiation and matrix crack formation lifetime. This provides a simple and comprehensive lifetime approximation for non-proportional, multi-axial loading, as is the case for rolling contact fatigue.

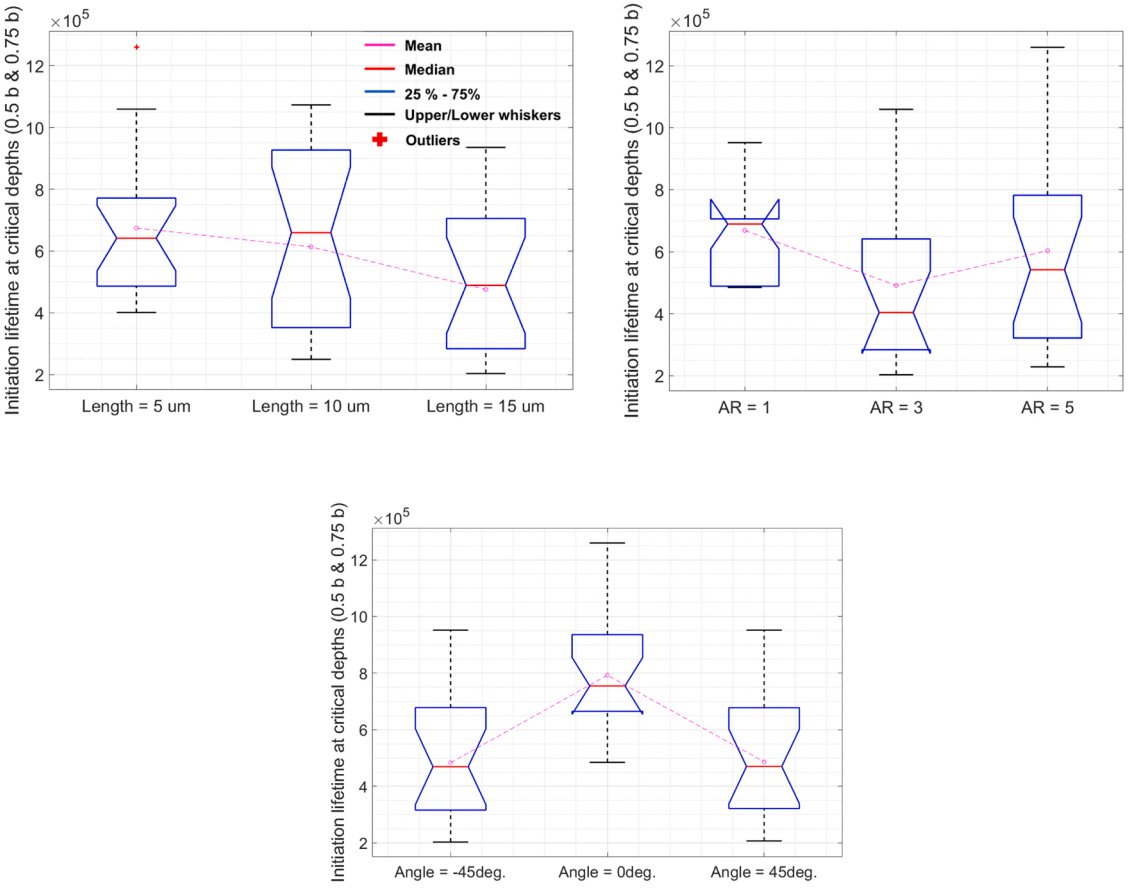


Fig. 16. Damage initiation lifetime trend for different inclusion characteristics (taken at 0.5 b & 0.75 b).

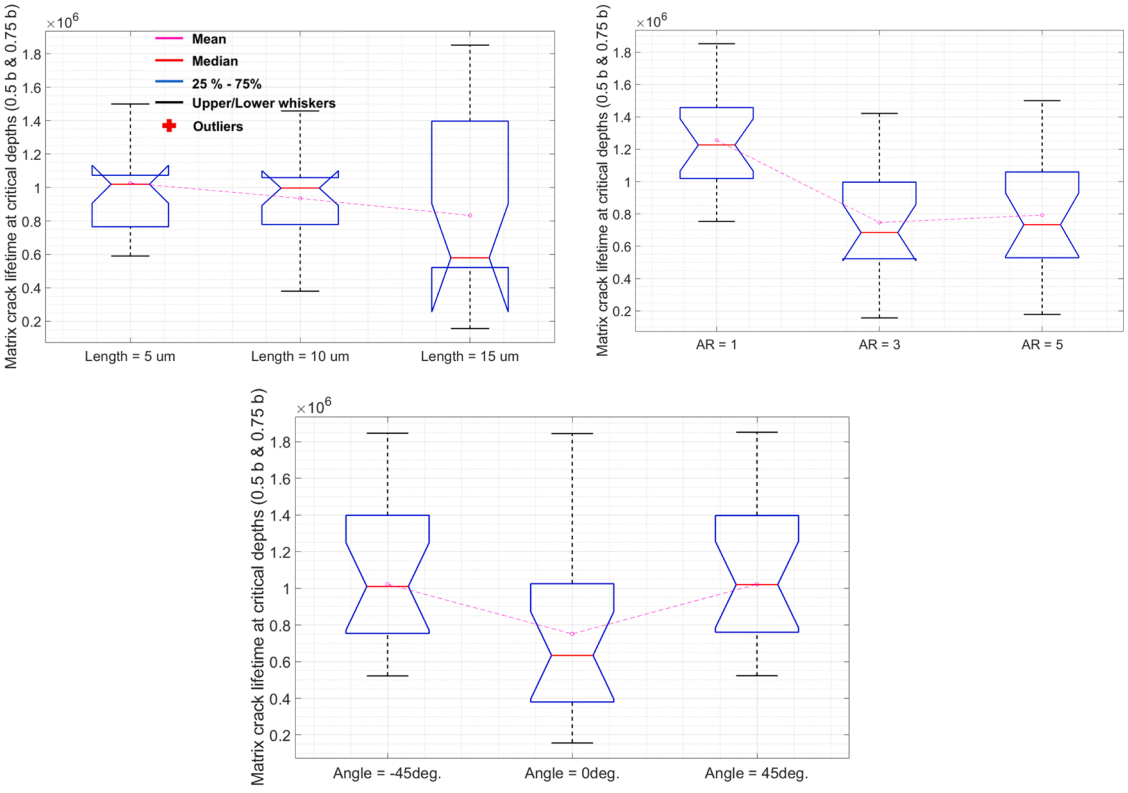


Fig. 17. Matrix crack formation lifetime trend for different inclusion characteristics (taken at 0.5 b & 0.75 b).

From this extensive numerical investigation, the following conclusions can be formulated:

a. Damage builds up around an inclusion leading to debonded interfaces very early on. Inclusions already being stress raisers in matrix, in addition that, these debonded interfaces act as an initial crack and raise the stress concentration leading to accelerated damage initiation.
b. Critical depth where maximum subsurface stresses occur was found to be a function of surface traction. Under normal operating conditions, the earliest crack initiation was observed at a critical depth of 0.5 *b* which is where the maximum shear stress reversals occur. But in case of events which can lead to high surface traction between roller and raceway, this traction effect shifts the maximum subsurface stresses closer to the surface changing the critical depth from 0.5 *b* to 0.25 *b*.
c. As far as role of inclusion size effects goes, the larger the inclusion length, the higher the damage it experiences and takes least lifetime to initiate and develop crack. As far as AR is concerned, sharper inclusions initiate damage earlier than the round ones. For inclusion orientation, although inclined inclusions initiates earlier, flat inclusions develop matrix crack quicker.

Based on these findings, a hypothesis explaining how a crack initiates from a NMI was put forward. All findings from the numerical work were compared with the experimental investigations to understand the mechanism behind crack initiation from a NMI.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

The data that has been used is confidential.

Acknowledgements

The authors gratefully acknowledge the financial support from VLAIO (Flemish Agency for Innovation and Entrepreneurship) and SIM (Strategic Initiative Materials) through project SBO MaSiWEC (HBC.2017.0606).

References

[1] Slack TS, Raje N, Arakere NK, Sadeghi F. A review of rolling contact fatigue. J Tribol 2009;131:041403. https://doi.org/10.1115/1.3209132.
[2] Moghaddam S.M. Analytical and experimental investigation of microstructural alterations in bearing steel in rolling contact fatigue. (2016). Open Access Dissertations. 682. ⟨http://docs.lib.purdue.edu/open_access_dissertations/682⟩.
[3] Lund T. Sub-surface initiated rolling contact fatigue- influence of non-metallic inclusions, processing history, and operating conditions. J ASTM Int 2010;7(5).
[4] Grabulov A. Fundamentals of Rolling Contact Fatigue.(2010). Doctoral thesis. ⟨http://resolver.tudelft.nl/uuid:f7da30ab-2bdc-4d3f-8a26–6c577b33a7c8⟩.
[5] Vegter E., Krock H., Kadin Y., Ocelík V. Nonmetallic Inclusion Bonding in Bearing Steel and the Initiation of White-Etching Cracks. In: Beswick J.M., editor. Bearing Steel Technologies: 11th Volume, Advances in Steel Technologies for Rolling Bearings, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428–2959: ASTM International; 2017, p. 519–532. ⟨https://doi.org/10.1520/STP160020160145⟩.
[6] Zerbst U, Madia M, Klinger C, Bettge D, Murakami Y. Defects as a root cause of fatigue failure of metallic components. II: Non-metallic inclusions. Eng Fail Anal 2019;98:228–39. https://doi.org/10.1016/j.engfailanal.2019.01.054.
[7] Atkinson HV, Shi G. Characterization of inclusions in clean steels: a review including the statistics of extremes methods. Prog Mater Sci 2003;48:457–520. https://doi.org/10.1016/S0079-6425(02)00014-2.
[8] Al-Tameemi HA, Long H, Dwyer-Joyce RS. Damage characterisation of white etching cracks in a black oxide coated wind turbine gearbox bearing. Wear 2019; 432–433:102923. https://doi.org/10.1016/j.wear.2019.05.038.
[9] Bruce T. Analysis of the Premature Failure of Wind Turbine Gearbox Bearings. (2016). PhD thesis, University of Sheffield.
[10] Cerullo M, Tvergaard V. Micromechanical study of the effect of inclusions on fatigue failure in a roller bearing. Int J Struct Integr 2015;6:124–41. https://doi.org/10.1108/IJSI-04-2014-0020.
[11] Al-Tameemi HA, Long H. Finite element simulation of subsurface initiated damage from non-metallic inclusions in wind turbine gearbox bearings. Int J Fatigue 2020; 131:105347. https://doi.org/10.1016/j.ijfatigue.2019.105347.
[12] Sraml M. Numerical procedure for predicting the rolling contact fatigue crack initiation. Int J Fatigue 2003;25:585–95. https://doi.org/10.1016/S0142-1123(03)00019-7.
[13] Sauvage P. On an Extension of the Fatemi and Socie Equation for Rolling Contact in Rolling Bearings.2018. In: Proceedings of the 7th International Conference on Fracture Fatigue and Wear, pp 438–457.
[14] Mahdavi H, Poulios K, Kadin Y, Niordson CF. On the effect of microplasticity on crack initiation from subsurface defects in rolling contact fatigue. Int J Fatigue 2022:106870. https://doi.org/10.1016/j.ijfatigue.2022.106870.
[15] Schäfer B, Sonnweber-Ribic P, ul Hassan H, Hartmaier A. Micromechanical modeling of fatigue crack nucleation around non-metallic inclusions in martensitic high-strength steels. Metals 2019;9:1258. https://doi.org/10.3390/met9121258.
[16] Lai J., Kadin Y. A better understanding of material imperfections | Evolution. Evolution Online 2018. ⟨https://evolution.skf.com/a-better-understanding-of-material-imperfections/⟩ (accessed June 13, 2022).
[17] Jiang Y, Sehitoglu H. A model for rolling contact failure. Wear 1999;224:38–49. https://doi.org/10.1016/S0043-1648(98)00311-1.
[18] Ringsberg JW. Life prediction of rolling contact fatigue crack initiation. Int J Fatigue 2001;23:575–86. https://doi.org/10.1016/S0142-1123(01)00024-X.
[19] Liu Y, Stratman B, Mahadevan S. Fatigue crack initiation life prediction of railroad wheels. Int J Fatigue 2006;28:747–56. https://doi.org/10.1016/j.ijfatigue.2005.09.007.
[20] Melander A. A finite element study of short cracks with different inclusion types under rolling contact fatigue load. Int J Fatigue 1997;19:13–24. https://doi.org/10.1016/S0142-1123(96)00045-X.
[21] Stadler K., Lai J., Vegter R.H. A Review: The Dilemma With Premature White Etching Crack (WEC) Bearing Failures. In: Beswick J.M., editor. Bearing Steel Technologies: 10th Volume, Advances in Steel Technologies for Rolling Bearings, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428–2959: ASTM International; 2014, p. 1–22. ⟨https://doi.org/10.1520/STP158020140046⟩.
[22] Bruce T, Rounding E, Long H, Dwyer-Joyce RS. Characterisation of white etching crack damage in wind turbine gearbox bearings. Wear 2015;338–339:164–77. https://doi.org/10.1016/j.wear.2015.06.008.
[23] Ravi G., De Waele W., Hertelé S. Numerical Methodology to Predict Subsurface Crack Initiation from Non-metallic Inclusions Due to Rolling Contact Fatigue. In: Abdel Wahab M., editor. Proceedings of the 8th International Conference on Fracture, Fatigue and Wear, Singapore: Springer Singapore; 2021, p. 455–471. ⟨https://doi.org/10.1007/978–981-15–9893-7_33⟩.
[24] Ravi G, Daems P-J, Nikolic K, De Waele W, Helsen J, Petrov R, et al. An interdisciplinary framework to predict premature roller element bearing failures in wind turbine gearboxes. Forsch Ing 2021;85:229–40. https://doi.org/10.1007/p0010-021-00463-0.
[25] Standard Specification for High-Carbon Anti-Friction Bearing Steel n.d. ⟨https://www.astm.org/a0295_a0295m-14r20.html⟩ (Accessed June 11, 2022).
[26] Barrow ATW, Kang J-H, Rivera-Díaz-del-Castillo PEJ. The $\epsilon \rightarrow \eta \rightarrow \theta$ transition in 100Cr6 and its effect on mechanical properties. Acta Mater 2012;60:2805–15. https://doi.org/10.1016/j.actamat.2012.01.046.
[27] Keller, Jonathan, Scott Lambert. 2019. Gearbox Instrumentation for the Investigation of Bearing Axial Cracking. Golden, CO: National Renewable Energy Laboratory. NREL/TP-5000–70639. ⟨https://www.nrel.gov/docs/fy18osti/70639⟩.
[28] Vegter R.H., Stadler K. Review on Crack Initiation and Premature Failures in Bearing Applications. In: Beswick J.M., editor. Bearing Steel Technologies: 12th Volume, Progress in Bearing Steel Metallurgical Testing and Quality Assurance, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428–2959: ASTM International; 2020, p. 1–25. ⟨https://doi.org/10.1520/STP162320190054⟩.
[29] Mahdavi H. Micromechanical Modeling of Rolling Contact Fatigue. 2021. Technical University of Denmark. DCAMM Special Report No. S298.
[30] Basan R, Marohnić T. Multiaxial fatigue life calculation model for engineering components in rolling-sliding contact. MATEC Web Conf 2018;188:02016. https://doi.org/10.1051/matecconf/201818802016.
[31] Zeng D, Xu T, Liu W, Lu L, Zhang J, Gong Y. Investigation on rolling contact fatigue of railway wheel steel with surface defect. Wear 2020;446–447:203207. https://doi.org/10.1016/j.wear.2020.203207.
[32] Lai J, Lund T, Rydén K, Gabelli A, Strandell I. The fatigue limit of bearing steels – Part I: A pragmatic approach to predict very high cycle fatigue strength. Int J Fatigue 2012;38:155–68. https://doi.org/10.1016/j.ijfatigue.2011.09.015.
[33] Materials Data for Cyclic Loading - 1st Edition n.d. ⟨https://www.elsevier.com/books/materials-data-for-cyclic-loading/boller/978–0-444–42871-4⟩ (accessed July 21, 2022).
[34] Shamsaei N, Fatemi A. Fatigue life predictions under general multiaxial loading based on simple material properties. SAE Int J Mater Manuf 2011;4:651–8.
[35] Shamsaei N, Fatemi A. Effect of hardness on multiaxial fatigue behaviour and some simple approximations for steels. Fatigue Fract Eng Mater Struct 2009;32:631–46. https://doi.org/10.1111/j.1460-2695.2009.01369.x.
[36] ISO 18265:2013. Metallic materials — Conversion of hardness values ⟨https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/05/38/53810.html⟩ (Accessed August 18, 2022).

[37] Bomidi JAR, Weinzapfel N, Sadeghi F, Liebel A, Weber J. An improved approach for 3D rolling contact fatigue simulations with microstructure topology. Tribol Trans 2013;56:385–99. https://doi.org/10.1080/10402004.2012.754072.
[38] Salajegheh N, McDowell DL. Microstructure-sensitive weighted probability approach for modeling surface to bulk transition of high cycle fatigue failures dominated by primary inclusions. Int J Fatigue 2014;59:188–99. https://doi.org/10.1016/j.ijfatigue.2013.08.025.
[39] Kadin Y, Sherif MY. Energy dissipation at rubbing crack faces in rolling contact fatigue as the mechanism of white etching area formation. Int J Fatigue 2017;96:114–26. https://doi.org/10.1016/j.ijfatigue.2016.11.006.