



Challenges in determination of microscopic degree of cleanliness in ultra-clean gear steels

D. Fuchs¹ · T. Tobie¹ · K. Stahl¹

Received: 4 August 2021 / Revised: 2 September 2021 / Accepted: 22 September 2021 / Published online: 11 February 2022
© The Author(s) 2022

Abstract

Shot-peened, case-hardened gears can fail in the tooth root due to crack initiation below the surface of the steel matrix. Here, the crack is initiated at a non-metallic inclusion in the center of a so-called fisheye. This failure type can lead to a reduced endurance fatigue limit of the gears. It is for this reason that, over the last decade, much effort has been invested by steel manufacturers to reduce the non-metallic inclusion content of gear steels so as to mitigate or even completely prevent such crack initiation. These ultra-clean gear steels were achieved by various measures in the steel production process. However, as a result, the remaining non-metallic inclusions are inhomogeneously distributed in the steel volume in terms of both size and location. However, due to the inhomogeneity of ultra-clean steels, the question arose if the values derived according to the standards are still representative of ultra-clean steel batches. The results show that the standards can still be applied, but more effort must be applied. To determine the degree of cleanliness, six microsections are currently evaluated according to steel test specification (SEP) 1571, method K. It is shown that an examination of 24 microsections starting from size class 0 seems beneficial to get more reliable and comparable results of the degree of cleanliness of these ultra-clean gear steels. In addition, it is shown that a high degree of cleanliness has been achieved for all steel batches investigated with the measures taken in the steel production process.

Keywords Degree of cleanliness · Non-metallic inclusion · Evaluation of inclusion · SEP 1571 · DIN 50602 · Ultrasonic immersion testing · Ultra-clean gear steel · Gear

1 Introduction

Recent trends toward green energy and the electrification of the powertrain have led to more stringent or completely new requirements for gearboxes. To meet higher torque demands, the tooth root bending strength of case-hardened gears can be increased by applying a shot-peening process. However, in such shot-peened gears, crack initiation can occur below the surface. This failure mechanism is called fisheye failure, which is manifested as a crack that initiates at a non-metallic inclusion in the center of the fisheye. Over the last decade, a great deal of effort has been expended by steel manufacturers in reducing the non-metallic inclusion content of gear steels. These ultra-clean

gear steels were achieved by various measures in the steel production process. The goal is to mitigate or even completely prevent the occurrence of fisheye failure in high-strength gears. The reduced number of non-metallic inclusions in the steel volume means that those inclusions that exist are distributed more inhomogeneously throughout the steel matrix in terms of their size and location.

The degree of cleanliness can be specified by determining the non-metallic inclusion content of a defined steel volume according to certain standards, such as ISO 4967 or steel test specification (SEP) 1571. These are established standards used in the gear industry for characterizing and comparing steel batches. To determine a characteristic value for the degree of cleanliness, six microsections are generally evaluated according to the standards.

To gain more knowledge and allow a probable correlation between the degree of cleanliness and the tooth root bending strength of gears requires due characterization of the inclusion content of the complete steel volume.

✉ D. Fuchs
fuchs@fzg.mw.tum.de

¹ Gear Research Center (FZG), Technical University of Munich, 85748 Garching, Germany

However, due to the inhomogeneity of ultra-clean gear steels, the question arises if the values derived based on six microsections are still representative of ultra-clean steel batches or if a greater effort, for example an evaluation of a higher number of microsections, must be made. In addition, it will be checked whether the measures taken in the steel production process have resulted in ultra-clean gear steels.

2 State of scientific knowledge

2.1 Influence of fisheye failure on fatigue strength

Figure 1 shows schematically an example of a fisheye fracture in a gear. A non-metallic inclusion in the center of the fisheye is responsible for the crack initiation. An optically dark area often surrounds the non-metallic inclusion. This fracture mechanism occurs foremost with higher numbers of load cycles and leads to a decrease in bending strength. It is therefore apparent that the degree of non-metallic content in the form of inclusions has a great influence on the strength of high-strength gears [1–3].

Tridello et al. [4] show that the combination of the tested specimen's volume and the inclusion distribution has an impact on the fatigue strength. Consequently, more and more effort is invested nowadays in the process of making and characterizing steel [5]. The degree of cleanliness represents a way of describing the content of non-metallic inclusions in materials. Temmel et al. [6] show, however, that not all methods are suitable for characterizing steels with a low degree of sulfur. In the following, standardized methods for characterizing the degree of cleanliness are presented.

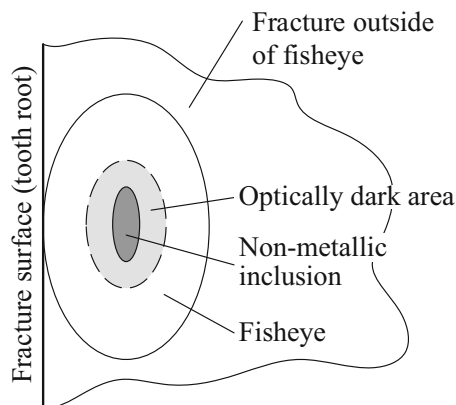


Fig. 1 Schematic of fisheye fracture in a gear

2.2 Methods for characterizing degree of cleanliness

The degree of cleanliness is generally divided into macroscopic and microscopic inclusions in accordance with common standards. Macroscopic inclusions are of a size greater than 0.03 mm^2 according to SEP 1571 and Deutsches Institut für Normung (DIN) 50602. The macroscopic degree of cleanliness is determined using the blue fracture test according to SEP 1584, the step turning test according to SEP 1580 or ultrasonic immersion testing according to SEP 1927 or ASTM A388/A388M-19. For the microscopic degree of cleanliness, on the other hand, the following standards apply: SEP 1571, DIN 50602, DIN EN 10247, ISO 4967, ASTM E45-13 and ASTM E2283-08.

The following is a brief description of the standards used in this publication:

- (1) Ultrasonic immersion testing according to SEP 1927: This method compares the test specimens with a reference block. The specimens are investigated in a water tank using an ultrasonic immersion search unit with a pulse repetition frequency of 10 MHz. The macroscopic inclusions are categorized according to their position, size and occurrence.
- (2) Evaluation of inclusions according to SEP 1571: The degree of cleanliness is determined on a particle basis by means of microsections. Usually (at least) six microsections are evaluated to determine a value for the degree of cleanliness. The standard differentiates between inclusions of type A (typically manganese sulfides), B (crumbled or elongated stringer aluminum oxides), C (silicon oxides), D (globular aluminum oxides) and D_{sulf} (calcium sulfides). The standard describes three methods:
 - a. Method M: maximum inclusion value
 - b. Method K: mean inclusion value
 - c. Method E: extreme value, analyzed based on size class (reference area: $100,000 \text{ mm}^2$)
 Methods M and K use six microsections for evaluation; method E uses at least 12 microsections, and 24 microsections are recommended.
- (3) Evaluation of inclusions according to DIN 50602. This standard is very similar to SEP 1571 and also employs methods M and K. Inclusions are divided into four types: SS (sulfide inclusion), OA (oxide stringer inclusion), OS (oxide inclusion in elongated form), and OG (oxide inclusion in globular form). The types are comparable to types A, B, C and D in SEP 1571.
- (4) Evaluation of inclusions according to ISO 4967. Standard diagrams are used to determine the degree

of cleanliness. The standard discriminates between five types: A (sulfide type), B (aluminate type), C (silicate type), D (globular oxide type) and DS (single globular type). A further distinction is made between fine and thick inclusions.

- (5) Extreme value analysis according to ASTM E2283-08. Six specimens are used according to the test method ASTM E45-13. The greatest maximum length value is recorded for each specimen. The procedure is repeated three times, resulting in 24 inclusion lengths. The greatest inclusion length is determined based on these 24 lengths and is expected to lie within a reference area of 150,000 mm².

2.3 Specifications for degree of cleanliness in ISO 6336

Part 5 of ISO 6336 categorizes gear steels into three material quality classes: ML, MQ and ME. The degree of cleanliness is an important differentiating factor. Table 1 lists the cleanliness requirements of case-hardened wrought steels. The higher the material quality class, the greater the requirements.

2.4 Influences on determination of degree of cleanliness

Most of the standards presented use (at least) six microsections to determine the degree of cleanliness. To limit the time and costs, six microsections are generally used in industrial practice. Murakami [7], for example, shows that the selection of the inspection plane has a great influence on the cleanliness value as shown in Fig. 2.

Table 1 Cleanliness requirements of case-hardened wrought steels according to ISO 6336, part 5

Requirement	Material quality class	
	ML	MQ, ME
Degree of cleanliness	No specification	Cleanliness in accordance with ISO 4967, procedure in accordance with method A, inspected area of approximately 200 mm ²
Ultrasonic test according to ASTM A388	No specification	Required

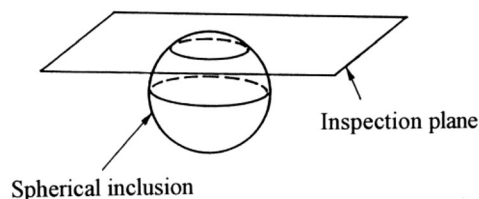


Fig. 2 Influence of inspection plane on apparent inclusion size [7]

3 Summary of current state of knowledge

This state of scientific knowledge shows that the degree of cleanliness has a considerable impact on the fatigue strength of shot-peened gears. Nowadays, ultra-clean materials are more often used to achieve higher tooth root fatigue strengths. In this case, the few remaining inclusions are distributed inhomogeneously in the material in terms of size and location. According to the current standards for determining the degree of cleanliness, only six microsections are evaluated. Therefore, the question arises if ultra-clean gear steels can still be evaluated using the standards currently applied in industrial practice. Despite an extensive literature research, no investigations into this topic were found.

4 Aim of investigation

In the course of Forschungsvereinigung Antriebstechnik e.V. (FVA) research project 293 IV [8], which builds on the results and conclusions of [2, 9–14], extensive experimental investigations were performed. The focus was on the very high cycle fatigue range of shot-peened, and case-hardened gears made out of ultra-clean gear steels. Compressive residual stresses are introduced through the shot-peening process, which can lead to higher load carrying capacities.

However, fisheye fractures at non-metallic inclusions can occur in the tooth root fillet of shot-peened, and case-hardened gears in the very high cycle fatigue range. Therefore, in the framework of this research project, ultra-clean gear steels were used to examine, if by using such ultra-clean gear steels, a crack initiation below the surface at non-metallic inclusions can be prevented. As a result, higher load carrying capacities are expected in the very high cycle fatigue range. One of the main goals of this research project was to correlate the load carrying capacity of shot-peened and case-hardened gears with the microscopic degree of cleanliness of ultra-clean gear steels.

However, a reliable correlation of the degree of cleanliness and the load carrying capacities can only be made with statistically validated values. As a result, the questions arose whether a reliable degree of cleanliness value for

these ultra-clean gear steels can be determined with the current procedures according to the standards, like SEP 1571, and whether high degrees of cleanliness have been achieved with the measures taken in the steel production process. This publication addresses these questions and deals with following influence factors: starting size class, characteristic value, number of microsections, respectively, samples and the influence of extreme value methods. In addition, further questions are clarified.

5 Region of interest

For the investigations on FZG back-to-back test rig and Pulsator test rig (see Fig. 3), gear sizes with a normal module $m_n = 1.5$ and 5 mm were used. The region of interest (ROI) for the cleanliness inspections was placed in the later tooth root fillet of the gear as shown in Table 2. The approximate area of the ROI in both cases is 210 mm^2 for each microsection. The required microsections were taken from a billet, which was divided into segments.

6 Steel batches and documentation of material

6.1 Steel batches

The gear steel batches investigated are MnCr-, CrNiMo-, NiMo- and NiCr-alloyed gear steels. Table 3 presents an overview of the steel batches, alloy systems, casting method, diameters of the steel bars, reduction ratios and features.

Four of the nine steel batches are from continuous casting, whereas the other steel batches are from ingot

casting. The diameter of the steel bars ranges from 100 to 140 mm, and the reduction ratios are all above the specification value of 5:1 for continuous casting according to ISO 6336, part 5. Steel batch OW5 shows the highest value of 27:1. Furthermore, the steel batches are all classified in the scatter band HH for a high hardenability according to DIN EN ISO 683, part 3. The feature of steel batches OW1, OW7 and S9 is a modified calcium treatment with additional recrystallization annealing. Steel batches S4 and S6 are open melted, whereas steel batch S8 is electroslag remelted. Steel batch OW4 has a low sulfur content, and steel batch OW5 has a low aluminum content. A modified rolling/forging process was used for steel batch OW3.

6.2 Chemical analysis

Chemical analysis and oxygen content data are listed in Tables 4 and 5. All steel batches are within the limits specified in DIN EN ISO 683, part 3, which are also listed in Table 4. Furthermore, all steel batches reveal oxygen contents that are below the maximum specification of 25×10^{-6} according to ISO 6336, part 5 as shown in Table 5.

7 Demonstration of effectiveness of measures taken to achieve ultra-clean gear steels

Various measures have been taken to produce ultra-clean gear steels as shown in Table 3. In the following, it will be checked whether the measures taken in the steel production process have resulted in ultra-clean steels.

All gear steels investigated in this paper are classified as material quality class ME. Ultrasonic immersion testing to

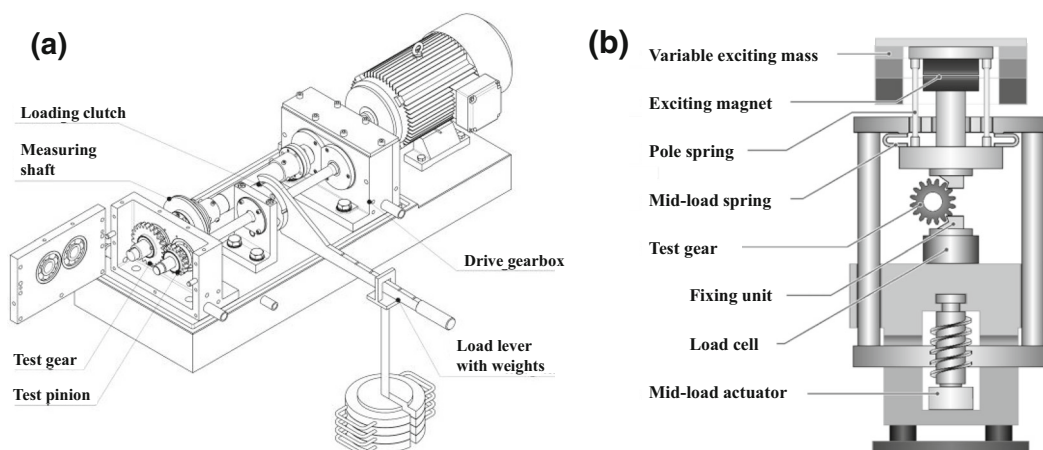


Fig. 3 Test rigs used for investigations on tooth root bending strength. **a** FZG back-to-back test rig (center distance $a = 91.5$ mm) according to DIN ISO 14635, part 1; **b** pulsator test rig [15]

Table 2 Definition of ROI

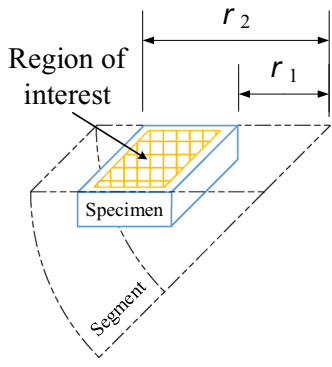
Schematic diagram	m_n / mm	Radius r_1 /mm	Radius r_2 /mm	Length/ mm	Area/ mm ²	Region of interest
	1.5	35	49	15	210	Later tooth root fillet region of gear
	5.0	49	63	15	210	Later tooth root fillet region of gear

Table 3 Overview of investigated test steel batches and their characteristics

Steel batch	Alloy system	Casting method	Diameter of steel bar/mm	Reduction ratio	Feature
OW1	20MnCr5	Ingot	125	8:1	Modified calcium treatment with additional recrystallization annealing
S4	20MnCr5	Continuous	105	8:1	Open melted
S6	20MnCr5	Continuous	105	8:1	Open melted
S8	20MnCr5	Ingot	100	8:1	Electroslag remelted
OW4	20MnCr5	Continuous	100	8:1	Low sulfur content
OW3	18CrNiMo7-6	Continuous	140	8:1	Modified rolling/forging process
OW5	18CrNiMo7-6	Ingot	140	27:1	Low aluminum content
OW7	20NiMo9-7	Ingot	130	12:1	Modified calcium treatment with additional recrystallization annealing
S9	18NiCr5-4	Ingot	110	17:1	Modified calcium treatment with additional recrystallization annealing

determine the macroscopic degree of cleanliness was therefore mandatory and was performed according to SEP 1927 (except for steel batches S4, S6 and S8, due to lack of raw material). This is in deviation to ISO 6336, part 5, which specifies ultrasonic immersion testing according to ASTM A388/A388M-19. However, the ultrasonic immersion testing according to SEP 1927 is a well-established practice among German and European industrial gear manufacturers and is usually used instead of the procedure according to ASTM A388/A388M-19. The determination procedure was performed at 10 MHz. The region of interest was defined as 6–35 mm from the surface. In the case of steel batch OW4, four specimens were examined; in all other cases, two.

Only in the specimens of steel batches OW3 and OW4, echoes could be detected by ultrasonic immersion testing

using the chosen parameters, see Table 6 and Fig. 4. It should be noted that the echoes in steel batch OW4 are located foremost in the core region of the bar and not near the later tooth root fillet. Steel batch OW3 shows echoes distributed over the entire region of interest investigated.

Figure 5 provides a first impression of the microscopic degree of cleanliness of these ultra-clean steel batches. The values were determined according to ISO 4967, method A. The steel batches OW7 and S9 show the best (lowest) cleanliness index. Steel batches OW1, S4, S6 and OW3 show values equal to or higher than five, while the values for the other steel batches are in between. All steel batches are below the threshold values for material quality ME based on ISO 4967, method A according to part 5 of ISO 6336 for case-hardened wrought steels.

Table 4 Chemical analysis of steel batches and limits according to DIN EN ISO 683, part 3 and for steel batch 20NiMo9-7 according to material inspection document of steel manufacturer

Steel batch	Alloy system	Chemical composition/mass%									
		C	Mn	Cr	Ni	Mo	S	Al	Cu	P	Si
OW1	20MnCr5	0.20	1.17	1.15	0.12	0.05	0.019	0.026	0.18	0.018	0.23
S4	20MnCr5	0.21	1.20	1.14	0.15	0.04	0.028	0.030	0.14	0.009	0.15
S6	20MnCr5	0.18	1.25	1.07	0.22	0.06	0.012	0.023	0.09	0.012	0.26
S8	20MnCr5	0.18	1.12	1.15	0.19	0.05	0.006	0.019	0.12	0.016	0.16
OW4	20MnCr5	0.21	1.31	1.25	0.16	0.02	0.006	0.031	0.12	0.010	0.17
OW3	18CrNiMo7-6	0.20	0.54	1.74	1.56	0.29	0.011	0.025	0.20	0.011	0.26
OW5	18CrNiMo7-6	0.19	0.54	1.65	1.42	0.27	0.001	0.009	0.10	0.007	0.27
OW7	20NiMo9-7	0.20	0.23	0.38	2.24	0.67	0.001	0.082	0.16	0.004	0.08
S9	18NiCr5-4	0.18	0.78	1.15	1.41	0.09	0.001	0.033	0.13	0.008	0.31
20MnCr5	Maximum	0.22	1.40	1.30	–	–	0.035	–	–	0.025	0.40
	Minimum	0.17	1.10	1.00	–	–	–	–	–	–	–
18CrNiMo7-6	Maximum	0.21	0.90	1.80	1.70	0.35	0.035	–	–	0.025	0.40
	Minimum	0.15	0.50	1.50	1.40	0.25	–	–	–	–	–
18NiCr5-4	Maximum	0.21	0.90	1.20	1.50	–	0.035	–	–	0.025	0.40
	Minimum	0.16	0.60	0.90	1.20	–	–	–	–	–	–
20NiMo9-7	Maximum	0.21	0.30	0.40	2.35	0.70	0.002	–	–	0.025	0.10
	Minimum	0.18	0.22	0.35	2.25	0.67	–	–	–	–	–

Table 5 Total oxygen content ($w_{O_{total}}$) of steel batches (10^{-6})

	MnCr-alloyed		CrNiMo-alloyed		NiMo-alloyed	NiCr-alloyed	
	S4, S6, S8	OW4	OW3	OW5	OW7	S9	
OW1	Not determined		9	14	18	6	13

Table 6 Existence of echoes

Steel batch	OW1	S4, S6, S8	OW4	OW3	OW5	OW7	S9
Echoes	✗	Not determined	✓	✓	✗	✗	✗

All steel batches can be classified in the category “ultra-clean gear steels”. For this reason, all steel batches are used in the following to derive the factors influencing the determination of the microscopic degree of cleanliness.

8 Systematic determination of influence factors on determination of microscopic degree of cleanliness according to SEP 1571 of ultra-clean gear steels

The determination of the microscopic degree of cleanliness was done in the following according to SEP 1571. This is in deviation to ISO 6336, part 5, which specifies for gear steels with material quality class ME threshold values based on ISO 4967, method A. However, the determination

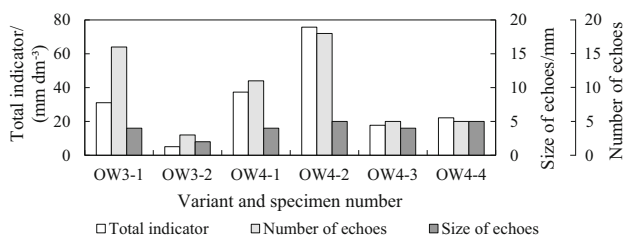


Fig. 4 Results of ultrasonic immersion testing according to SEP 1927

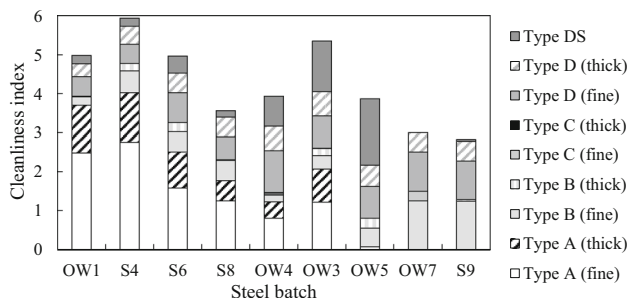


Fig. 5 Stacked bar diagram of degree of cleanliness according to ISO 4967, method A of steel batches investigated

of the microscopic degree of cleanliness according to SEP 1571 is a well-established practice among German and European industrial gear manufacturers and is usually used instead of the procedure according to ISO 4967, method A.

In the following, the influences on the individual methods according to SEP 1571 are examined in more detail. For SEP 1571, method K, the influence of starting size class, characteristic value, number of microsections and number of samples is investigated. The influence of number of microsections and samples is also investigated for SEP 1571, method M. Finally, method E according to SEP 1571 is compared to method M. SEP 1571 is used here as an example standard for the degree of cleanliness. It is assumed that the derived conclusions can also be applied in full or at least in large part to other standards, such as ISO 4967.

8.1 SEP 1571, method K

For gear steels, the determination of the microscopic degree of cleanliness usually starts at size class 4 in industrial practice to limit time and costs. However, the question is, if ultra-clean steels can still be evaluated when starting from size class 4 or more effort is required.

Figure 6 shows characteristic values for the surface area according to SEP 1571, method K, from size class -2 up to size class 4 for steel batches OW3 and OW4. For the evaluation, the results of four laboratories were used in the following. It can be seen that the steels show comparable values for size class 4 both for oxide and sulfide inclusions.

However, with lower starting classes, more and more differences between the steel batches are visible. For oxide inclusions, the tendencies are visible from size class 3 on and for sulfide inclusions from size class 1. Differentiation becomes much more pronounced from grade 2 or 1 onward. Therefore, it seems helpful to choose a starting size class of 1 or below for comparing ultra-clean gear steels.

According to the industrial practice of gear manufacturers, an overall total characteristic value is usually given for the degree of cleanliness according to SEP 1571, method K. However, in this case, the values for oxide and sulfide inclusions are combined, and no separate examination is possible. For the following evaluation, further microsections of four laboratories were used. This leads to slightly different values in comparison to the values in Fig. 6.

Figure 7 shows a comparison of overall total characteristic and total characteristic values. The values for steel batches OW5, OW7 and S9 are determined solely by oxide inclusions. Steel batches S8, OW4 and OW3 show a mix of sulfide and oxide inclusions. Therefore, for a more detailed differentiation of different steel batches, it seems useful to consider the characteristic values for oxide and sulfide inclusions separately.

According to the standards presented in Sect. 2, in general one sample ($\cong 6$ microsections; \cong means the mathematical symbol for “equal by definition”) is used to determine the degree of cleanliness. However, it is often the case in industrial practice that less than six microsections are analyzed for one sample, to save time and costs. Therefore, the following includes a check as to whether six microsections are necessary for ultra-clean steels. Furthermore, due to the limited area investigated with a single sample, the question arises if one sample only is representative of ultra-clean steels. Steel batch OW3 is used as an example, because of its broad database.

Figure 8 shows an overview of one sample evaluation for sulfide inclusions based on six microsections, using steel batch OW3 as an example. From size class 4 onward, there is hardly any visible difference between the microsections. However, starting from size class 0, the difference is great. The result of the sample examination is therefore strongly dependent on the number of microsections. In addition, the standards specify a minimum area of $100 \mu\text{m}^2$ for each microsection. The results in Fig. 8 are based on microsections with an area of $210 \mu\text{m}^2$. Even with this larger microsection area, great differences are visible.

Figure 9 shows the characteristic values of the surface area according to SEP 1571, method K for sulfide and oxide inclusions, using the example of steel batch OW3 and based on the results from several laboratories. The size classes 4 and 0 are compared in Fig. 9a and b. In the

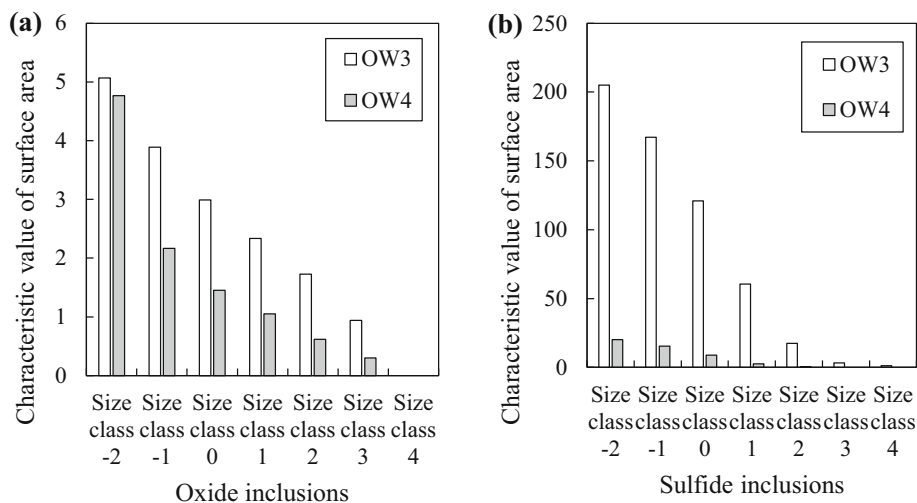


Fig. 6 Comparison of total characteristic values according to SEP 1571, method K for different size classes of oxide (a) and sulfide (b) inclusions using steel batches OW3 and OW4 as examples

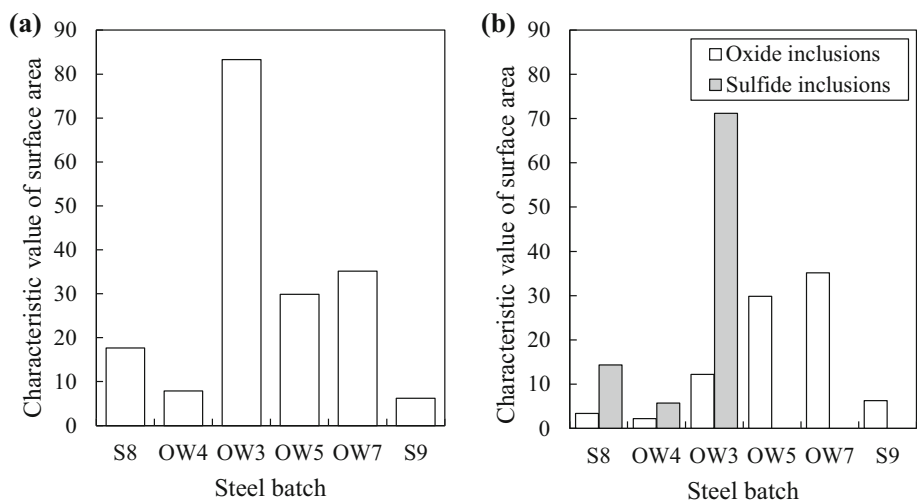


Fig. 7 Overall total characteristic value (a) and total characteristic value (b) starting from size class 0 according to SEP 1571, method K

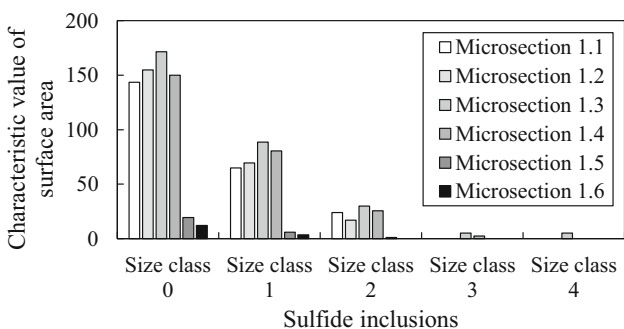


Fig. 8 Detailed analysis of sulfide inclusions in sample 1 of steel batch OW3 according to SEP 1571, method K

evaluation of sulfide inclusions, it can be seen that starting from size class 4, there are no noteworthy differences (range of 1.37) between the different sample numbers, as

shown in Fig. 9a. Though, significant differences are evident when comparing the values starting from size class 0 as shown in Fig. 9b. The range and standard deviation are much higher. The same tendency is also visible with oxide inclusions as shown in Fig. 9c and d. However, the numbers and differences are much smaller, because the inclusions in steel batch OW3 are predominantly sulfides.

8.2 SEP 1571, method M

According to method M of SEP 1571, the largest inclusion of each inclusion type is evaluated for each of the six microsections. The mean value of these six microsections gives the maximum inclusion value according to method M for one sample. Figure 10a shows the mean values according to method M for 1 to 36 microsections for each

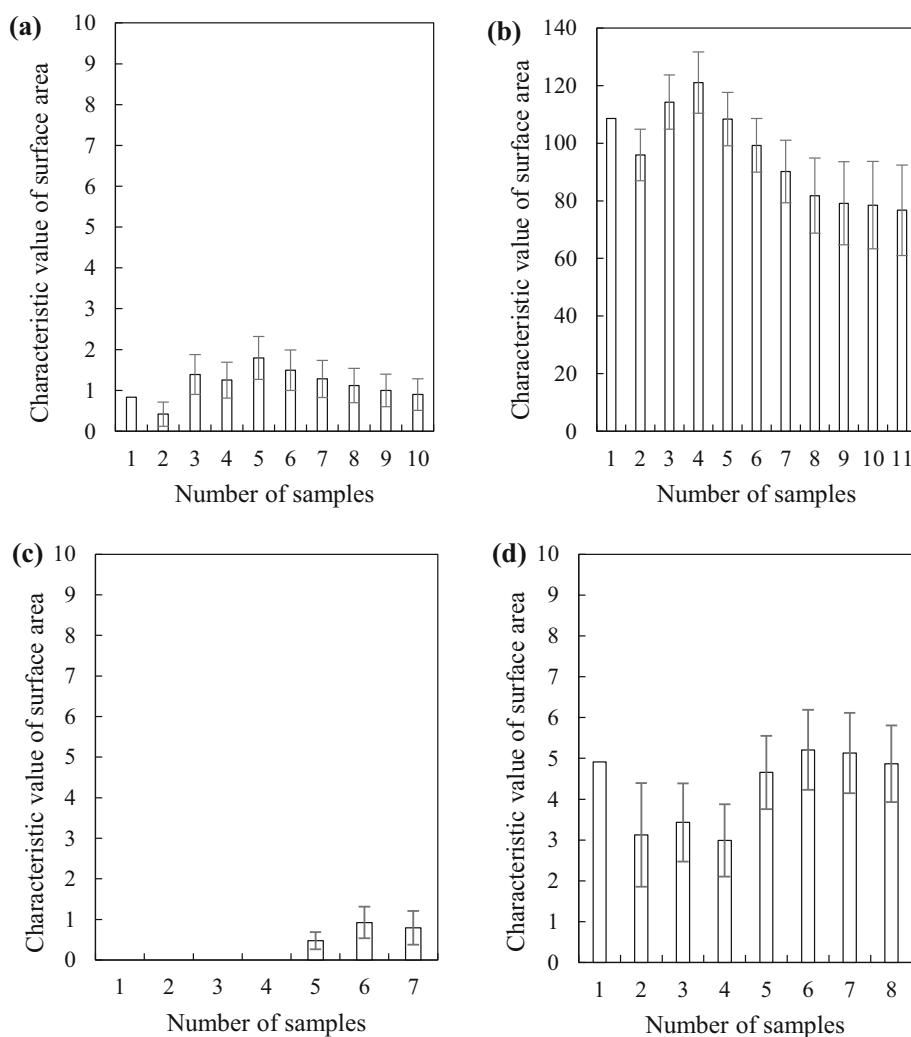


Fig. 9 Characteristic values of surface area according to SEP 1571, method K for sulfide inclusions starting from size class 4 (a) and 0 (b), and for oxide inclusions, starting from size class 4 (c) and 0 (d) using example of steel batch OW3

inclusion type present. It can be seen that after approximately 12 to 18 microsections, the maximum inclusion size class remains nearly constant. Figure 10b shows the mean values of 2–6 samples (one sample \cong six microsections) for each inclusion type present. It can be seen that all sample numbers show different values and standard deviations. However, except for D_{sulf} inclusion types, there is almost no visible influence on the maximum inclusion size class.

8.3 SEP 1571, method E

It was shown that above six or eight samples, respectively, the evaluations show more reliable values for the degree of cleanliness according to SEP 1571, method K, starting from size class 0. However, this number of samples is not practical for each steel evaluation in industrial practice. A better approach could be to use a combination of the

previous methods and an extreme value method, like SEP 1571, method E. Method E of SEP 1571 uses at least 12 microsections (\cong 2 samples). However, the recommended number of microsections is 24 (\cong 4 samples). An extreme value for an inclusion size is determined based on a Gumbel distribution, in which the largest inclusion in each individual microsection is taken. Based on an extrapolated virtual test surface area of 10,000 mm², a value can be derived for a maximum expected inclusion size or inclusion size class, respectively.

The evaluation according to method E is based on the size of the largest inclusion in each individual microsection, which means that an extended evaluation according to SEP 1571, method M can also be made. Figure 11a compares the maximum inclusion size class of both methods. Particularly with inclusion type B, the size class increases from two to four as a result of the extreme value analysis. The value obtained using method E is therefore more

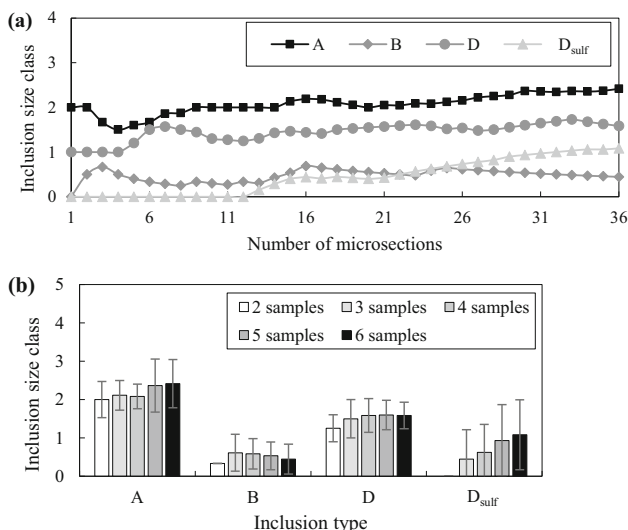


Fig. 10 Mean values for 1 to 36 microsections according to SEP 1571, method M (a), and mean value and standard deviation of 2 to 6 samples (b) combined using steel batch OW3 as an example

conservative, because it predicts larger inclusion sizes. Figure 11b presents the steel batches S4, S6, S8, OW4 and OW3 in the form of a stacked bar chart. The number of microsections evaluated for these steel batches was 24. An evaluation based on this representation of results seems helpful when comparing ultra-clean steel batches.

9 Further investigations

Further questions are clarified in the following. Using a statistical method, the minimum number of samples for SEP 1571, method K, is determined, when limiting the

number of samples to a maximum of 24. The value comparability of SEP 1571 and DIN 50602 is investigated in more detail. The influence of the gear steel alloy system and slight differences in inclusion content is investigated as well. Next, it is investigated, if the characteristic values of SEP 1571, method K and ISO 4967, method A, are comparable. The extreme values methods according to SEP 1571, method E and ASTM E2283-08 are compared, before the results from the laboratories are contrasted as a final point.

9.1 Statistical method of determining minimum number of samples for SEP 1571, method K

In Sect. 8, it was shown that six or eight samples, respectively, would be needed to deliver more reliable cleanliness values for ultra-clean gear steels. For the evaluations according to SEP 1571, methods M and K, (at least) one sample (\cong 6 microsections) should be used. For method E, two or four samples are recommended (\cong 12 or 24 microsections, respectively) in the standard. In industrial practice, the time and costs are important factors. Therefore, this section examines whether a reliable value can be delivered by up to four samples (\cong 24 microsections). The following uses sulfide inclusions as an example, starting from size class 0 of steel batch OW3. To eliminate probable laboratory influences, all values are from a single laboratory.

Figure 12a shows the individual values for each microsection. Random microsection combinations are shown based on these values, as shown in Fig. 12b. After combining approximately 18 microsections, the deviations between the mean values are within an acceptable range. It

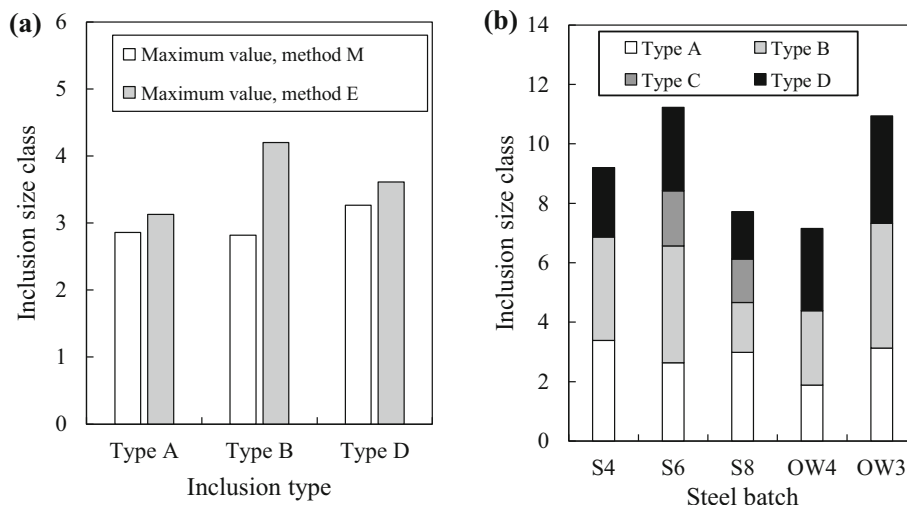


Fig. 11 Maximum inclusion size class according to SEP 1571, methods M and E using steel batch OW3 as an example (a), and stacked bar chart of inclusion types present according to SEP 1571, method E (b) for various steel batches (24 microsections)

seems that at least three samples ($\cong 18$ microsections) are helpful to get reliable cleanliness values, when limiting the number of microsections to 24.

9.2 Value compatibility between SEP 1571 and DIN 50602

According to its preamble, SEP 1571 is intended to be the value-compatible successor to DIN 50602. This value compatibility was previously confirmed by an interlaboratory test by the steel institute VdEh, which publishes the SEP 1571 specification. However, in the interlaboratory test, steels with a common degree of cleanliness and clean steels for bearing applications were used. This publication will therefore also examine whether value compatibility also exists for ultra-clean gear steels.

Figure 13a compares the overall total characteristic value according to method K. A good correlation can be seen between the two standards for evaluations starting from size classes 0, 1 and 2. Figure 13b and c shows the total characteristic values for oxide and sulfide inclusions, respectively. Again, there is a good correlation between the values obtained with method K of SEP 1571 and DIN 50602. Figure 14 compares the maximum inclusion sizes with method M, using steel batch OW3 as an example. This evaluation is based on six specimens ($\cong 36$ microsections). Some slight differences can be seen. However, except for the inclusion type A/SS, the same maximum inclusion size class is always given.

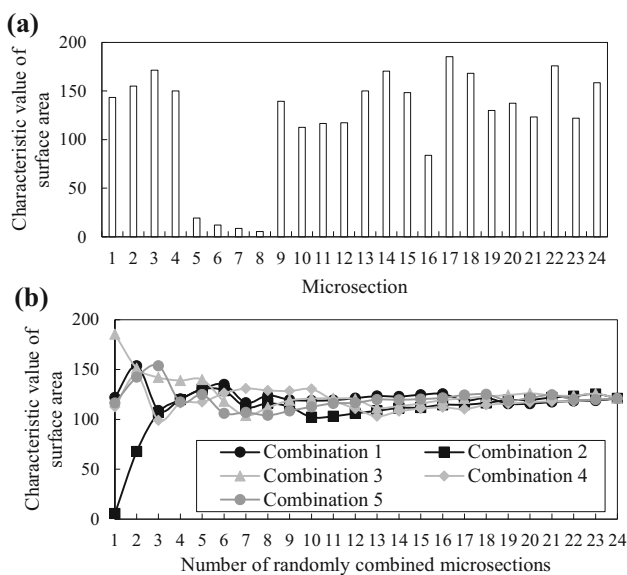


Fig. 12 Individual values for each microsection (a) and mean values for random microsection combinations (b) according to SEP 1571 method K, using example of steel batch OW3 (starting from size class 0)

9.3 Influence of gear steel alloy system and slight differences in inclusion content

So far, steel batch OW3 (18CrNiMo7-6, see Table 3) was used, due to its extended database compared to the other steel batches. The following now evaluates whether the conclusions drawn are also valid for other steel alloy systems and slightly different degrees of cleanliness. Figure 15 shows the results for steel batches OW1, OW5, OW7 and S9. To limit the evaluation effort, size class 2 was chosen as the starting size class for steel batch OW1, due to its higher non-metallic inclusion content. Steel batches OW7 and S9 show the lowest non-metallic inclusion content of all steel batches. The results for each steel batch are based on four samples ($\cong 24$ microsections). Virtually no differentiation is possible using values from size class 4 onward. However, starting from size classes below 4, differences between the steels batches are visible. As a result, it can be noted that the already drawn conclusions are also valid for other steel alloy systems and slightly different degrees of cleanliness.

9.4 Comparison of characteristic values of SEP 1571, method K and ISO 4967, method A

Part 5 of the ISO 6336 gear standard presents the specifications governing the cleanliness of gear steels. Here, the degree of cleanliness shall be determined according to ISO 4967, method A. Although the values of SEP 1571, method K and ISO 4967, method A are not directly comparable, the expectation is that comparing steel batches according to these standards should at least show similar tendencies.

Figure 16 compares the tendencies of SEP 1571, method K and ISO 4967, method A. Figure 16a compares the total characteristic values for each standard. Inclusions categorized as fine and thick are considered here. The same tendency can be seen with the steel batches OW1, S4, S6, OW7 and S9. When only comparing the steel batches S8, OW4, OW3 and OW5, a similar tendency is also apparent, but with an offset. Therefore, in Fig. 16b, only the inclusions categorized as thick according to ISO 4967, method A are compared to SEP 1571, method K. It can be seen that the tendencies are more similar, but differences are still present. However, it should be borne in mind that the database for this comparison is limited.

9.5 Comparison of extreme value methods according to SEP 1571, method E and ASTM E2283-08

Both standards use different approaches to determine a maximum feature:

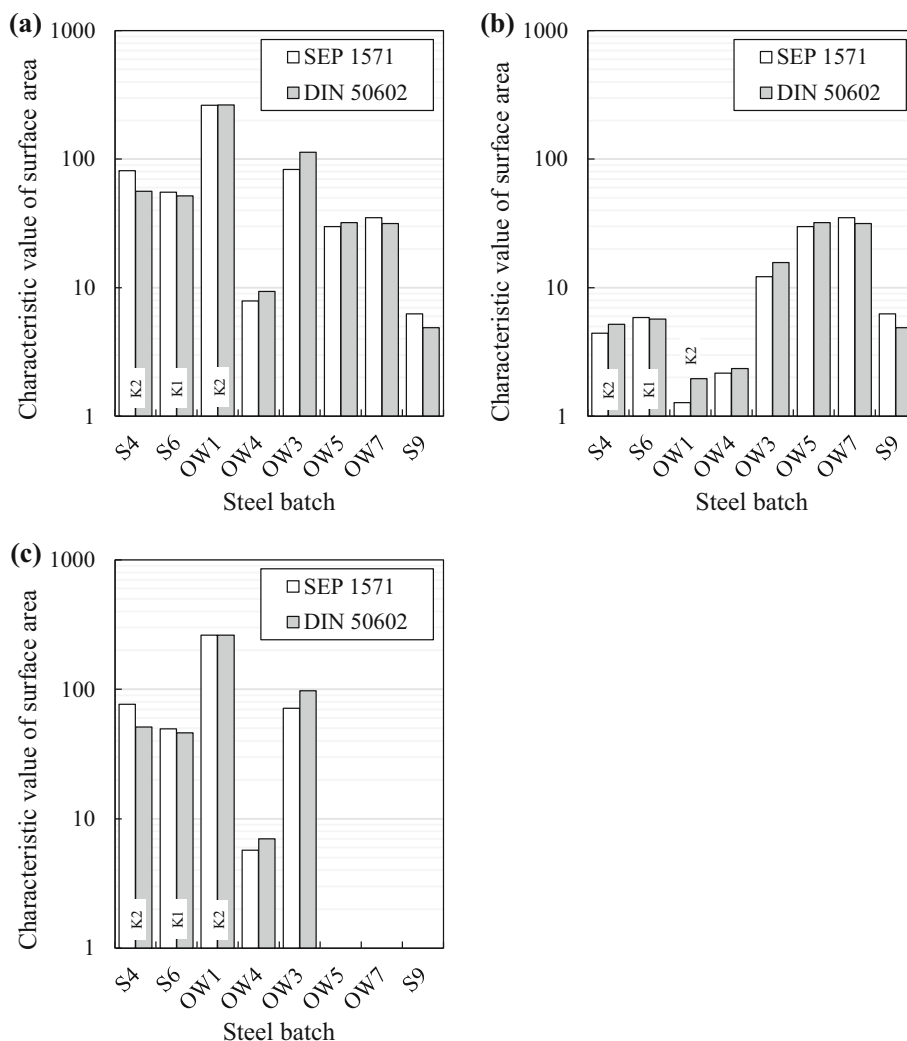


Fig. 13 Overall total characteristic value according to SEP 1571, method K, starting from size class 0 and from size class 1 (K1) or 2 (K2) (a) and total characteristic value for oxide (b) and sulfide (c) inclusions using method K of SEP 1571

1. ASTM E2283-08 determines a **maximum inclusion length** using a two-parameter (Gumbel) extreme value distribution. The method of moments or the method of maximum likelihood is given for estimating the extreme value distribution parameters. The stated 95% confidence interval for the maximum inclusion length is based on a probability of 99.9%.
2. Method E of SEP 1571 determines a **maximum inclusion size** class or maximum inclusion area using the method of maximum likelihood. The 95% confidence intervals are calculated by the Workman-Hotelling method [16].

Direct comparison is therefore not possible. However, a check is conducted in the following to ascertain whether at least the same tendencies are visible for inclusion types A and D of steel batches OW3 and OW4 as examples. Figure 17 compares the two standards. According to ASTM

E2283-8, inclusion type A shows higher values than inclusion type D for both steel batches. With SEP 1571, method E, the opposite tendency is apparent. However, both standards show higher values for steel batch OW3 than for steel batch OW4.

9.6 Comparison of results from laboratories

This publication applied the results from six laboratories. All laboratories used separate samples, which must be taken in consideration in the comparison of the laboratories. This approach was chosen to broaden the database and to consider the inhomogeneity of the material. The following compares the results of the laboratories, and steel batches OW3 and OW4 are used as an example. Table 7 gives an overview of the number of microsections from each laboratory evaluated in this subsection.

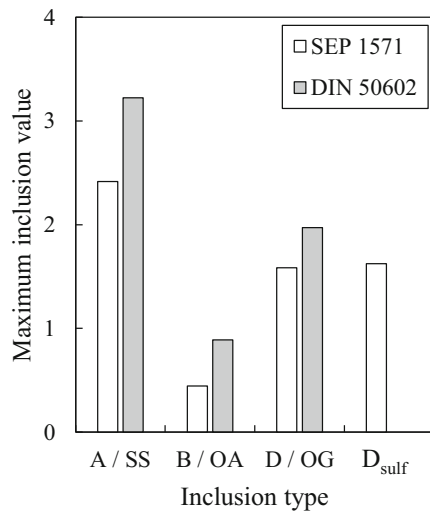


Fig. 14 Maximum inclusion value according to SEP 1571, method M, using OW3 as an example (note that there is no equivalent of D_{sulf} in DIN 50602)

Figures 18 and 19 show the characteristic value for oxide and sulfide inclusions according to method K of SEP 1571 and according to DIN 50602, respectively. It can be seen that the values from each laboratory differ slightly. Starting from size class 4 almost no differences can be seen between the laboratories in both steel batches, whereas starting from size class 0 some differences are visible. However, bearing in mind that the ultra-clean material is inhomogeneous, and different microsections were used, the values are all still within the same range.

10 Discussion of results and recommendations

A high degree of cleanliness has been achieved for all variants. The individual measures taken in the steel production process appear to be effective and result in ultra-clean gear steels. Steel batches OW7 and S9 show the highest degree of cleanliness. Both steel batches had a modified calcium treatment with additional recrystallization annealing. Furthermore, both steel batches were NiCr, and NiMo alloyed instead of MnCr or CrNiMo alloyed. However, the influence of the higher degree of cleanliness in combination with these both alloy systems on the tooth root bending strength is still to be investigated.

Normally, the location of the evaluation area for the examination of the degree of cleanliness is specified in the steel mill. This can vary depending on the dimension and manufacturing route of the steel batch. However, for the correlation of the tooth root bending strength with the degree of cleanliness it is crucial that the local degree of cleanliness is determined. It seems beneficial to determine

the degree of cleanliness in the later most stressed area. Therefore, ROI for the examinations of the degree of cleanliness was specified according to the region of the later tooth root fillet of the gear.

A determination of the macroscopic degree of cleanliness is required according to ISO 6336, part 5 for wrought steels according to material quality ME. All gear steels in this publication are classified according to this material quality. Therefore, an ultrasonic immersion testing was performed. Only in the specimens of steel batches OW3 and OW4, macroscopic inclusions could be detected by ultrasonic immersion testing using the chosen parameters. An ultrasonic immersion testing seems not to be an appropriate method for differentiating the microscopic degree of cleanliness of ultra-clean gear steels. However, an ultrasonic immersion testing should always be performed for gear steels to ensure that no macroscopic inclusion is present in the steel batch and is even required for material qualities MQ and ME according to ISO 6336, part 5.

10.1 SEP 1571

SEP 1571 is used here as an example standard for the degree of cleanliness. It is assumed that the derived conclusions can also be applied in full or at least in large part to other standards, such as ISO 4967.

The investigations showed for SEP 1571, method K that starting from size class 4 is not suitable for comparing ultra-clean steels. Starting from size class 4, no differentiation between ultra-clean gear steels can be made. This is because smaller inclusions are present in ultra-clean gear steels than in common gear steels. The examination of ultra-clean steels should start (at least) at size class 1, although starting from size class 0 seems beneficial. It is also possible to start from size classes below zero, but the increased effort needed to determine the values might be impractical in industrial practice; however, they should be considered in research projects, where the degree of cleanliness is a key research topic. In addition to the overall total characteristic value, the total characteristic value for oxide and sulfide inclusions should be stated as well. Especially, if the degree of cleanliness is determined by different inclusion types.

In an evaluation of ultra-clean gear steels according to SEP 1571, method K, no noteworthy difference is visible even with a high number of samples, starting from size class 4. However, from size class 0, a not negligible influence of the number of samples is apparent. For sulfide inclusions, values that are more reliable only appear from eight samples onward and for oxide inclusions only from six samples onward. This is because the non-metallic inclusions are inhomogeneously distributed in terms of

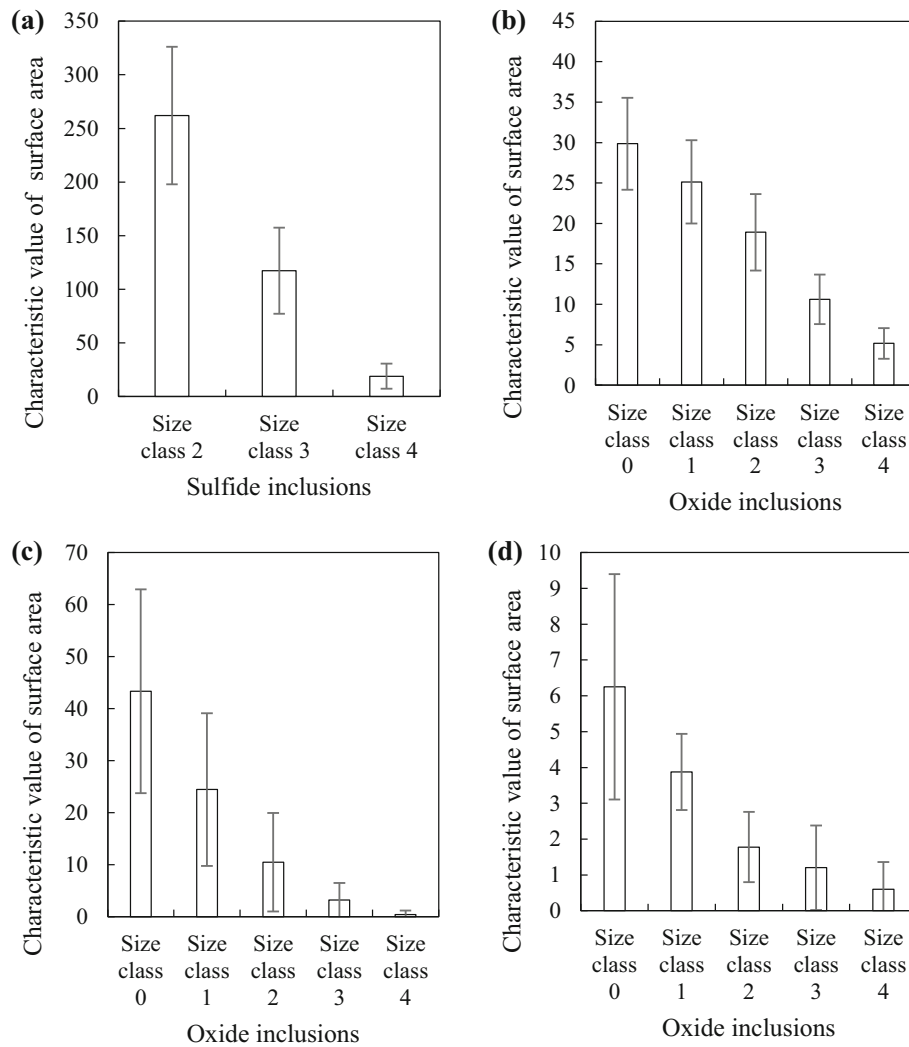


Fig. 15 Mean value and standard deviation of characteristic value of surface area according to SEP 1571, method K for steel batch OW1 (MnCr-alloyed) (a), steel batch OW5 (CrNiMo-alloyed) (b), steel batch OW7 (NiMo-alloyed) (c) and steel batch S9 (NiCr-alloyed) (d)

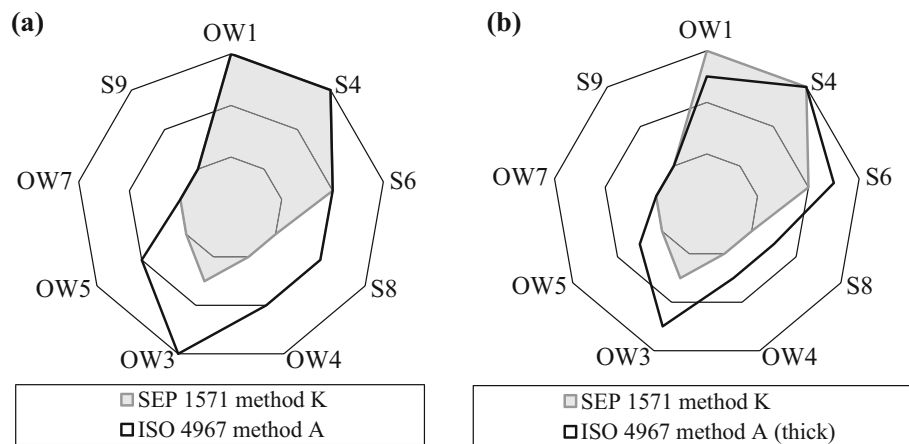


Fig. 16 Comparison of tendencies of SEP 1571, method K and ISO 4967, method A (a) and SEP 1571, method K and only inclusion categorized as thick according to ISO 4967, method A (b)

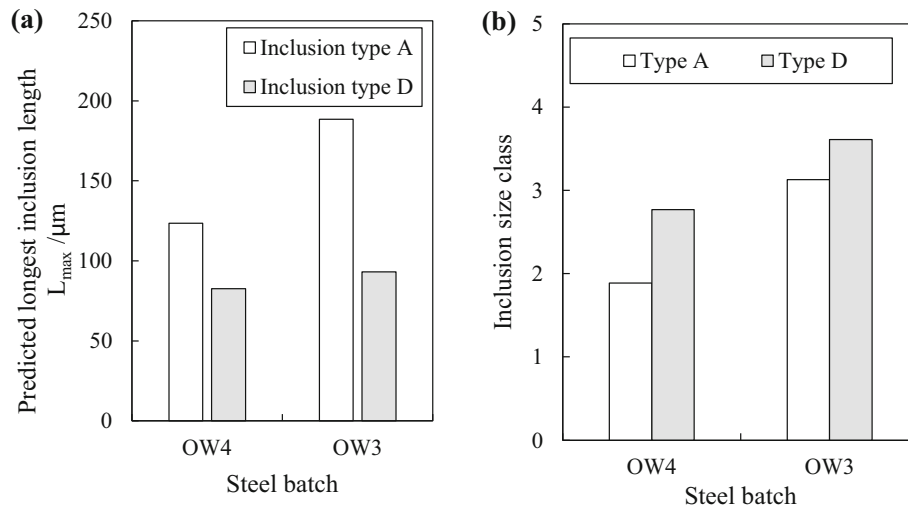


Fig. 17 Comparison of extreme value methods ASTM E2283 (a) and SEP 1571, method E (b)

Table 7 Overview of number of microsections evaluated from each laboratory

Steel batches	Laboratory					
	A	B	C	D	E	F
OW3	20	8	24	18	6	6
OW4	14	6	24	18	6	6

their size and location in the steel volume. In addition, a minimum of six microsections seems beneficial for each sample, and each microsection should be (at least) $200 \mu\text{m}^2$.

There is no strong influence of the number of samples on the maximum inclusion size according to SEP 1571, method M, if evaluating six microsections for one sample. However, it seems beneficial that (at least) two samples ($\cong 12$ microsections) are evaluated to determine the maximum inclusion size. For a comparison of ultra-clean steel batches, method E of SEP 1571 represents a more conservative approach than method M. Furthermore, a stacked bar chart facilitates comparison of ultra-clean steels according to method E.

10.2 Further investigations

It was shown that the determination of the degree of cleanliness of these ultra-clean gear steels should be based on six or more samples. However, this approach is not expedient in industrial practice. Four samples ($\cong 24$ microsections) are usually examined for SEP 1571, method E. When limiting the number of microsections to a maximum of 24, to limit time and costs, it seems beneficial that at least three samples ($\cong 18$ microsections) are examined

to determine the characteristic value of the surface area according to SEP 1571, method K. However, 24 microsections seems more beneficial.

The standard SEP 1571 is a value-compatible successor to DIN 50602 also for ultra-clean gear steels. There is a good correlation between the values obtained with method K of SEP 1571 and DIN 50602. Some slight differences can be seen when comparing these methods. However, except for the inclusion type A/SS, the same maximum inclusion size class is always given.

The already drawn conclusions are also valid for other steel alloy systems and slight differences in non-metallic inclusion content.

It is not possible to make a direct comparison of the values of SEP 1571, method K and ISO 4967, method A. However, it seems that a comparison of SEP 1571, method K and only the inclusion categorized as thick according to ISO 4967, method A may display a similar tendency. However, this has to be verified in further investigations. The overall tendency is similar when comparing the steel batches with ASTM E2283-08 and SEP 1571, method E. However, differences are apparent when comparing the inclusion types.

Comparing the results of different laboratories shows that the values from each laboratory differ slightly.

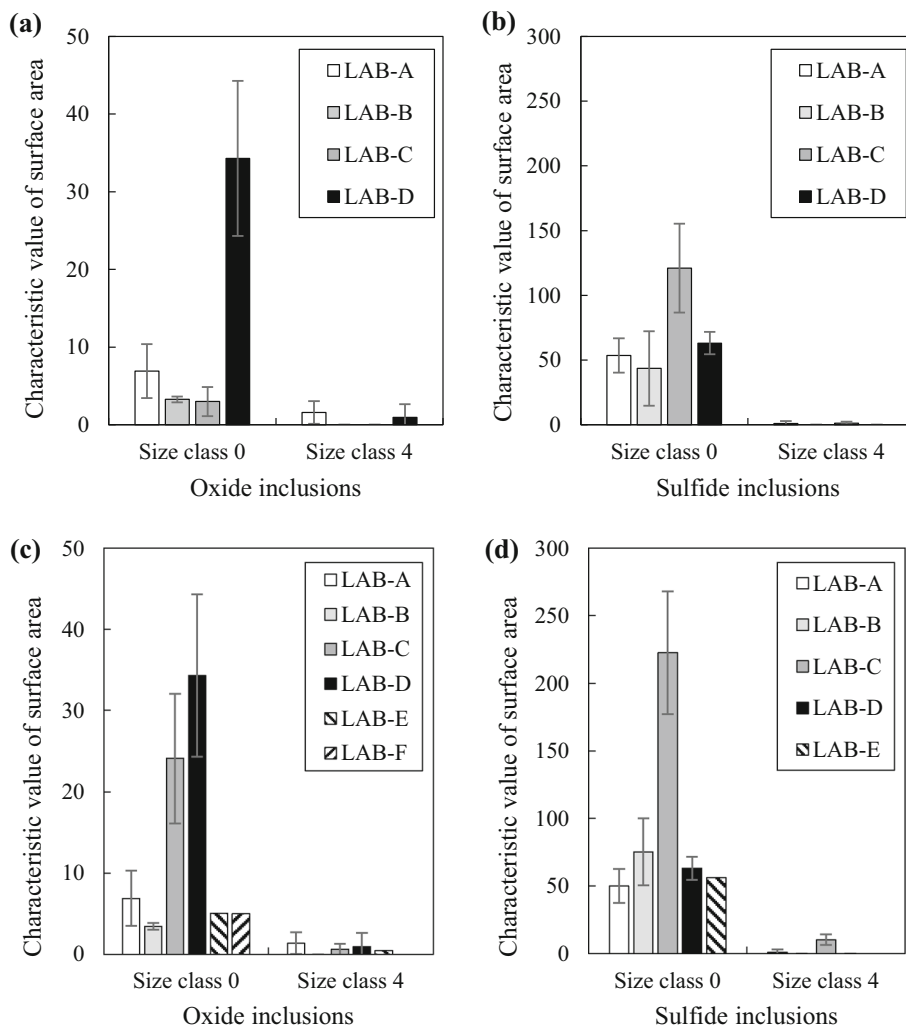


Fig. 18 Characteristic value according to method K of SEP 1571 of oxide (a) and sulfide (b) inclusions, and according to method K of DIN 50602 of oxide (c) and sulfide (d) inclusions of steel batch OW3 and comparison of laboratories

However, bearing in mind that the ultra-clean material is inhomogeneous and different microsections were used, the values are all still within the same range. It can be stated that the six laboratories, which carried out the cleanliness studies, achieved comparable results. Whether this also applies to other laboratories must be examined more in detail in each individual case. As a conclusion for industrial practice, it can be stated that no distinction should be drawn between steel batches with cleanliness values in the same range due to the inhomogeneity of ultra-clean gear steels. However, in further investigations, limit values should be elaborated to allow better comparison and differentiation of gear steels.

10.3 Recommendations

For industrial application of these results, it is important in the gear industry to make recommendations for direct

application and further improvement. A possible approach to characterize ultra-clean gear steels, based on the results presented here, would be as follows:

1. ROI should be specified according to the region of the later tooth root fillet of the gear.
2. Ultrasonic immersion testing according to SEP 1927 should be used to ensure that no macroscopic inclusions are present.
3. Determination of the microscopic degree of cleanliness should be performed according to SEP 1571, method K:
 - a. 4 single samples (\cong 24 single microsections)
 - b. Area of (at least) $200 \mu\text{m}^2$ for each microsection
 - c. Starting from size class 0
 - d. Separate statement of oxide and sulfide inclusions
 - e. Statement of standard deviation

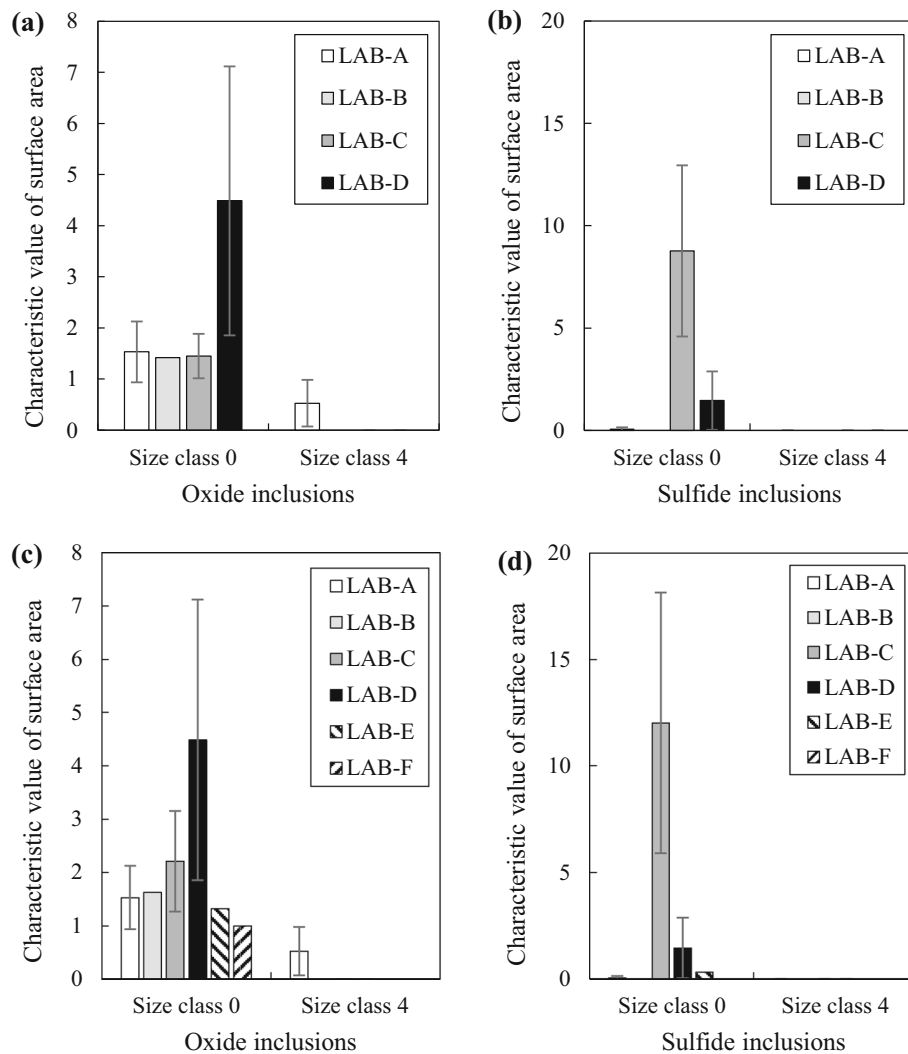


Fig. 19 Characteristic value according to method K of SEP 1571 of oxide (a) and sulfide (b) inclusions, and according to method K of DIN 50602 of oxide (c) and sulfide (d) inclusions of steel batch OW4 and comparison of laboratories

- Determination of the extreme value should be in accordance with SEP 1571, method E.

Should there be any irregularities in the characterization, repeated grinding and polishing of the microsections to gain 48 microsections in total is recommended.

11 Conclusions

- A high degree of cleanliness has been achieved for all steel batches investigated with the measures taken in the steel production process. All steel batches can be classified in the category “ultra-clean gear steels”.
- An examination of 24 microsections starting from size class 0 seems beneficial (SEP 1571, method K) to get more reliable and comparable results of the degree of cleanliness.

- Should there be any irregularities in the characterization, repeated grinding and polishing of the microsections to gain 48 microsections in total is a simple way to expand the database.

Acknowledgements The underlying research work was funded in equal proportions by the “Arbeitsgemeinschaft industrieller Forschungsvereinigungen e.V. (AiF)”, the German Federal Ministry of Economics and Technology (BMWi, IGF no. 16662 N) and the FVA. The results presented in this paper were taken from the FVA research project 293 III “Späte Zahnfußbrüche/ Reinheitsgrad”. More detailed information on the influence of non-metallic inclusions is given in the final report.

Funding Open Access funding enabled and organized by Projekt DEAL. The underlying research work was funded in equal proportions by the “Forschungsvereinigung der Arbeitsgemeinschaft der Eisen und Metall verarbeitenden Industrie e.V. (AVIF)” (A 305 / S 0024/10235/16) and the FVA. The results presented in this paper were taken from the FVA research project 293 IV “Späte Zahnfußbrüche/

Reinheitsgrad II". More detailed information on the influence of non-metallic inclusions, especially in the very high cycle range, is given in the final report.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- [1] A. Tridello, D. Paolino, G. Chiandussi, M. Rossetto, *Key Eng. Mater.* 665 (2015) 49–52.
- [2] S. Schurer, Influence of non-metallic inclusions in ultra-clean materials on the tooth root load-carrying capacity, Technical University of Munich, Munich, Germany, 2016.
- [3] A. Melander, M. Rolfsson, A. Nordgren, B. Jansson, H. Hedberg, T. Lund, *Scand. J. Metall.* 20 (1991) 229–244.
- [4] A. Tridello, D. Paolino, G. Chiandussi, M. Rossetto, *Procedia Structural Integrity* 2 (2016) 1117–1124.
- [5] R. Werl, G. Klösch, W. Winkler, A. Pissenberger, M.W. Egger, S. Aigner, J. Pühringer, S. Michelic, C. Bernhard, W. Schützenhöfer, R. Schneider, C. Schüller, *Berg Huettenmaenn Monatsh* 157 (2012) 194–203.
- [6] C. Temmel, B. Karlsson, K. Torresvoll, *Practical Metall.* 46 (2009) 123–136.
- [7] Y. Murakami, *Metal fatigue*, Academic Press, New York, USA, 2019.
- [8] D. Fuchs, T. Tobie, K. Stahl, Investigations into the defect failure of gears and its influence on the gear load carrying capacity, Forschungsvereinigung Antriebstechnik e.V. (FVA), Frankfurt, Germany, 2021.
- [9] N. Bretl, Influences on the tooth root load carrying capacity of case-hardened gears in the range of high load cycles, Technical University of Munich, Munich, Germany, 2010.
- [10] A. Stenico, material-mechanical investigations of the tooth root load carrying capacity of case-hardened gears Technical University of Munich, Munich, Germany, 2007.
- [11] D. Fuchs, S. Schurer, T. Tobie, K. Stahl, *Forschung im Ingenieurwesen* 83 (2019) 579–587.
- [12] N. Bretl, S. Schurer, T. Tobie, K. Stahl, B.R. Höhn, Investigations on tooth root bending strength of case hardened gears in the range of high cycle fatigue, in: American Gear Manufacturers Association Fall Technical Meeting, 2013, pp. 52–59.
- [13] K.J. Winkler, S. Schurer, T. Tobie, K. Stahl, *Proceed. Inst. Mech. Eng. Part C: J. Mech. Eng. Sci.* 233 (2019) 7338–7349.
- [14] D. Fuchs, S. Schurer, T. Tobie, K. Stahl, On the determination of the bending fatigue strength in and above the very high cycle fatigue regime of shot-peened gears, *Forschung im Ingenieurwesen*, 2021. <https://doi.org/10.1007/s10010-021-00499-2>.
- [15] S. Schurer, T. Tobie, K. Stahl, Load carrying capacity increase in the tooth root due to ultra-clean steels, Frankfurt, Germany, 2021.
- [16] E. Mullins, *Statistics for the quality control chemistry laboratory*, Royal Society of Chemistry, Cambridge, UK, 2003.