



# Gear root bending strength: statistical treatment of Single Tooth Bending Fatigue tests results

## Conference Proceedings

Luca Bonaiti<sup>1</sup> · Francesco Rosa<sup>1</sup> · Prasad Mahendra Rao<sup>1</sup> · Franco Concli<sup>2</sup> · Carlo Gorla<sup>1</sup>

Received: 26 March 2021 / Accepted: 25 October 2021 / Published online: 15 November 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature 2021

### Abstract

Gear tooth failure due to tooth root bending fatigue is one of the most dangerous failure modes in gears. Therefore, the precise definition of gear bending fatigue strength is a key aspect. Furthermore, in order to calculate the service life of a gear component, an accurate estimation of the S-N curve is required. This curve must properly account for the slope of the fatigue strength region, the slope of the region ahead the fatigue knee, as well as the position of the knee itself. Hence, in order obtain a curve at different reliability levels, a proper estimation of the dispersion associated with the experimental points is required. Usually, Single Tooth Bending Fatigue (STBF) tests are used to investigate the gear load carrying capacity with respect to the bending failure mode. Then, starting from the gear specimen test data, the S-N curve that has to be used in the rating method has to be determined. Maximum Likelihood Estimation (MLE), a statistical tool capable of considering also interrupted tests (e.g. runouts) has been used to estimate, in the most reliable way, the S-N curve from experimental points. Furthermore, if the STBF test is performed in a symmetrical configuration, i.e. two teeth loaded at the same time, also the survival of one of the two teeth represents an information that can be taken into account thanks to MLE.

Moreover, additional effects must be considered to obtain the S-N curve of a real component starting from STBF tests. In reality, the load history and the statistical behaviour are different, since in a meshing gear the strength is determined by its weakest tooth, while in a STBF test the failing tooth is predetermined. The latter effect is considered by means of the statistic of extremes, which enables the estimation of the strength of the weakest tooth and therefore of the gear. This paper describes in detail the proposed calculation method and explains its application to determine the S-N curve in practical cases.

---

✉ Luca Bonaiti  
luca.bonaiti@polimi.it

<sup>1</sup> Department of Mechanical Engineering, Politecnico di Milano, Milan, Italy

<sup>2</sup> Faculty of Science and Technology, Free University of Bolzano, Bolzano, Italy

## Zahnfußbiegefestigkeit: Statistische Behandlung der Ergebnisse von den STBF-Versuchen

### Zusammenfassung

Versagen von Verzahnungen aufgrund von Zahnfußbiegeermüdung ist eine der gefährlichsten Versagensarten von Zahnradern. Daher ist die genaue Definition der Biegewechselfestigkeit von Getrieben ein wichtiger Aspekt. Zur Berechnung der Lebensdauer einer Komponente eines Getriebes ist die genaue Schätzung der Wöhlerkurve erforderlich. Diese Kurve muss die Neigung des Ermüdungsfestigkeitsbereichs, die Neigung des Bereichs der Ermüdungsknie sowie die Position des Knies selbst richtig berücksichtigen. Um eine Kurve mit unterschiedlichen Zuverlässigkeitsniveaus zu erhalten, ist daher eine zuverlässige Schätzung der Streuung in Verbindung mit den experimentellen Punkten erforderlich. Üblicherweise werden Single-Tooth Bending-Fatigue (STBF)-Versuche verwendet, um die Getriebetragefähigkeit in Bezug auf den Biegeversagensmodus zu untersuchen. Ausgehend von den Prüfdaten des Zahnradmusters muss dann die Wöhlerkurve bestimmt werden, die im Bewertungsverfahren verwendet werden muss. Die Maximum-Likelihood-Schätzung (MLE), ein statistisches Werkzeug, das auch unterbrochene Tests verwaltet: wird der STBF-Versuch in einer symmetrischen Konfiguration durchgeführt, d. h. zwei Zähne gleichzeitig belastet, stellt auch das Überleben eines der beiden Zähne eine Information dar, die dank der MLE berücksichtigt werden kann.

Darüber hinaus müssen zusätzliche Effekte berücksichtigt werden, um ausgehend von STBF-Versuche die Wöhlerlinie eines realen Bauteils zu erhalten. In Wirklichkeit sind die Belastungshistorie und das statistische Verhalten unterschiedlich, da bei einem Zahnrad die Festigkeit durch den schwächsten Zahn bestimmt wird, während bei einem STBF-Versuch der ausfallende Zahn vorgegeben wird. Letzterer Effekt wird durch die Extremwertstatistik berücksichtigt, die eine Abschätzung der Festigkeit des schwächsten Zahnes und damit des Zahnrades ermöglicht. Dieses Veröffentlichen beschreibt ausführlich das vorgeschlagene Berechnungsverfahren und erläutert seine Anwendung zur Bestimmung der Wöhlerlinie in praktischen Fällen.

### 1 Introduction

Tooth root bending fatigue is considered as one of the most dangerous gear failure mode because, typically, such kind of a failure results in the interruption of the power flow within the gearbox. Standards, such as ISO 6336-3 [1] and ANSI/AGMA 2001 [2], provide designers with an analytical framework, to assess a gear train, in respect to this failure mode, by comparing the maximum tooth root stress with the limit value of the gear itself. Data of several typical materials are present within the standards (e.g. [2, 3]). Both ISO 6336-3 [1] and ANSI/AGMA 2001 [2] present indication about the complete S-N curve, with the slopes associated to both, the limited life region and the long life one. Those curves are defined at 1% gear failure probability (or 99% reliability) for tooth root bending fatigue. ANSI/AGMA 2001 [2] presents corrective coefficients to “shift” the S-N curve so that it assures a certain reliability level, different from the typical. Designers working with the ISO 6336 series standard can rely on the work of Hein et al. [4] where a corrective coefficient  $Y_Z$  has been proposed to assess tooth root bending fatigue at different reliability levels; a coefficient  $Z_Z$  has been proposed for the phenomenon of pitting.

However, assuring a certain reliability requires two different things. On the one hand, it requires a proper estimation of the load; topic covered by standards such as ISO 6336-6 [5]. On the other hand, it also requires an accurate definition of the complete gear S-N curve, both in term of

the slope of the limited life region and the long life region, as well as the associated dispersions [6]. Here, it would be better to rely on data specific to the current industrial scenario rather than relying on standardized data as they are not representative of the gears investigated. Therefore, extensive gear testing is needed even more when it is required to assure a certain reliability level after a given service life.

The available testing methodologies for the tooth root bending fatigue strength are running gears tests (e.g. [7, 8]), Single Tooth Bending Fatigue (STBF) tests (e.g. [9–15]), and tests on notched specimens (e.g. ISO 6336-3 annex A [1, 16–18]). Amongst them, due to its higher efficiency and capability to directly work with gears, STBF tests are the most adopted.

However, STBF tests results cannot be directly applied to assess a gear pair. Indeed, as there is no rolling/sliding contact between gears, therefore, the loading condition of an STBF test rig is not representative of the real case.

Two different approaches for the treatment of STBF test data exist, so that they can be used within the existing calculation methods. They are proposed within FVA report no. 304 [9] and within the works of Rao and Mc Pherson [8, 11]. Both the authors highlight two main issues that must be considered in order to treat STBF test data:

1. The testing and service loading conditions are different. In the real case, the force varies both, the application point, and its magnitude, while in STBF the force is a si-

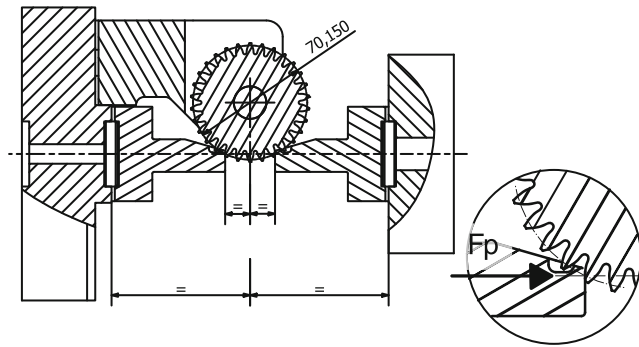


Fig. 1 STBF apparatus

- nusoidal load whit a fixed amplitude and a given application point.
- The testing and service failures are different. In STBF test, a failed test corresponds to the breakage of some tested teeth, while in the meshing case, the failure of the specimen (i.e. the gear itself) corresponds to the failure of the weakest tooth.

While a paper by the authors investigates the effect of different loading conditions by means of an high cycle multiaxial fatigue criteria [19], this paper aims to shed more light on this latter aspect.

Stahl and Rao faced the problem with two different approaches. On the one hand, Stahl, K [9]. proposed several coefficients, based on experimental evidence: two coefficients (i.e.  $f_{p \rightarrow Z}$  and  $f_{Z \rightarrow ZR}$ ) to translate STBF results to the case of a meshing gear, and a further one (i.e.  $f_{1\%ZR}$ ) to reach the desired reliability level, which in this case is  $R=99\%$  (i.e. a probability of failure of 1%); a further coefficient  $f_{kor}$  based on Rettig experimental evidence [20] is used to deal with the different loading conditions. On the other hand, Rao and Mc Pherson [8, 11] define a simple statistical relation between the teeth and the gear, and by means of a ‘fit by eye’, they translate STBF test data, corrected to the real case load, to the desired reliability level. Allowable Stress Range (ASR) diagrams are used to deal with the different loading conditions.

In order to reduce the need of predetermined experimental factors, an approach based on the Maximum Likelihood Estimation (MLE) and the statistic of extremes is proposed in this paper. MLE is an estimation technique typically applied in the determination of the S-N curve, as it has great advantages of being able to consider, at the same time, different typology of data. Exact data (e.g. failures) and censored points are considered together, within the same

calculation procedure. The latter are points of which the exact value is unknown; still, it is possible to say that their exact value would be greater (right censored, e.g. runouts) than a certain value, smaller (left censored, e.g. a failure occurred before the inspection), or between two values (interval censored, e.g. a failure occurred between two inspections) [6, 21, 22]. Indeed, MLE is a flexible tool; e.g., as its calculation procedure is independent from the test sequence, MLE allows to find endurance limits also where the staircase method cannot be used optimally for some reasons. MLE has been applied for the estimation of the S-N curve of specimens (e.g. [23–26]) as well as in other components, e.g. welded structures (e.g. [27, 28]). Regarding gears, Krantz, T. [29] applied MLE for the estimation of the pitting S-N curve.

Within this work, MLE is applied to estimate the experimental S-N curve, that is the S-N curve of the teeth, and its associated dispersion. Implementing the statistic of extremes, this first curve is then rearranged, by means of a mathematical passage, to estimate the gear S-N curve. In other words, MLE is used to estimate the teeth distribution (i.e. the parent distribution), which is then used to estimate the gear distribution by defining the distribution of the weakest tooth among the  $z$  gear teeth. Different curves, at different reliability levels, can be estimated by simply calculating the corresponding percentile.

## 2 Experimental procedure

In the previous years, the authors have investigated the tooth root load carrying capacity of several types of gears via the STBF approach. Some examples include: gears for aeronautical applications [30, 31], small planetary gearboxes [32], wind turbine gearboxes [33, 34], generic applications [35] as well as additively manufactured gears [36, 37]. All the aforementioned tests have been performed on a mechanical resonance pulsator, equipped for each experimental campaign with a specific fixture for the gear specimen. Fig. 1 shows the equipment scheme adopted for the experimental campaign that has been taken as an example in this article. The equipment consists of two anvils, a fork and a centering pin. The anvils are the component that, during testing, are responsible of transmitting the load to the gear teeth and, due to the same principle involved within the Wildhaber measurements, they load the teeth in a symmetric way. That is, the contact between the anvils and the teeth occurs at the same diameter for both the spanned teeth. The last two com-

Table 1 Main gear geometry data

$z$	Mn	$\alpha_n$	b	$\beta$	$d_a$	$d_b$	$d_f$	$\rho_{fp}$	$X$
–	Mm	°	Mm	°	Mm	Mm	Mm	Mm	Mm
30	2.2	17	20	0	73	59.743	63.116	0.810	0.55

ponents (i.e. fork and centering pin) are responsible for the correct positioning of the gear between the anvils during the preliminary mounting phase. During the test, the pin is removed, and a minimum compressive load is kept on the teeth to avoid undesired displacements; all the tests are performed at a load ratio  $R=0.1$ . The experimental procedure is described in detail in the aforementioned papers (i.e. [30–37]).

The experimental campaign taken as example for the statistical elaboration of STBF test results has been performed on case hardened gears made of 20MnCr5 (additional information can be found in [38]); no mechanical surface treatment has been applied. A total number of 34 experimental points have been collected, 9 runouts at 5 million cycles have been observed. Table 1 summarizes the spur gear main data. If the results of the aforementioned experimental campaign have to be used to design/assess a gear pair with a different module, the size factor  $Y_x$  must be taken into account.  $Y_x$  typical values can be found in [1, 39, 40].

Method B of ISO 6336-3 [1] has been used to calculate the relationship between the pulsator applied load,  $F_p$  and the tooth root bending stress. As the standards always consider two meshing gears, the force/stress relationship has been calculated starting from a gear pair, defined in a way such as, for the gear specimen, the outer point of single contact is the pulsator point of load application. In other words, the outer single contact diameter  $d_{en}$  corresponds to the diameter at which the designed anvil/teeth contact occurs (as shown in Fig. 1,  $d_{en} = 70.150\text{ mm}$ ). For the sake of brevity, details and brief information on the whole procedure can be found in the papers cited at the beginning of this section (i.e. [30–37]). A correction factor  $f_{\text{korrr}} = 0.9$  has been used to translate stress data from the STBF test loading condition to the one of the meshing gear case [9, 20].

### 3 Maximum Likelihood: estimation of the tooth S-N curve

Several models have been developed to describe the S-N curve of a specimen (or of a component). They have different levels of complexity, from the more common ones, such as the ones described in [23, 24], to the more complex models, such as the decoupled model [25] and the random fatigue limit model [26]. Due to its simplicity, the model proposed by Spindel et al. [23] has been selected. As per this model, the S-N curve can be described as:

$$\log\left(\frac{N}{N_e}\right) = \frac{1}{2}(k + k_1) \log\left(\frac{S}{S_e}\right) + \frac{1}{2}(k - k_1) \left| \log\left(\frac{S}{S_e}\right) \right| \tag{1}$$

where  $N$  is the number of cycles,  $S$  is the applied load,  $N_e$  and  $S_e$  represent the knee position,  $k$  is the slope of the limited life region and  $k_1$  is the slope of the long-life region. An easy way to fit Eq. 1 to the experimental data is to consider  $S$ , or directly  $\log(S)$ , as the independent variable and  $N$ , or  $\log(N)$ , as the dependent variable and then fit the data by using rearranging Eq. 1 as:

$$\log(N) = \log(N_e) + \frac{1}{2}(k + k_1) (\log(S) - \log(S_e)) + \frac{1}{2}(k - k_1) |\log(S) - \log(S_e)| \tag{2}$$

However, to fit such a curve to experimental data, runouts must be considered as a point with the same meaning of a failure. That is, we are losing the information that such experimental points do not represent a failure, but a point which has survived to  $N_i$  cycle under a certain load  $S_i$ . To avoid loss of such information, the parameters of Eq. 1 can be estimated via MLE. For this, it is assumed that experimental points are log-normally distributed, that is the logarithmic to the base 10 of the points follow a normal distribution, of which the PDF  $f(x)$  and CDF  $F(x)$  are described as:

$$\begin{cases} f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \\ F(x) = \int_{-\infty}^x f(x)dx = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ = \frac{1}{2} \left( 1 + \text{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right) \right) \end{cases} \tag{3}$$

where  $\text{erf}$  is the error function.

MLE is based on finding the distribution parameters, in this case  $\mu$  and  $\sigma$ , which are more likely to represent the experimental points. In other words, MLE estimates the distribution parameters that maximize the likelihood  $\mathcal{L}$  [6].

The case examined in this paper presents two kinds of data. The first ones are observed values, that are represented by specimens whose lifetime, at a certain load level, is known (i.e. a failure). The second ones are right-side censored data, that are represented by points whose lifetime, at a certain load level, is greater than a certain level, because the specimen was not broken (f.i. a runout). Therefore,  $\mathcal{L}$  can be defined as:

$$\mathcal{L} = \prod_{i=1}^n (f_{x_i;\mu,\sigma} dx)^{\delta_i} \cdot \prod_{i=1}^n (1 - F_{x_i;\mu,\sigma})^{1-\delta_i} \tag{4}$$

where  $\delta_i$  is a constant equal to 1 when the  $i^{\text{th}}$  observation is an observed value, and null in the case of a right-side censored data. To simplify the estimation procedure, the natural logarithm of  $\mathcal{L}$  is considered:

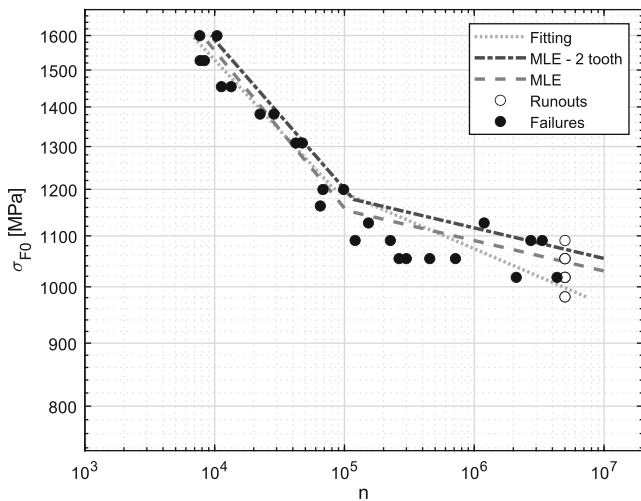


Fig. 2 Teeth S-N curve estimation comparison

$$\begin{aligned} \uparrow = & \sum_{i=1}^n \delta_i (\ln f_{x_i}) + \sum_{i=1}^n \delta_i n \ln (d_x) \\ & + \sum_{i=1}^n (1 - \delta_i) (\ln (1 - F_{x_i})) \end{aligned} \quad (5)$$

where  $\ell$  is the so-called log-likelihood,  $\ln f_{x_i}$  represents exact data (failures), and  $\ln (1 - F_{x_i})$  right censored data (runouts).  $\sum_{i=1}^n \delta_i n \ln (d_x)$  is no more considered as it is a constant term and, therefore, it does not affect the maximization procedure. Anyhow, as only minimization procedures are included in the usual calculation software, distribution parameters are estimated by minimizing  $-\ell$ .

As proposed by Nelson [22], it is possible to use MLE for curve fitting, independently of the chosen PDF, by considering the location parameter as a function of the independent variable; also dependency of the scale parameter on the independent variable can be considered too. In our case, we set the mean value  $\mu$  equal to Eq. 2. In [23], Spindel et al. propose a constant standard deviation in the  $\log (\sigma_{F_0})$  direction. Considering that the developed model uses  $N$  as the dependent variable, and that the model presents 2 different slopes, the standard deviation in the  $\log (\sigma_{F_0})$  direction is “rotated” in the  $\log (N)$  direction. Therefore, two different standard deviations in the  $\log (N)$  directions (i.e.,  $\sigma_{1, \log (N)}$  and  $\sigma_{2, \log (N)}$ ) are obtained by multiplying  $\sigma_{\log (\sigma_{F_0})}$  by the absolute value of  $k$  (or  $k_1$ ).

However, one information obtained from the test has not been yet considered. As described in Sect. 2, the adopted STBF test configuration is symmetric. That is, the tested teeth are loaded at the same diameter, sharing the same applied load, resulting in the same nominal tooth root bending stress. Indeed, each experimental point (i.e., failure or runout) ‘hides’ an information about the survival of the other loaded tooth. In other words, due to the test configuration, each test interrupted by the failure of one of the two teeth is a failure and a survival at that number of cycles,

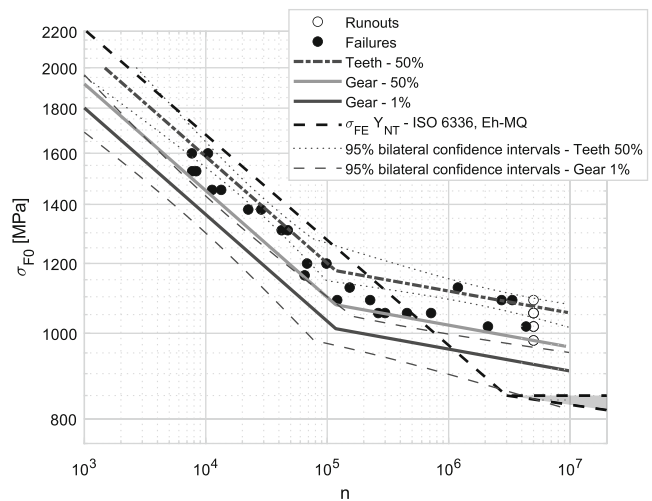


Fig. 3 Estimated gear S-N curve

at which the failure of the weakest tooth (among the tested pair) occurs. On the other hand, the case of a runout represents two runouts. In the case of a failure, an observed value and a right-side censored data are present (instead of a single observed value), while, in the case of a runout, two right-side censored data are present. The case of a test stopped due to the failure of both the loaded teeth will simply be considered as the failure of two teeth. Anyhow, as it implies that both teeth have the same lifetime, it is a very rare case, so it is not discussed in the below mathematical development. Other authors, like Wallin, K. [27] and Marquis, G. and Mikkola, T. [28], deal with the same problem of this topic in the field of welded joints. As in the adopted STBF test configuration, more than one possible failure point is present in their tested welded structure. Also, those authors solved the issue of having several specimens under testing at the same time and under the same load by adopting similar considerations. To consider this information, a further term, representing right censored data, is added to Eq. 5:

$$\begin{aligned} \uparrow = & \sum_{i=1}^n \delta_i (\ln f_{x_i}) + \sum_{i=1}^n (1 - \delta_i) (\ln (1 - F_{x_i})) \\ & + \sum_{i=1}^n (\ln (1 - F_{x_i})) \end{aligned} \quad (6)$$

in this way, it is possible to consider that a right-censored term (the “second” tooth) is always present.

Fig. 2 shows a comparison of the three different procedures to estimate the S-N curve. The curve obtained by curve fitting (i.e., fitting) and those estimated by mean of MLE as in Eq. 5 (i.e., MLE and MLE—2 teeth), in the region before the knee, estimate a similar slope; indeed, both are estimating a line intercepting a comparable set of points. However, as the fitting does not consider the physical meaning of runouts, the slope associated to the region ahead of the knee is not comparable. This results in a different es-



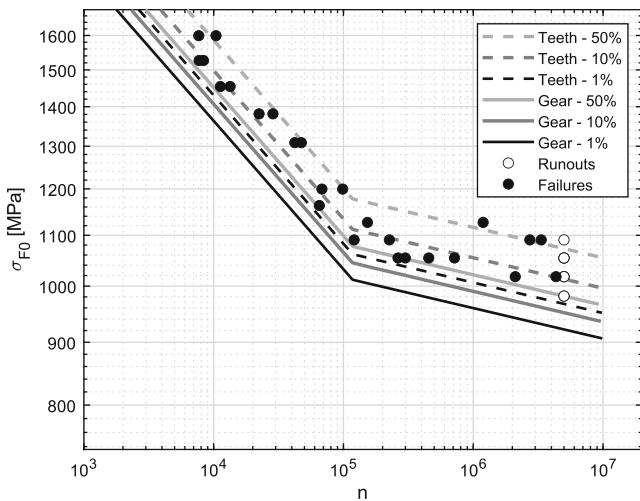


Fig. 4 Estimated S-N curves

timation of the knee. Furthermore, the curve obtained with Eq. 6 likelihood formulation results to be shifted upwards due to the consideration of the survived teeth.

#### 4 Statistic of extremes: from tooth to gear

Once the S-N curve of the teeth has been estimated (and therefore, the CDF), STBF result elaboration is necessary to make it representative of the gear. To do this, it is necessary to consider that, in the real case scenario, where gears are meshing and rotating, gear failure due to tooth root bending fatigue corresponds to the failure of the weakest gear tooth [8, 11, 20]. Therefore, starting from the estimated teeth S-N curve, the gear S-N curve is defined as the one describing the weakest tooth amongst the  $z$  gear teeth. In other words, the CDF of the teeth has to be elaborated in order to define the gear CDF, that is the CDF of the weakest tooth among the  $z$  gear teeth. It is worth underlying that in 1987, when Rettig studied the effect of the different loading conditions between STBF tests and running gears, he adopted the statistic of extreme in order to translate STBF/teeth results to the gear [20].

By means of a mathematical passage [6], statistic of extremes allows to define the CDF of the minimum value  $X_{(1)(x)}$  over  $n$  extractions of  $X$  (the distribution of which is known) as:

$$F_{X_{(1)(x)}} = 1 - (1 - F_{X(x)})^n \tag{7}$$

where  $F_{X(x)}$  is the parent CDF and  $F_{X_{(1)(x)}}$  is the extreme CDF. In other words, Eq. 7 expresses the distribution of the smallest value obtained from  $n$  extractions from a population described by the parent distribution. In the case of gear tooth root bending fatigue, we aim to estimate the distribu-

tion of the tooth with the smallest resistance (i.e., the gear CDF  $F_{gear}$ ) over  $z$  gear teeth. For the sake of simplicity, in the case of tooth root bending fatigue, Eq. 7 can be seen as:

$$F_{gear} = 1 - (1 - F_{tooth})^z \tag{8}$$

The above equation implies that the gear load carrying capacity depends on  $z$ . Thus, considering CDF/PDF properties and typical shapes, under the assumption that the teeth statistical parameters do not depend on  $z$ , Eq. 8 implies that the difference between the SN curves for two gears with different teeth numbers, is almost proportional to the logarithm of  $z$ .

According to ISO 6336-5 [3], the typical failure probability for gear failure due to tooth root bending fatigue is 1% (i.e. a reliability of 99%), hence it is required to define the curve at the required reliability level. In other words, it is required to estimate the load carrying capacity with the 99% percent probability of being exceeded of the weakest tooth over  $z$  teeth. In order to do so, the corresponding percentile is calculated for different stress levels, i.e., finding for a given stress level, the  $N$  value that allows  $F_{gear}$  to be equal to the desired failure probability.

Fig. 3 reports the estimated S-N curve at different reliability levels. The final curve (i.e., the curve at 1% failure probability for the gear) is calculated as the 1% percentile of the gear PDF. For comparison, the 50% percentile gear curve, as well as the initial 50% percentile teeth curve are included. 95% bilateral confidence intervals have been calculated using the quadratic approximation [6, 22].

It is worth underlining that the estimated fatigue knee is around 120,000 cycles, a number of cycles different than the one proposed by ISO 6336-3 (i.e., 3 million of cycles). There are two reasons behind such a difference. Firstly, the knee estimated here is simply an intersection between the two lines defined in Eq. 1; in other words, it is a curve parameter. It does not have the same meaning of the fatigue knee obtained looking also at the fatigue limit because the fatigue limit has statistical meanings that are not included within the Spindel model. Secondly, ISO 6336-3 describes general gear material properties, and it is reasonable to assume that certain gear will show a different position of the fatigue knee; indeed, [8, 9, 11, 20] show fatigue knee positions different from the one proposed by ISO.

#### 5 Results and conclusion

Even if ISO 6336-3 [1] and ISO 6336-5 [5] allow to estimate the gear tooth root bending fatigue strength by means of STBF tests, they do not give all the details to elaborate STBF test results. Here, a statistical framework through which analyse STBF test data in order to translate them to

the gear specimens, in meshing condition, has been proposed. With this methodology, the data elaboration is based only on the experimental data and does not require predefined coefficients.

Fig. 4 show the estimated percentiles for the teeth as well as the gear. As can be seen, not considering the effect of the different failure conditions (i.e. predetermined for STBF and weakest for meshing gears) implies an overestimation of the tooth root bending fatigue curve. The curve referred to the gear shows a lower carrying capacity than the one referred to the teeth. Furthermore, meshing gear percentiles, shown in Fig. 4, are closer to each other rather than those of STBF. Indeed, extreme value PDFs are narrower than the parent one [6].

## 6 Nomenclature

The nomenclature is shown in Table 2.

**Acknowledgements** The authors express their gratitude prof. Stefano Beretta for his invaluable support during the validation of this statistical model.

**Conflict of interest** L. Bonaiti, F. Rosa, P. M. Rao, F. Concli and C. Gorla declare that they have no competing interests.

## References

- ISO 6336-3:2019—Calculation of load capacity of spur and helical gears—Part 3: Calculation of tooth bending strength. .
- ANSI-Standard, “Fundamental Rating Factors and Calculation Methods for Involute Spur and Helical Gear Teeth,” Am. Gear Manuf. Assoc. Alexandria, 2004.
- ISO-6336-5, “Calculation of Load capacity of Spur and Helical Gears, Part 5: Strength and quality of materials,” Int. Stand. Organ. Geneva, 2006.
- Hein M, Geitner M, Tobie T, Stahl K, Pinnekamp B (2018) Reliability of gears—determination of statistically validated material strength numbers
- ISO 6336-6:2019, “Calculation of Load capacity of Spur and Helical Gears, Part 6: Calculation of service life under variable load,” Geneva, CH, 2019.
- Beretta S (2010) Affidabilità delle costruzioni meccaniche: Strumenti e metodi per l’affidabilità di un progetto. Springer, Berlin Heidelberg
- Hong IJ, Kahraman A, Anderson N (2020) A rotating gear test methodology for evaluation of high-cycle tooth bending fatigue lives under fully reversed and fully released loading conditions. Int J Fatigue 133:105432
- Rao SB, McPherson DR (2003) Experimental characterization of bending fatigue strength in gear teeth. Gear Technol 20(1):25–32
- Stahl K (1999) Lebensdauerstatistik. Abschlussbericht: Forschungsvorhaben Nr. 304. FVA,
- Benedetti M, Fontanari V, Höhn B-R, Oster P, Tobie T (2002) Influence of shot peening on bending tooth fatigue limit of case hardened gears. Int J Fatigue 24(11):1127–1136
- McPherson DR, Rao SB (2008) Methodology for translating single-tooth bending fatigue data to be comparable to running gear data. Gear Technol. 25:42–51

**Table 2** Nomenclature

$\alpha_n$	Normal pressure angle
$\beta$	Helix angle
$\delta_i$	MLE coefficient
$\mu$	Mean
$\rho_{fp}$	Root radius
$\sigma$	Standard deviation
$\sigma_{F0}$	Nominal tooth root stress
$b$	Facewidth
$d_a$	Tip diameter
$d_b$	Base diameter
$d_f$	Root diameter
$d_{en}$	Outer single contact diameter
$f(x)$	Probability Density Function (PDF)
$F(x)$	Cumulative Distribution Function (CDF)
$F_P$	Pulsator load
$f_{1\%ZR}$	Stahl’s corrective coefficients (from gear 50% to meshing gear 1%)
$f_{korr}$	Stahl’s corrective coefficients for load differences
$f_{p \rightarrow Z}$	Stahl’s corrective coefficient (from 50% pulsator to tooth 50%)
$f_{Z \rightarrow ZR}$	Stahl’s corrective coefficients (from tooth 50% to meshing gear 50%)
$h_{aP}^*$	Addendum coefficient
$h_{fP}^*$	Dedendum coefficient
$k_{k1}$	Fatigue model coefficients
$\mathcal{L}$	Likelihood
$l$	Log-Likelihood
$m_n$	Normal module
$N_e$	Number of cycle at Spindel’s knee position
$S_e$	Stress at Spindel’s knee position
$X$	Profile shift
$Y_x$	Size factor
$Y_Z$	Hein’s reliability factor for tooth root bending
$z$	Number of teeth
$Z_Z$	Hein’s reliability factor for pitting

- SAE (2017) J1619, single tooth gear bending fatigue test. SAE International.
- Güntner C, Tobie T, Stahl K (2017) Influences of the residual stress condition on the load-carrying capacity of case-hardened gears. Am. Gear Manuf. Assoc. AGMA Tech. Pap. 17FTM20
- Winkler KJ, Schurer S, Tobie T, Stahl K (2019) Investigations on the tooth root bending strength and the fatigue fracture characteristics of case-carburized and shot-peened gears of different sizes. Proc Inst Mech Eng Part C J Mech Eng Sci 233(21–22):7338–7349
- Koenig J, Hoja S, Tobie T, Hoffmann F, Stahl K (2019) Increasing the load carrying capacity of highly loaded gears by nitriding. MATEC Web Conf 287:2001
- Medlin D, Cornelissen B, Matlock D, Krauss G et al (1999) Effect of Thermal Treatments and Carbon Potential on Bending Fatigue Performance of SAE 4320 Gear Steel. SAE Technical Paper 1999-01-0603. <https://doi.org/10.4271/1999-01-0603>
- Spice J, Matlock D, Fett G (2002) Optimized Carburized Steel Fatigue Performance as Assessed with Gear and Modified Brugger Fatigue Tests. SAE Technical Paper 2002-01-1003. <https://doi.org/10.4271/2002-01-1003>

18. Vilela Costa L, Corrêa de Oliveira D, Wallace D et al (2020) Bending Fatigue in Low-Pressure Carbonitriding of Steel Alloys with Boron and Niobium Additions. *J. of Materi Eng and Perform* 29:3593–3602. <https://doi.org/10.1007/s11665-020-04757-2>
19. Bonaiti L, Bayoumi ABM, Concli F, Rosa F, Gorla C (2021) Gear root bending strength: a comparison between Single Tooth Bending Fatigue Tests and meshing gears. *J Mech Des.* <https://doi.org/10.1115/1.4050560>
20. Rettig H (1987) Ermittlung von Zahnfußfestigkeits-Kennwerten auf Verspannungsprüfständen und Pulsatoren—Vergleich der Prüfverfahren und gewonnenen Kennwerte. *Antriebstechnik* 26:51–55
21. Nelson WB (2003) *Applied life data analysis* vol 521. John Wiley & Sons,
22. Nelson WB (2009) *Accelerated testing: statistical models, test plans, and data analysis* vol 344. John Wiley & Sons,
23. Spindel JE, Haibach E (1981) Some considerations in the statistical determination of the shape of SN curves. In: *Statistical analysis of fatigue data* ASTM International.
24. Beretta S, Clerici P, Matteazzi S (1995) The effect of sample size on the confidence of endurance fatigue tests. *Fatigue Fract Eng Mater Struct* 18(1):129–139
25. Lorén S (2003) Fatigue limit estimated using finite lives. *Fatigue Fract Eng Mater Struct* 26(9):757–766
26. Pascual FG, Meeker WQ (1999) Estimating fatigue curves with the random fatigue-limit model. *Technometrics* 41(4):277–289
27. Wallin K (1999) The probability of success using deterministic reliability. In: *European structural integrity society*, vol 23. Elsevier, , pp 39–50
28. Marquis G, Mikkola T (2002) Analysis of welded structures with failed and non-failed welds based on maximum likelihood. *Weld World* 46(1):15–22
29. Krantz TL (2002) The influence of roughness on gear surface fatigue. Case Western Reserve University Cleveland, Ohio
30. Gasparini G, Mariani U, Gorla C, Filippini M, Rosa F (2008) Bending fatigue tests of helicopter case carburized gears: Influence of material, design and manufacturing parameters. In: *American Gear Manufacturers Association (AGMA) Fall Technical Meeting*, pp 131–142
31. Gorla C, Rosa F, Conrado E, Concli F (2017) Bending fatigue strength of case carburized and nitrided gear steels for aeronautical applications. *Int J Appl Eng Res* 12(21):11306–11322
32. Gorla C, Conrado E, Rosa F, Concli F (2018) Contact and bending fatigue behaviour of austempered ductile iron gears. *Proc Inst Mech Eng Part C J Mech Eng Sci* 232(6):998–1008
33. Gorla C, Rosa F, Conrado E, Albertini H (2014) Bending and contact fatigue strength of innovative steels for large gears. *Proc Inst Mech Eng Part C J Mech Eng Sci* 228(14):2469–2482
34. Gorla C, Rosa F, Concli F, Albertini H (2012) Bending fatigue strength of innovative gear materials for wind turbines gearboxes: effect of surface coatings. In: *ASME 2012 International Mechanical Engineering Congress and Exposition*, pp 3141–3147
35. Conrado E, Gorla C, Davoli P, Boniardi M (2017) A comparison of bending fatigue strength of carburized and nitrided gears for industrial applications. *Eng Fail Anal* 78:41–54
36. Bonaiti L, Concli F, Gorla C, Rosa F (2019) Bending fatigue behaviour of 17-4 PH gears produced via selective laser melting. *Proc Struct Integr.* <https://doi.org/10.1016/j.prostr.2020.02.068>
37. Concli F et al (2021) Bending fatigue behavior of 17-4ph gears produced by additive manufacturing. *Appl Sci* 11(7):3019. <https://doi.org/10.3390/app11073019>
38. Colombo S (2019) *Advanced Statistical Models for Bending Fatigue Characterization of Gears*. Politecnico di Milano, Milan, Italy. <http://hdl.handle.net/10589/165633>
39. Dobler D-IA, Hergesell IM, Stahl IK (2016) Increased tooth bending strength and pitting load capacity of fine-module gears. *Gear Technol* 33(7):48–53
40. Concli F (2021) Tooth root bending strength of gears: dimensional effect for small gears having a module below 5 mm. *Appl Sci* 11(5):2416