Experimental Conformity Level for comparison between endurance tests and life calculation models

Sébastien Blachère[1], Antonio Gabelli[1] and Guillermo E. Morales-Espejel[1,2]

[1]SKF Research and Technology Development, Houten, The Netherlands
[2] Université de Lyon, INSA-Lyon CNRS LaMCoS UMR5259 F69621, France

Abstract

Rolling bearing fatigue life is a stochastic process generally represented by a Weibull-like statistical distribution. The typical reliability indicator taken as characteristic performance of rolling bearings is the L_{10} life, i.e. durability for 10% failure probability among a large bearing population. For a specific bearing under specific operating conditions, calculation models are available to compute the values of L_{10}. Calculation models must also be compared to test data and the degree of conformity between the calculated life and the experimental life must be assessed. This article offers a new statistical measure, defined as Experimental Conformity Level (ECL), able to quantify the way a calculated life L_{10} fits with the estimated L_{10} from test data. The ECL combines the deviation between the estimated L_{10} from testing and the calculated L_{10}, with the precision of the experimental data. This gives a premium to the ECL value in case the fit is related to a large data set leading to precise estimations of the experimental L_{10} used in the assessment.

Keywords: Fatigue, Weibull statistics, Bearing Life, Life estimation

1. Introduction

Rolling bearings are machine elements that are subjected to Rolling Contact Fatigue (RCF) and usually operate under high rotation frequencies. This type of fatigue is categorized as Very High Cycle Fatigue (VHCF). Typically, rolling bearings reach the end of life by fatigue damage originated from the surface or the subsurface [1] in the rolling contact. It is also well known that seemingly identical bearings, running under the same operating conditions, have significantly different individual endurance lives. This occurs because the random presence of inhomogeneities in the material microstructure, surface finishing defects and geometrical tolerances have a very significant effect on the endurance of an individual bearing. This is why the fatigue life of an individual bearing is usually treated as a random variable [2]. Early models and also more recent bearing life models [1, 3, 4, 5, 6, 7] apply a combination of physical principles (i.e. RCF, Tribology) and statistics, usually based on the Weibull statistical model [8]. These models attempt to predict the number of revolutions for a given probability of survival of a population of seemingly equal bearings running under seemingly equal operating conditions. Following this approach, the L_{10} life rating of an individual bearing is the number of revolutions that the bearing will attain or exceed with a probability of survival or reliability of 90%. Within the framework of good economic sense, it was established in the past [3, 4, 9] that 90% reliability is indeed a suitable reliability level that can be verified by testing. This is usually done by performing endurance testing on a population sample of rolling bearings [10]. The objective of the current article is to introduce a new statistical method able to quantify the degree of conformity between endurance test data and the L_{10} predicted using bearing life calculation models.

2. Life statistical models

To model the randomness of physical phenomena like the fatigue of materials or mechanical product life, the Weibull statistical distribution is often used. It was introduced in the setting of material strength by Waloddi Weibull [2] and extended to a wide range of experimental data [8]. The 2-parameter Weibull distribution, denoting (η,β) its 2 parameters, is widely used together with its special case, the exponential distribution. The 2-parameter Weibull distribution turns into an exponential distribution when the shape parameter β equals to 1.

In both definitions, L denotes the random variable standing for the Life duration. The distributions are given with their two most common expressions, the more mathematical form with η (or λ for the exponential) as a scale parameter, and the more engineering form, using the 10th life percentile L_{10} as a scale parameter. A life percentile L_p is the time that p% of a large homogeneous population will not survive. Equivalently, L_p is the time that (100 − p)% of a large homogeneous population will survive.

Exponential Distribution

Weibull 1-parameter is the exponential distribution:

$$P(L > x) = \exp(-\lambda x) \text{ with } \lambda\ (exponential\ scale\ parameter) > 0$$

Weibull 2-parameter Distribution

Weibull 2-parameter distribution is:

$$P(L > x) = exp\left(-\left(\frac{x}{\eta}\right)^{\beta}\right) = 0.9^{\left(\frac{x}{L_{10}}\right)^{\beta}}$$

with η (scale parameter), β (shape parameter) > 0. By definition of a percentile, L_{10} being the 10th percentile, it corresponds to x such that $P(L > x) = 0.9$. Therefore,

$$L_{10} = \eta \times (-\ln 0.9)^{1/\beta}$$

This leads to the engineering formula for the Weibull 2-parameter distribution:

$$P(L > x) = exp\left(-\left(\frac{x}{\eta}\right)^{\beta}\right) = 0.9^{\left(\frac{x}{L_{10}}\right)^{\beta}}$$

with β (shape parameter), L_{10} (10th Life percentile) > 0

3. Life percentile estimation

For the Weibull 2-parameter distribution, the classical method used to estimate the parameters is the Maximum Likelihood Estimation (MLE). This method is known to be biased (see for instance [6]), this bias being non-negligible for the small sample size used in testing, less than 30 items typically. A recognized median bias correction technique (for the MLE estimation) was developed to obtain accurate estimates together with confidence bounds. The current bias correction method in life analysis of mechanical components uses correction factors computed from Monte Carlo simulations and applied to non-censored data [Non-censored data means that all bearings are run until failures] or Type II censored data [Type II censored data means that bearings are run in parallel until a fixed number of failures is reached and then all the running ones are stopped]. For a complete explanation of this bias correction techniques, see [11, 12, 13, 14]. See also the more recent article [15] referring to software able to proceed with such bias correction and also [16] focusing on improving this bias correction technique for test data including general censoring scenarios.

Any parameter estimation comes with a confidence interval showing the interval within which the target parameter lies with a chosen confidence level. The width of the confidence interval is a good indicator of the precision of the estimation.

The classical confidence interval for L_{10} is [$L_{10,5}$, $L_{10,95}$]. The levels 5 and 95 in the subscript correspond to the level of confidence associated with the calculation. In 90% of the case the interval [$L_{10,5}$, $L_{10,95}$] contains the true target L_{10} value.

Similarly $L_{10,50}$ can be computed from test data and called the median estimate of L_{10}.

The confidence interval gives then a key information on the precision of the L_{10} estimation. A wide confidence interval means that there is a high uncertainty around this estimation (like when you make a poll for an election asking only 10 people). A narrow confidence interval means that there is high precision around this estimation (like when you make the election poll asking 10,000 people chosen within a representative random sample).

Generally, this precision is measured via the ratio between the upper and lower bounds. Indeed, in the latter example, if having more or better data helps to get $L_{10,5}$ = 300 Mrevs and $L_{10,95}$ = 600 Mrevs instead of 100 and 800, the precision improved from a factor of 8 (800/100) to a factor of 2 (600/300). Although 5 and 95 are classical confidence levels, any other values can be used. For instance, 10 and 90 are also sometimes used leading to the interval [$L_{10,10}$, $L_{10,90}$].

The 2-parameter Weibull distribution has a second parameter β, shape parameter, which needs also to be estimated from the test data leading then to similar confidence bounds and intervals as for the L_{10}: β_5, β_{10}, β_{50}, β_{90} and β_{95}. The estimations of the shape parameters β are also biased and the bias correction techniques also applies to β. See again [11, 12, 13, 14] for more details and formulas.

4. Experimental Conformity Level (ECL)

A traditional use of confidence intervals like [$L_{10,10}$, $L_{10,90}$] is a comparison with calculated L_{10} values from life models. Such calculated L_{10} will be denoted L_{10}(calc). A classical method to compute the experimental confidence is to fit a Gaussian distribution on the confidence interval. The method is simply to take confidence bounds as percentiles of a Gaussian distribution (actually two distributions, one below the median estimate and one above the median estimate). This is an engineering method not based on statistical method.

This method is only expressing how safe the test is with respect to the calculated L_{10}(calc) but without judging potential underestimation. It has also the drawback of not considering the estimated value of the shape parameter β. This impact will be explained below using Figure 1.

We then introduce a new statistical quantity, called: **"Experimental Conformity Level (ECL)".** This parameter aims to quantify the conformity between the calculated life L_{10}(calc) and the result from the tests (including the confidence intervals on the L_{10} and the β). This parameter is linked to the experimental confidence (Gaussian) but provides new features that can be illustrated as follows:

- It gives a premium, respectively a penalty, for narrow, respectively wide, confidence intervals on L_{10}
- It takes into account the estimated β and its associated confidence bounds

The second point is of importance especially when the estimated β is high because, in such a case, a value slightly different from the L_{10} can correspond to a much lower reliability level as shown in the subsequent examples.

Example: If L_{10}=100 Mrevs and Beta = 1.1 are supposed to be known, then 150 Mrevs corresponds to the true L_{15}, but if Beta = 2, then 150 Mrevs corresponds to the true L_{21}. This is illustrated in Figure 1.

Therefore, at a high β, an identical quantitative error on the L_{10} value that is calculated is more detrimental for the final reliability of the product given to the customer. Therefore a wider confidence interval is more detrimental at high β than at low β.

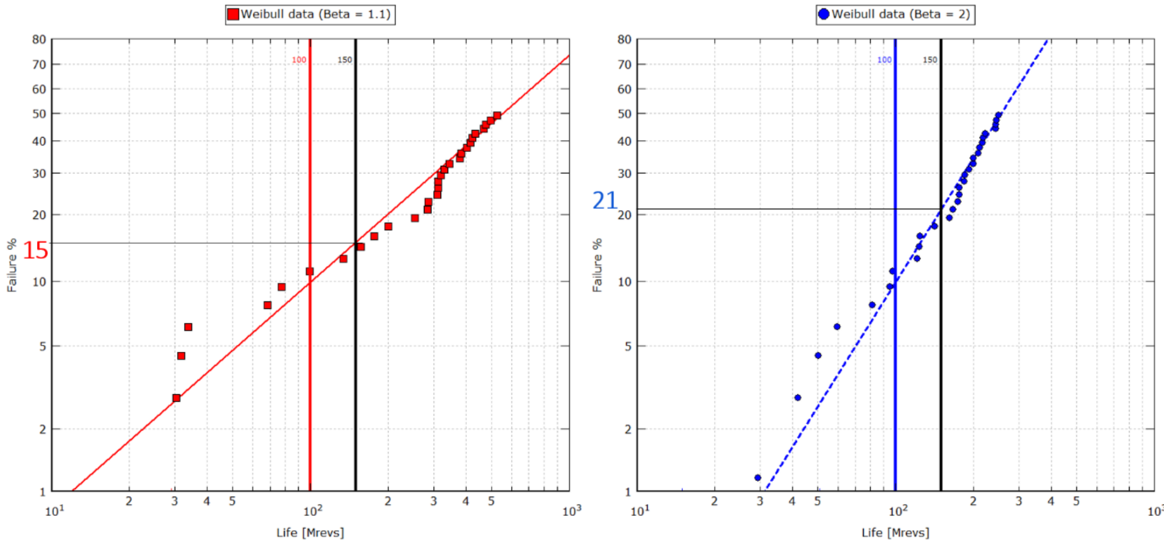

Figure 1- L10 estimation sensitivity to Beta (1.1 and 2)

The formula of the ECL is built by computing two failure percentages P1 and P2. These failure percentages are computed using L_{10} and β values taken from the confidence intervals. The values are chosen to be conservative. P1 measures the risk to have a calculated life too high compared to the true life. This risk is evaluated using β_{90} (to be conservative) in order to reflect the sensitivity to β illustrated in Figure 1. P2 measure the risk to have a calculated life too low compared to the true life. The conservative approach is taken for P1 since the associated risk is more detrimental.

Assume that the true L_{10} equals the $L_{10,10}$ and the true β equals the β_{90}, then the calculated L_{10}(calc) corresponds to the true L_{P1}:

$$P1 = 100 \times \left(1 - 0.9^{\left(\frac{L_{10}(\text{calc})}{L_{10,10}}\right)^{\beta_{90}}}\right)$$

Each value of P1 is associated with a percentage X% by:

- $P1 \leq 15 \Rightarrow X = 100\%$
- $15 < P1 < 25 \Rightarrow X = (25 - P1) \times 10\%$
- $P1 \geq 25 \Rightarrow X = 0\%$

The objective of the value X is to give a penalty when the calculated L_{10}(calc) risks to lead to too high life percentile. This risk being computed from the estimated L_{10} and β.

The extreme values (15 and 25) are chosen to reflect acceptable risks when looking at actual reliability levels. Between those extreme values, X is simply linearly interpolated.

Assume now that the true L_{10} equals the $L_{10,50}$ and the true Beta slope equals β_{50}, then the calculated L_{10} corresponds to the true L_{P2}:

$$P2 = 100 \times \left(1 - 0.9^{\left(\frac{L_{10}(\text{calc})}{L_{10,50}}\right)^{\beta_{50}}}\right)$$

Each value of P2 is associated with a percentage Y% by:

- $P2 \leq 3 \Rightarrow Y = 0\%$
- $3 < P2 < 8 \Rightarrow Y = (P2 - 3) \times 20\%$
- $P2 \geq 8 \Rightarrow Y = 100\%$

The objective of the value Y is to give a penalty when the calculated L_{10}(calc) could lead to too low life percentile. This risk being computed from the estimated L_{10} and β.

The extreme values (3 and 8) are chosen to reflect acceptable risks when looking at actual reliability levels. Between those extreme values, Y is simply linearly interpolated.

The final ECL is defined as

$$ECL = Max\{(X + Y - 100), 0\}\%$$

combining values from the lower and upper true-life percentiles corresponding to the calculated L_{10}. This means that having confidence, from the test results, that the calculated L_{10} is actually between the true L_8 and the L_{15} leads to an ECL of 100%. Also, if the calculated L_{10} has a risk to be less than the true L_3 or higher than the true L_{25}, then the ECL becomes 0%. The intermediate cases are linearly interpolated between the latter extreme cases.

The motivation behind taking 90% confidence in the calculation of P1 ($L_{10,10}$ and β_{90}) and 50% confidence in the calculation of P2 ($L_{10,50}$ and β_{50}) is to put more weight on the most conservative (business-wise) case.

In order to interpret the ECL, a high ECL percentage (above 90%) will then guaranty strong and trustful conformity between the test results and the calculated life. This can be applied either to test data or field data.

5. Discussion

The ECL is a novel method to assess at the same time the accuracy of a life estimation form a life test and the fit between the test result with a calculated life. The lack of confidence can come from two sources: either because the test has large confidence intervals (too few tested samples, poor Weibull fit…) or because the calculated life does not fit with the test results (estimated life from the test). Each of these two sources will penalize the ECL value.

In order to better understand the added value that the ECL could bring to the statistical analysis of test data, we present 3 examples of endurance tests where the data has been normalized so that the $L_{10,50}$ is always 100 (see Figure 2)

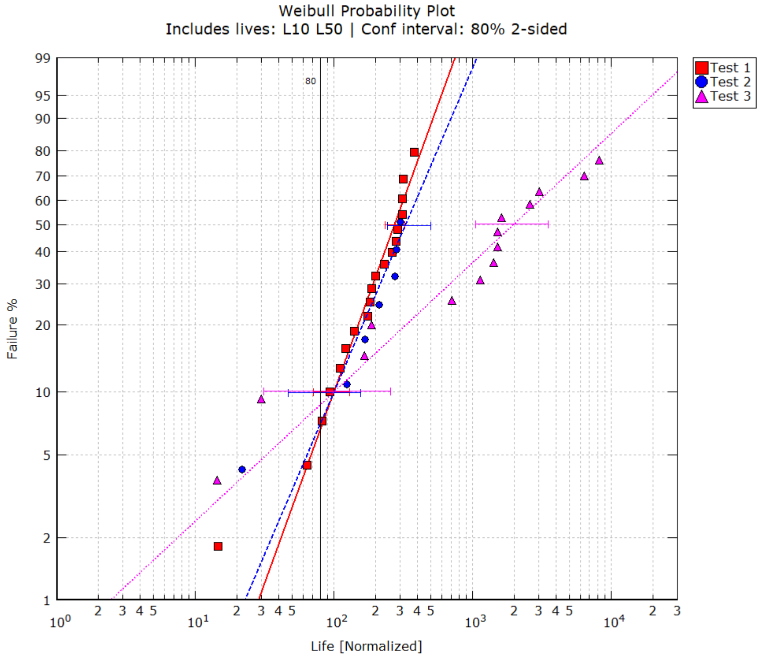

Figure 2 – Weibull plot for three tests (normalized)

Test 1 has many failures and a high beta. Test 2 has a limited number of failures and still a high beta. Test 3 has many failures and a low beta.

If we assume L_{10}(calc)=80 (a conservative value but rather close to the $L_{10,50}$=100), the ECL can be calculated for each of the 3 tests, see Table 1 that shows all the data for the calculation and the last row shows the calculated ECL for each test.

Table 1. ECL calculated for each test of Figure 2.

Parameter	Test 1	Test 2	Test 3
β, 90 %	2.36	2.38	0.79
$L_{10,10}$	71.14	46.7	31.35
β, 50%	1.87	1.56	0.63
$L_{10,50}$	100	100	100
$L_{10}(calc)$	80	80	80
Calculated ECL	**74%**	**0%**	**52%**

The use of the ECL allows to conclude that Test 1 ensures a very high conformity between the test and the calculated life. This is due to the very narrow confidence interval on L_{10}. The calculation gives P1=13 and P2=6.7, so the use of L_{10}(calc) is not leading to any significant risk of overestimation or underestimation of the life.

Test 2 is not giving any conformity, although it has 7 failures and a reasonable confidence interval width. The reason is that the beta is high (illustrated by a high slope on the Weibull plot). Such high beta means that a small shift in life calculation can have a big impact on the reliability. The calculation gives P1=31.6 and P2=7.2. This means that selling L_{10}(calc)=80 as correct, there is a risk that this value corresponds to the $L_{31.6}$ instead of the L_{10}. So, when a customer is expecting 10% failures maximum at a designed time, he/she may get 31.6% failures, 3 times more! In such case, more test data must be obtained to have a better estimation of the life.

Test 3 ensures limited conformity. This is partially due to the wide confidence interval, but the low beta (illustrated by a low slope on the Weibull plot) is forcing this large width. The computation of the ECL allows to balance the impact of the beta and the impact of the limited sample size. In the case of Test 2, we tested many samples and have got many failures. Therefore, we essentially obtained the inherent width for the confidence interval. The calculation gives P1=19,8 and P2=8.7, which means that the error in terms of life percentile that can be made by using L_{10}(calc) remains reasonable.

To complete the analysis, we could study Test 4 with fewer failures and sill a low beta value (see Figure 3). This will increase the uncertainty (and then the width of the confidence interval) loosing then any conformity (ECL=0%).

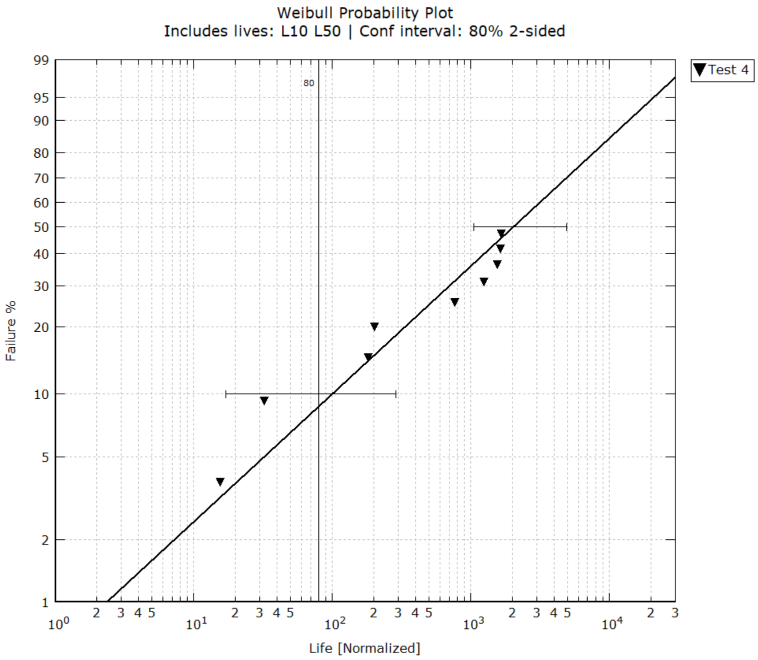

Figure 3 - Test 4 (Normalized / low beta / less failures)

6. Conclusions

A new statistical measure, the Experimental Conformity Level (ECL) has been introduced to quantify the way a calculated life L_{10}(calc) fits with experimental data. The ECL weights the deviation between the estimated L_{10} and the calculated L_{10}(calc) using the confidence bounds on both the L_{10} and the β. This gives a premium to the ECL value when we deal with large set of test data leading to high precision in the estimations of the L_{10} obtained from testing.

The ECL calculation takes into account the estimated value of the Weibull shape parameter Beta and this gives a weighted measure of the fit with the experimental data and overcomes the potential misinterpretation regarding the actual deviation between the calculated and the estimated life. Indeed, identical deviations will have different reliability consequences when they are related to test results with significant different Beta values (see Figure 1).

The ECL is a new statistical measure that provides the following advantages:

- Quantitative statement on how well a life calculation model correlates to the experiments
- Ability to rank different life calculation models based on actual experimental data
- Proven robustness to compensate for different values of the shape parameter Beta of tests

Acknowledgement

The authors wish to thank Mr. Bernie van Leeuwen, Director of the SKF Research and Technology Development Center for his kind permission to publish this article.

References

[1] G. Morales-Espejel, A. Gabelli and A. de Vries, "A Model for Rolling Bearing Life with Surface and Subsurface Survival—Tribological Effects," *Trib. Trans.,* vol. 58, pp. 894-906, 2015.

[2] W. Weibull, "Statistical theory of the strength of materials," *Ing. Vetenskaps Akad.,* pp. 147-151, 1939.

[3] G. Lundberg and A. Palmgren, "Dynamic Capacity of Rolling Bearings," *Acta Polytechnica, Mechanical Engineering Series,* vol. 1, no. 3, pp. 1-52, 1947.

[4] G. Lundberg and A. Palmgren, "Dynamic Capacity of Roller Bearings," *Acta Polytechnica, Mechanical Engineering Series,* vol. 2, no. 4, pp. 96-127, 1952.

[5] E. Ioannides and T. Harris, "A New Life Model for Rolling Bearings," *Journal of Tribology,* vol. 107, pp. 367-378, 1985.

[6] A. Zaretsky, STLE Life Factors for Rolling Bearings, 2nd ed., 1999., Park Ridge, IL, : Society of Tribologists and Lubrication Engineers, 1999.

[7] ISO/TR, "ISO/TR 1281-2 Rolling Bearings – Explanatory Notes on ISO 281, Part 2 Modified rating life calculation, based on a systems approach to fatigue stresses," ISO, 2008.

[8] W. Weibull, "A statistical distributron function of wide applicability," *J. Appl. Mech.,* vol. 18, no. 3, pp. 293-297, 1951.

[9] T. Tallian, "Weibull distribution of Rolling contact fatigue life and deviations therefrom," *ASLE transactions,* vol. 5, pp. 183-196, 1962.

[10] G. Morales-Espejel and A. Gabelli, "Rolling bearing performance rating parameters: Review and engineering assessment," *Proc. IMechE, Part C, J. of Mech. Eng. Sci.,* vol. 234, no. 15, pp. 3064-3077, 2020.

[11] L. Bain and C. Antle, "Estimation of parameters in the Weibull distribution.," *Technometrics,* vol. 9, no. 4, pp. 621-627, 1967.

[12] D. R. Thoman, L. J. Bain and C. E. Antle, "Inferences on the parameters of the Weibull distribution," *Technometrics,* vol. 11, no. 3, pp. 445-460, 1969.

[13] J. Mc Cool, "Evaluating Weibull endurance data by the method of maximum likelihood," *ASLE transactions,* vol. 13, no. 3, pp. 189-202, 1970.

[14] J. Mc Cool, "Using Weibull regression to estimate the load-life relationship for rolloing bearings," *ASLE transactions,* vol. 29, pp. 91-101, 1986.

[15] J. Mc Cool, "Software for Weibull inference.," *Quality Engineering,* vol. 23, pp. 253-264, 2011.

[16] Blachere, "A new bias correction technique for weibull parametric estimation.," *Quality Engineering Application and Research, IQF,* 2015.