BEARING WORLD

www.bearingworld.org

Journal

Volume 6_2021

Editors: G. Poll _ A. Grunau _ C. Kunze

Published by
FVA sharing drive innovation
VDMA Verlag

Imprint

Bearing World Journal
Volume 5, December 2020

Published by:
Forschungsvereinigung Antriebstechnik e.V. (FVA)
Lyoner Straße 18
60528 Frankfurt am Main
Germany
www.fva-net.de

© 2022
VDMA Services GmbH
Lyoner Straße 18
60528 Frankfurt am Main
Germany
www.vdma-verlag.com

All rights reserved, particularly the right of duplication and disclosure, as well as translation. No part of the work may be reproduced in any form (print, photocopy, microfilm or any other method) without written consent from the publisher or saved, processed, duplicated or disclosed.

Online-ISSN 2566-4794

Volume 6_2021

Dear reader,

Globalization increasingly requires more and more international networking between research and development engineers. In response to this, the German Research Association for Drive Technology (FVA) launched the first Bearing World conference in 2016. With that inaugural meeting, the FVA initiated a very fruitful international dialogue in which researchers and developers from universities and bearing manufacturers came together with users and experts from the industry. The Bearing World conference usually is held every two years; more than 280 experts from 18 countries met at the last Bearing World conference in 2018 in Kaiserslautern, Germany, to share the latest research findings in the world of bearings.

The Bearing World Journal, which is published annually, serves to foster exchange between international experts during non-conference years by featuring peer-reviewed, high-quality scientific papers on rolling element bearings as well as plain bearings. As an international expert platform for publishing cutting-edge research findings, the journal intends to contribute to technological progress in the field of bearings.

We are now starting to prepare the 2022 edition of Bearing World Journal and are looking forward to new contributions from the scientific and industrial communities. We would like to thank all authors for their fascinating contributions to Bearing World Journal No. 6.

_ **Prof. Dr.-Ing. Gerhard Poll,** Initiator, Head of international Scientific Board
_ **Dr.-Ing. Arbogast Grunau,** President of the FVA Management Board
_ **Christian Kunze,** Editor-in-chief

Please send the paper you intend to publish in the next issue of the Bearing World Journal via e-mail as Word document to FVA (**submission@bearingworld.org**).
In addition please attach a PDF document.

Bearing World Scientific Board

Scott Bair, Georgia Institute of Technology, USA
Prof. Harry Bhadeshia, University of Cambrigde, Great Britain
Prof. Stefan Björklund, KTH Royal Institute of Technology, Stockholm, Sweden
Prof. Benyebka Bou-Said, Institut National Des Sciences Appliquées (INSA) Lyon, France
Prof. Bernd Bertsche, Universität Stuttgart, Germany
Prof. Ludger Deters, TU Magdeburg, Germany
Prof. Duncan Dowson, University of Leeds, Great Britain (†)
Prof. Rob Dwyer-Joyce, University of Sheffield, Great Britain
Prof. Michel Fillon, Université de Poitiers, France
Prof. Sergei Glavastkih, KTH Royal Institute of Technology, Stockholm, Sweden
Prof. Irina Goryacheva, Russian Academy of Sciences, Russia
Prof. Feng Guo, Qingdao Technological University, China
Prof. Martin Hartl, Brno University of Technology, Czech Republic
Prof. Stathis Ioannides, Imperial College London, Great Britain
Prof. Georg Jacobs, RWTH Aachen University, Germany
Prof. Motohiro Kaneta, Brno University of Technology, Czech Republic
Prof. Michael M. Khonsari, Louisiana State University, USA
Prof. Ivan Krupka, Brno University of Technology, Czech Republic
Prof. Roland Larsson, Luleå University of Technology, Sweden
Prof. Antonius Lubrecht, Institut National Des Sciences Appliquées (INSA) Lyon, France
Prof. Piet Lugt, SKF Nieuwegin; University of Twente, Enschede, Netherlands
Prof. Jianbin Luo, State Key Laboratory of Tribology, Tsinghua University, China
Prof. Guillermo Morales-Espejel, INSA Lyon, France
Prof. Anne Neville, University of Leeds, Great Britain
Prof. Hiroyuki Ohta, Nagaoka University of Technology, Japan
Prof. Gerhard Poll, Leibniz University Hanover, Germany
Prof. Martin Priest, University of Bradford, Great Britain
Prof. Farshid Sadeghi, Purdue University, Lafayette, Indiana, USA
Prof. Richard Salant, Georgia Institute of Technology, USA
Prof. Bernd Sauer, TU Kaiserslautern, Germany
Prof. Ian Sherrington, University of Central Lancashire, Great Britain
Prof. Hugh Spikes, Imperial College London, Great Britain
Prof. Gwidon Stachowiak, Curtin University Australia, Australia
Prof. Kees Venner, University of Twente, Enschede, Netherlands
Prof. Philippe Vergne, Institut National Des Sciences Appliquées (INSA) Lyon, France
Prof. Fabrice Ville, Institut National Des Sciences Appliquées (INSA) Lyon, France
Prof. Sandro Wartzack, Friedrich-Alexander-University Erlangen-Nürnberg, Germany
Prof. John A. Williams, University of Cambridge, Great Britain
Prof. Hans-Werner Zoch, IWT Stiftung Institut für Werkstofftechnik, Bremen, Germany

Contents

The Weibull Distribution and the Problem of Guaranteed Minimum Lifetimes

Prof.Dr.-Ing.habil. Harald Rosemann IMKT (Leibniz University Hannover)
Josephine Kelley IMKT (Leibniz University Hannover)
Prof.Dr.-Ing. Gerhard Poll IMKT (Leibniz University Hannover)

Abstract

For service life tests, a shifted Weibull distribution, also known as the translated or three-parameter Weibull distribution, is commonly used. The shifted Weibull distribution promises completely fault-free operation until time $t = L_0$, in other words, in the early stage the process is deterministic. Only after this phase does the distribution allow random behavior, i.e. from the time $t = L_0$ on, the process is stochastic. This model, which is based on two consecutive time periods of quite different nature, is at odds with the idea of a continuously progressing fatigue, wear or decay process as long as there are no influences from outside. To replace this arguably inconsistent model, variants of the Weibull distribution of purely stochastic nature are proposed and investigated that start with a reduced probability of failure before transitioning to normal Weibull behavior.

1 Introduction

Materials wear and fatigue, and, as a result, failures occur. Individual failures as a consequence of fatigue or wear occur at unpredictable, statistically distributed times. It is often assumed that the service lifetimes are distributed according to the Weibull distribution, as this is the distribution that yields the highest target values in parameter estimation using optimization methods such as the maximum likelihood procedure. The original Weibull distribution is defined by two parameters.

Attempts have been made to develop a modified variant of the Weibull distribution by introducing a third parameter in order to describe failure behavior that is initially infrequent. This variant is in constant use, which is clear from some of the first entries from an internet search for the term 'Weibull distribution'. The additional third parameter, also known as threshold, accounts for a minimum initial operating time, during which an (alleged) absolute and total absence of failure is guaranteed. In the following, we consider whether this assumption is justified or should be replaced by a more stringent approach.

2 The problem

2.1 The Weibull distribution with two parameters

For many service life tests, the original Weibull distribution with two parameters can suitably represent the observed values. In general, F(t) denotes the cumulative distribution function of a time-dependent random variable and W(t) specifically denotes the Weibull cumulative distribution function:

$$F(t) = W(t) = \begin{cases} 1 - e^{-\left(t/T\right)^{\beta}}, & t \geq 0,\ \beta > 0,\ T > 0 \\ 0, & t < 0 \end{cases} \qquad (1)$$

An important characteristic is that, in the exponential function, the time t itself is raised to the power β. The parameter T is called the characteristic time; regardless of the value of β, one always has $W(T) = 1 - {}^{1}/_{e} \approx 0.632$.

At $t = 0$, the cumulative distribution function $W(t)$ is equal to zero and begins to increase monotonically as a function of t, approaching the value 1 for large t. From the values of the cumulative distribution function, one attains the probability that a failure occurs at or before time t. With $W(t) = 0$ for $t < 0$, the distribution shows that the effect cannot occur before the cause, i.e. a failure can only be expected after the start of the damage-inducing loading; this fundamentally excludes the possibility of failure before the damage-inducing loading, and, indeed, the probability of a negative service lifetime is zero.

Instead of the characteristic value T, one commonly uses the L_{10}-lifetime and algebraically manipulates Eqn. (1) into:

$$F(t) = W(t) = \begin{cases} 1 - e^{\ln(0.9)\left[\frac{t}{L_{10}}\right]^{\beta}}, & t \geq 0,\ \beta > 0,\ L_{10} > 0, \\ 0, & t < 0 \end{cases} \tag{2}$$

Once again, there is a value independent from β that the cumulative distribution function depends on: by definition, $W(L_{10}) = 0.1$ and so L_{10} gives the time up to which 10% of failures are to be expected.

2.2 The shifted Weibull distribution (translated or 3-parameter Weibull distribution)

For certain applications, one discovers that the initial number of failures is lower than predicted by the standard Weibull distribution. This deviation is attributed to processes such as wear, deterioration, or fatigue, which usually require a certain amount of time for damage to develop into failure. For this reason, Snare [1] and later on Bergling [2], used a third parameter L_0, also known as threshold, in the evaluation of roller bearing lifetimes to shift the cumulative distribution function to the right, according to

$$F(t) = W(t) = \begin{cases} 1 - e^{\ln(0.9)\left[\frac{t-L_0}{L_{10}-L_0}\right]^{\beta}}, & \begin{cases} t \geq L_0, \\ \beta > 0, \\ L_{10} > L_0 \geq 0 \end{cases} \\ 0, & t < L_0 \end{cases} \tag{3}$$

to obtain a 'better' fit to the data points for early failures. When plotted, this correction can be visually judged to be adequate. Also, if the superiority of a parameter set is to be judged using the target value that arises from the optimization of an estimation process such as the maximum likelihood method, then the three-parameter Weibull distribution should indeed be preferred to the two-parameter Weibull distribution. On the one hand, this is the argumentation in favor of the three-parameter Weibull distribution.

2.3 The conflict

On the other hand, however, shifting the original Weibull distribution to get the curve of Eqn. (3) introduces a new phase into the model. It is valid for $t < L_0$ and is of purely deterministic nature; the second phase, valid for $t \geq L_0$, is stochastic. These two domains of fundamentally different nature share the predefined, non-random border at $t = L_0$.

In the first part, the model ensures that there are no failures before $t = L_0$. An event in this region representing a failure can not occur and is labeled as 'impossible' by definition of Eqn. (3). Strictly spoken, such a fundamental statement cannot be deduced or validated purely from observation, regardless of the number of data points. Even though an estimator $\widehat{L}_0$ for a sample exists and can be computed according to Park [3], this does not on its own prove the existence of a failure-free period of time L_0.

From a numerical point of view, one hardly notices a difference between 'exactly zero' and very, very small, say one billionth or even less. Qualitatively, on the other hand, the 'impossible event' is fundamentally different from one with a low probability. The first is based on abstract definition, the other is a matter of the real world; in the first case, one can be completely unconcerned, in the other one, precautionary measures may become necessary.

Additionally, this model necessitates an exogenous 'timer setting' that triggers the transition to the second phase after which the ongoing fatigue or wear processes are allowed to develop into a failure.

This is an unsatisfactory situation as there is a conflict. On the one hand, one has the best distribution (among the ones tested), while on the other hand, the statement and core assumptions of the distribution do not apply to the continuously progressing process that generates the observed values. A pragmatic way to resolve this issue would be to consider the Weibull distribution with $L_0 > 0$ an approximation. Nevertheless, one must be prepared to fend off any outside claims that one has guaranteed safety from premature failures. There is a dilemma with only one possible resolution: to find a distribution that yields even higher target values in parameter estimation, that can also be interpreted without any problems.

3 New approach

3.1 Hyperbola instead of the straight lines

The question therefore becomes whether it is possible to find an intermediate solution that preserves the Weibull character and allows for delayed failure behavior without permitting any misinterpretation. It is useful to simplify the equations by using the $(L_{10} - L_0)$-normalized variables $t' = t/(L_{10} - L_0)$ and $L'_0 = L_0/(L_{10} - L_0)$. We then can write what is different in each distribution as auxiliary functions of t' as $g_2(t') = t'$ and $g_3(t') = t' - L'_0$, respectively; the index is counting the parameters. The functions $g(t')$ are both the basis which is taken to the power β in the cumulative distribution function of Weibull.

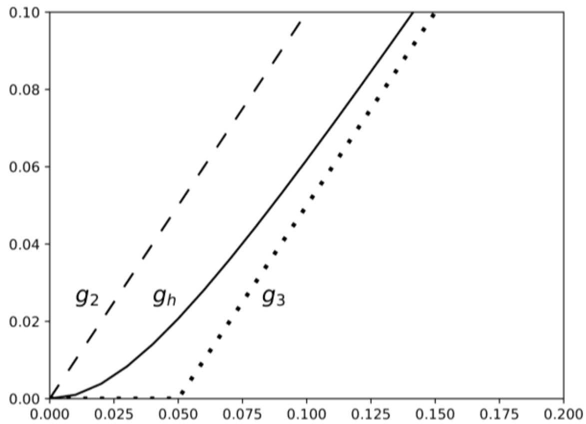


Fig. 1: $g(t')$ over t', $L'_0 = 0.05$

These two functions that depend on t' and L'_0 are shown in Fig. 1 as two parallel lines with g_2 on the left as a dashed line, and shifted by $L'_0 = 0.05$ to the right as g_3, which is represented by a dotted line. In the area between the two lines, we may draw another curve. This curve should increase monotonically from the value 0 at $t' = 0$ and approach the line g_3 for large t'. By taking the same name L'_0 for a similar parameter, an obvious choice would be the branch of a hyperbola, i.e.

$$g_h(t') = -L'_0 + \sqrt{t'^2 + L'^2_0}, \quad t' \geq 0,\ L'_0 \geq 0 \tag{4}$$

which is represented by the continuous line in Fig. 1. Near $t' = 0$ the function $g_h(t')$ behaves like $t'^2/2L'_0$, i.e. it begins with a horizontal tangent.[1]

3.2 Comparison of the cumulative distribution functions

The three versions of $g(t')$ lead to three Weibull distribution functions via $W(g(t'))$, where each $g(t')$ replaces the original t'; we apply the notation W_2 to mean $W(g_2(t'))$ for each $g(t')$. Figures 2 and 3 show the curves with linear coordinates on the left and Weibull coordinates on the right, which shows the original Weibull distribution as a straight line. For these calculations, $\beta = 1.35$ was chosen.

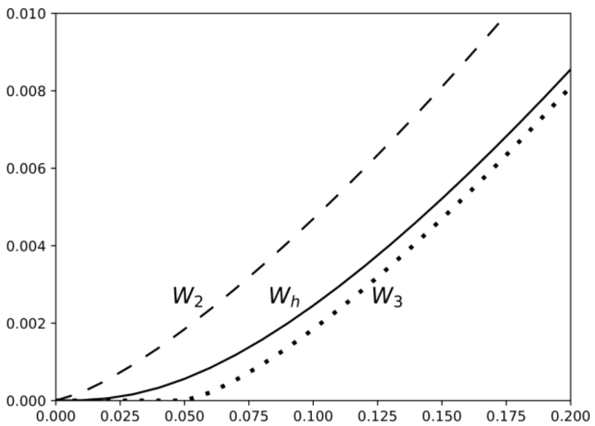


Fig. 2: $W(g(t'))$ over t', $L'_0 = 0.05$, linear coordinates

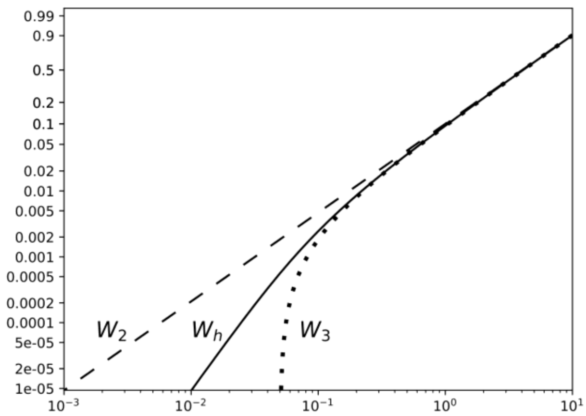


Fig. 3: $W(g(t'))$ over t', $L'_0 = 0.05$, Weibull coordinates

The desired sensible behavior is clearly visible. On the left in Fig. 2, W_h remains close to 0 longer than the original W_2 and in the further course it approaches W_3 more and more. In the Weibull diagram on the right, W_h begins steeper than W_2 but not as abruptly as W_3, which starts at the fixed value $t' = L'_0$.

Thus, early failures are less likely by the hyperbola approach according to Eqn. (4) than for the original Weibull distribution W_2 but not completely impossible before $t' = L'_0$ as it is for W_3. For larger values of t', the curves W_h and W_3 merge as a consequence of Eqn. (4), which can also be seen in the representation with Weibull axes. Fig. 2 with undistorted axes shows only the section with small t'; when these axes are expanded to $t' = 10$ as was done for the Weibull coordinates, one would not be able to distinguish the curves, especially for large t'.

[1] If, on the other hand, one wants to represent particularly frequent early failures rather than delayed ones, one may use a different hyperbola branch that increases quickly at $t' = 0$, just like the square root function: $g_h(t') = \sqrt{t'^2 + 2t'L'_0}$

The stated goal has been achieved since a useful replacement has been found. It is of continuously stochastic nature without a deterministic portion. Using initially small probabilities, it can represent delayed failures. There is no necessity for assumptions of a guaranteed lifetime L_0.

4 Extension of the hyperbola

4.1 Further replacement of the straight lines

Is the potential of the first approach now exhausted or can it be pursued further and expanded? The characteristic course of the hyperbola branch should be preserved; how can it be varied? By generalizing the square root and the second power, we arrive at

$$g_c(t') = -L'_0 + \left[\left(t' \right)^c + \left(L'_0 \right)^c \right]^{1/c}, \quad t' \geq 0, \quad L'_0 \geq 0, \quad c \geq 1 \tag{5}$$

with the new parameter c, the name of which is also used as an index for $g_c(t')$, denoting the modified approach.[2] The curve of $g_c(t')$ increases monotonically with t', as was the case with the first hyperbola in Eqn. (4); by replacing t' with $g_c(t')$ in the Weibull formula, the definition of a distribution is still fulfilled.

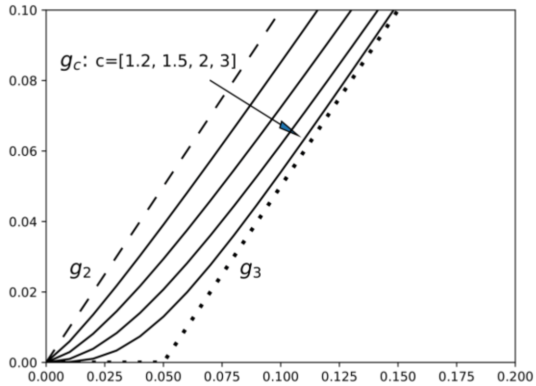


Fig. 4: $g(t')$ *over* t', $L'_0 = 0.05$

Figure 4 shows a sheath of continuous curves between the original straight lines, which are represented by dashed line and dotted line, respectively. The list shows the corresponding values for c, where the arrow is pointing in the direction of increasing values. For $t' = 0$, the curves increase with t', with almost horizontal tangent lines, like $t'^c / c L_0'^{c-1}$, and with increasing c they can thus lie along the time axis more closely and for a longer duration.

The new formula does not just fill the area between the first two straight lines, it also has the nice property of including the original Weibull distribution for $c = 1$, while the other shifted one is boundary case for $c \to \infty$.

---

[2] Values in the range $0 < c < 1$ generate more frequent early failures

4.2 Comparison of the cumulative distribution functions

The appearance of the corresponding cumulative distribution functions, on the left in equally divided coordinates and on the right with Weibull axes, now turns out as one might expect; between the two original curves, there are arbitrarily many intermediate variants. In Fig. 6 with Weibull coordinates, the curves run from the bottom almost straight up towards the line W_2 with varying curvature. Because the series expansion of $g_c(t')$ begins with order t'^c for small times t', the initial slope of the W_c in the Weibull coordinates is $c\beta$.

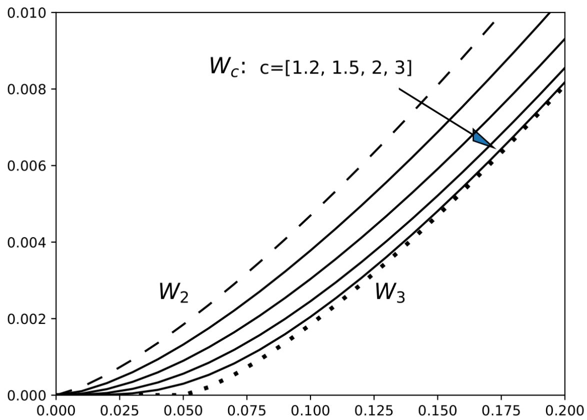


Fig. 5: $W(g(t'))$ over t', $L'_0 = 0.05$, linear coordinates

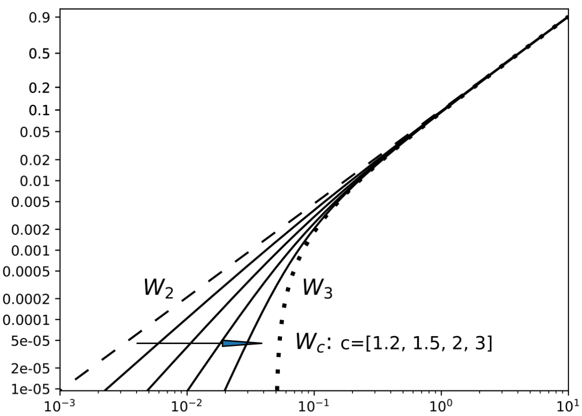


Fig. 6: $W(g(t'))$ over t', $L'_0 = 0.05$, Weibull coordinates

4.3 Special properties

As an example, Fig. 7 repeats the representation of the first hyperbola approach according to Eqn. (4). Additionally, a series of small circles shows the nearly linear initial slope of 2β and continues it to larger values. We see that this line, together with W_2, can be pieced together to conservatively approximate W_h. This is reminiscent of the old rule for the design of ball bearings, according to which the value of β should be increased to 1.5 for service lifetimes below L_{10}.[3]

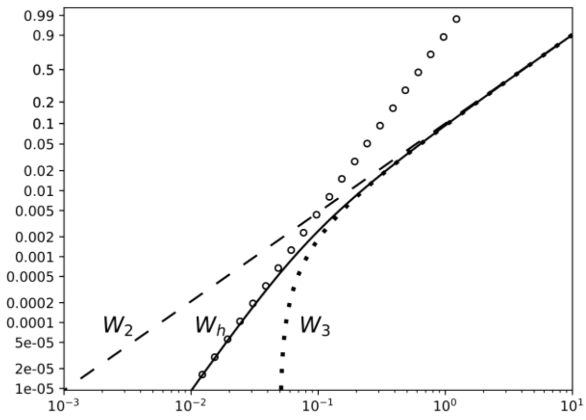


Fig. 7: $W(g(t'))$ and asymptote over t', $L'_0 = 0.05$, Weibull coordinates

[3] This modification is taken into account in the calculation of the reliability factor a_1 according to ISO 281 (2007 and previous versions) [4].

4.4 A short look at parameter estimation

For the original Weibull distribution with two parameters, one calculates the estimators $\hat{\beta}$ and $\hat{L}_{10}$ from measured service lifetimes. Every measurement has an influence on each of those two values. At most, the extreme failure times with low and high values have more influence on the result of the slope $\hat{\beta}$ in the Weibull coordinates and the intermediate values have more weight in the calculation of $\hat{L}_{10}$.

This changes for the four parameters of the extended approach. The new values L_0 and c arise on their own as the influence and efficacy in the initial range; as a result, their estimation $\hat{L}_{10}$ and $\hat{c}$ depend mainly on the times of the first early failure cases. This is related to a reduced dependence of both estimators $\hat{\beta}$ and $\hat{L}_{10}$ on the first early failure cases. A sufficiently large number of early failure cases is therefore necessary in order to estimate the new parameters accurately and reliably. If so far the number of early failures appeared to be sufficient to calculate the estimate $\hat{L}_0$ of the shifted Weibull distribution alone, such a number might now also be good enough to get usable values for $\hat{L}_0$ and $\hat{c}$ for the proposal. Moreover, typical values for certain special applications can be considered, such as the typical values of β equal to 1.11 for roller bearings primarily with point contacts versus β equal to 1.35 for cases with point and line contacts.

5 Conclusion

For continuously progressing wear and fatigue processes, the Weibull distribution with three parameters is not a suitable model for the distribution of service lifetimes as long as there are no external influences; it can only be viewed as a pragmatic approximation. In the approach presented here, the linear dependence on time t is replaced by a hyperbolic dependence. This new variant can represent delayed failure behavior in a fully stochastic model while avoiding difficulties with interpretation of the parameters, in particular with respect to guaranteed service lifetimes.

Acknowledgements

I would like to thank my highly esteemed colleagues M.Sc. Josephine Kelley and Prof.Dr.-Ing. Gerhard Poll at the IMKT, who translated this script from German and gave valuable support and encouragement for the preparation of this publication.

References

[1] Snare, B., Neuere Erkenntnisse über die Zuverlässigkeit von Wälzlagern; Die Kugellager-Zeitschrift, Heft Nr. 162 (1969) S. 3-7.

Remark: Figure 2 refers to an L_{10} of 15 Million rotations; in the text, however, it states: "Die Lager liefen . . . bei . . . einer Belastung, die nach dem Katalog einer L_{10}-Lebensdauer von 10 Millionen Umdrehungen entspricht." (The bearings ran for a load that, according to the catalogue, corresponds to an L_{10} service lifetime of 10 Million revolutions.)

[2] Bergling, G., Betriebszuverlässigkeit von Wälzlagern; Die Kugellager-Zeitschrift, Jahrgang 51, Heft Nr.188 (1976) S. 1-10.

Remark: Figure 3 (agrees with Figure 2 in [1]) shows an L_{10} of 15 Million revolutions; in the legend, a different value of 10 Million revolutions is stated.

[3] Park, C., A Note on the Existence of the Location Parameter Estimate of the Three-Parameter Weibull Model Using the Weibull Plot; Mathematical problems in engineering (2018), S. 1-6.

[4] ISO 281:2007, Rolling bearings - Dynamic load ratings and rating life.
(2007, and previous versions)

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.

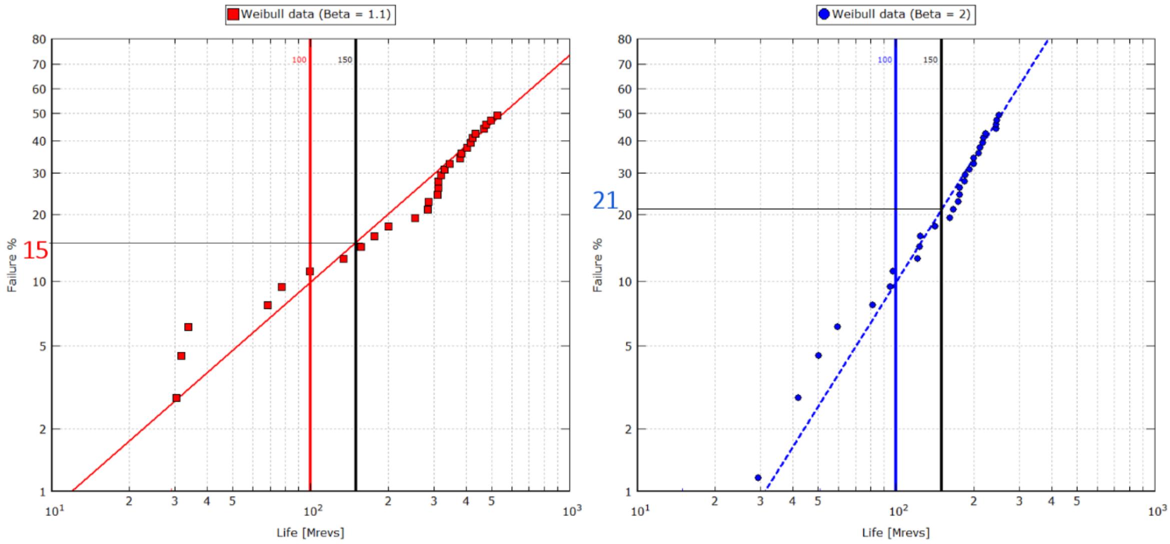


Figure 1- L10 estimation sensitivity to Beta (1.1 and 2)

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)

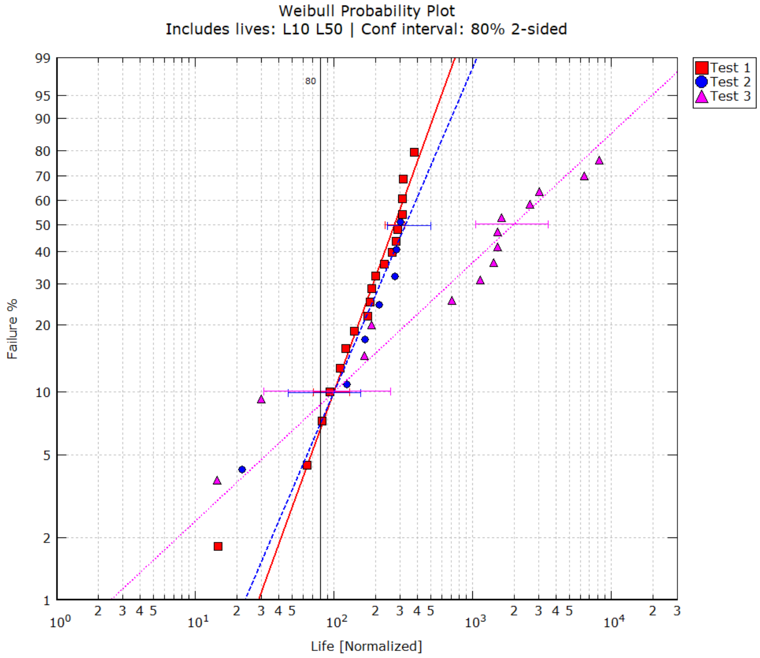


Figure 2 – Weibull plot for three tests (normalized)

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
$\beta, 90\ \%$	2.36	2.38	0.79
$L_{10,10}$	71.14	46.7	31.35
$\beta, 50\%$	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%).

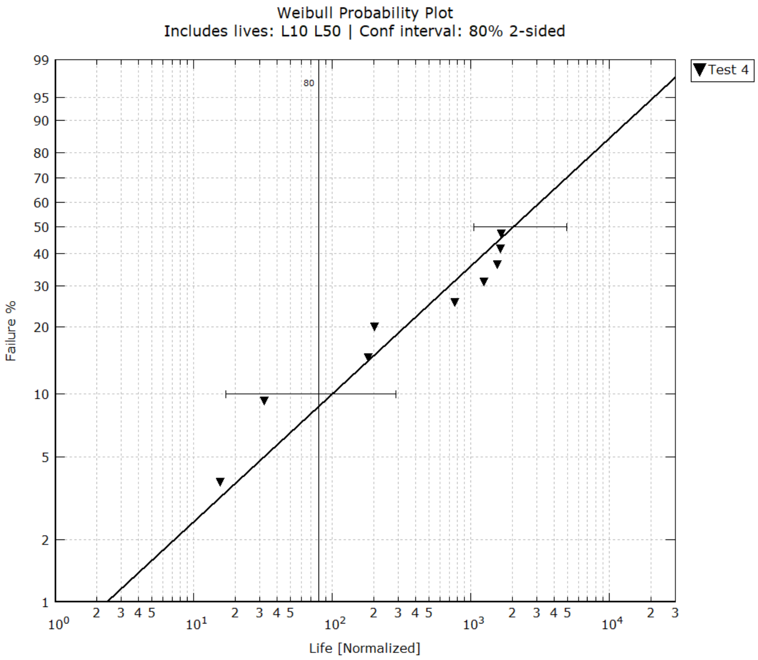


Figure 3 - Test 4 (Normalized / low beta / less failures)

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.

Extrapolating Rolling Bearing Life Data to Very High Reliabilities: Friend or Foe

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

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

*Corresponding author: guillermo.morales@skf.com

Abstract

Bearing life is assumed to encounter randomness driven by a Weibull-like statistical distribution. Typical reliability level taken as performance characterization of rolling bearings is the L_{10} life (10% of failure among a large bearings population). Business wise, quantitative estimates of higher reliability levels (below L_{10}) are of importance for an increasing number of applications, thus they need then to be investigated further. This article aims are twofold (i) to describe the various ways to give quantitative estimation of high reliability levels depending on the available information (test data, prior knowledge), (ii) the level of confidence needed together with the estimation techniques (extrapolation, confidence bounds) and the statistical model applied for the life distribution (Weibull 2, Weibull 3). Practical recommendations are also derived to offer guidelines and limitations when confronted to either realistic (size-wise) data sets or extrapolation requests from standard L_{10} calculations.

Bearings, Bearing Life, Reliability Analysis, Bearing Testing, Weibull Analysis

1. Introduction

Rolling bearings are machine components that are subjected to rolling contact fatigue (RCF) which is a type of very high cycle fatigue (VHCF). They might reach their end of life with damage originated from surface or sub-surface mechanisms [1]. It is well known that seemingly equal bearings running at seemingly equal operating conditions in a machine can produce very different individual lives, this is because small variations in the material micro-structure, geometry and surface finishing from manufacture, particles in the contact or small variations in the operating conditions can have a very large effect in the performance of the individuals. This is why the bearing life of an individual rolling bearing is considered as a random variable [2]. Pioneering and recent bearing life models used in industry [1, 3, 4, 5, 6, 7] apply a combination of physical principles (RCF, Tribology) and statistics, usually a Weibull statistical model [8]. These models attempt to predict the bearing life of populations of seemingly equal bearings under seemingly equal operating conditions with a certain reliability value. Thus the L_{10} life is the life that 90% of a large population of bearings will achieve (also named as 90% reliability).

It has been demonstrated in the past [3, 4, 9] that a good reliability level that can be verified accurately with endurance testing of rolling bearings in a frame of good economic sense, is indeed the 90 % reliability. The standard ISO 281 in its 1999 version included only up to 99 % reliability values. But in the 2007 version [7] this was increased until 99.95 %.

In a wider machine design perspective, so far only rolling bearings are designed considering quantitative reliability levels. Other machine components like gears or cam-followers do not yet benefit form a physics-probabilistic life calculation, only recently this idea for gears was again revived [10]. However the idea of using much higher reliability than 90% for bearings as response of complete system failures in the field could be dangerous. Since this is based on faith in the accuracy of higher reliability calculation values stated in the ISO 281 [11]. This aspect is already open for consideration in system design standards [12]. Therefore, it is important to investigate the accuracy and correctness of extrapolating reliability factors obtained in medium reliability values for the life of rolling bearings to very high reliabilities.

The article is structured as follows. Section 2 gives the background on the high reliability factor used in the ISO. Section 3 computes the precision of the different methods, also assessing their robustness towards assumptions. Section 4 gives practical recommendations in case a high reliability is requested. Details on the different statistical distributions used to model bearing life are given in Appendix A. Extrapolation factors that could be used for high reliability levels are discussed in Appendix B.

1.1. Objective of the Present Article

To investigate the accuracy and correctness of extrapolating reliability factors obtained in medium reliability values for the life of rolling bearings to very high reliability levels. Where quantitative boundaries for such extrapolations are derived, concrete recommendations are set and comparisons between different statistical distributions are done. This article can then serve as rules for such extrapolations based on statistical analysis and extensive Monte Carlo simulations. In the current literature, this aspect has been neglected and therefore, this aspect is novel and makes the contents of this article critically important.

Nomenclature:

Notation	Definition
a_{ISO}	Life modification factor, based on a systems approach of life calculation
L	Life random variable
η	Weibull statistics scale parameter
β	Weibull statistics shape parameter
L_p	General life percentile
L_{10}	10^{th} Life Percentile
L_0	Minimum life
α	Ratio L_0/L_{10}
$L_{10,X}$	X^{th} confidence bound on L_{10} (X% chance to have the true L_{10} below)

2. Derivation of the Reliability Factor for Life in Rolling Bearings

The ISO/TR 1281-2 [7] describes in more detail the introduction of the reliability factor for life calculation in rolling bearings called a_1, which is applied in the modified bearing life equation as follows:

$$L_p = a_1 a_{ISO} L_{10} \tag{4}$$

Allowing for the calculation of the L_p life (bearing life with $S = 100 - p$ [%] reliability) from the L_{10} life value (bearing life with 90% reliability). A table of values for a_1 respect to S is given up to value of $S = 99.95$, thus up to $L_{0.05}$. This represents very high reliability. Up to what point this is still valid or accurate? In this paper, answers to these questions are explored. For that it is necessary to understand the derivation given in [7] for this parameter.

Starting from a 3-parameter Weibull distribution (see Appendix A):

$$\frac{L_p - L_0}{L_{10} - L_0} = \frac{\left[Log\left(\frac{100}{100-p}\right)\right]^{1/\beta}}{\left[Log\left(\frac{100}{90}\right)\right]^{1/\beta}} \tag{5}$$

Based on the extensive data pooling presented in [9], a minimum life L_0 (life achieved with 100% reliability) is assumed in [7] with $L_0=\alpha \times L_{10}$. Equation (5) becomes then

$$\frac{L_p - \alpha \times L_{10}}{L_{10} - \alpha \times L_{10}} = \frac{\left[Log\left(\frac{100}{100-p}\right)\right]^{1/\beta}}{\left[Log\left(\frac{100}{90}\right)\right]^{1/\beta}} \tag{6}$$

And from (6), the ratio L_p/L_{10} can be derived:

$$\frac{L_p}{L_{10}} = a_1 = \alpha + (1-\alpha) \times \frac{\left[Log\left(\frac{100}{100-p}\right)\right]^{\frac{1}{\beta}}}{\left[Log\left(\frac{100}{90}\right)\right]^{\frac{1}{\beta}}} \tag{7}$$

From equation (7) and the extra assumption that the shape parameter β equals 1.5, the reliability factor a_1 is obtained in [7]. Therefore the life L_p becomes:

$$L_p = a_1 L_{10} \tag{8}$$

Notice that this derivation (from [7]) requires that L_p follows closely a 3-parameter Weibull distribution with β equals 1.5. In addition, [7] presents the calculation for 2 values of α, namely 0 and 0.05 and the final ISO 281 [11] applies α =0.05. Whether these assumptions are in general valid or not, can be questioned.

Indeed, the latter is essentially based on the extensive pooling (2520 bearings tested with 2230 failures) presented in [9]. In this reference [9], the reliability plot shows a bending respects to the straight lines (Weibull 2-parameters). This bending leads to an assumed L_0 around 0.004, while L_{10} is estimated as 0.1, the factor 0.05 is an approximation of the ratio 0.004/0.1 (using the values as in [9] or 0.4/10 as in Figure 1 with a different normalization). But, the final bending in [9] depends only on 2 failures (out of 2230). Later on, further pooling of test data [13, 14, 15] were in line with [9] (see Figure 1 with normalized pooled data). Nevertheless, these references are more than 30 years old and the increased performance of bearings and steels may affect different reliability levels in different ways. Typically, a better steel will improve the overall performance of the population (modification of a_{ISO}) but not the very early failures in the same range. The 0.05 factor may then become larger. Extensive pooling of data shown in Figure 3 demonstrates that the Weibull 3-parameters assumption may not be valid in general.

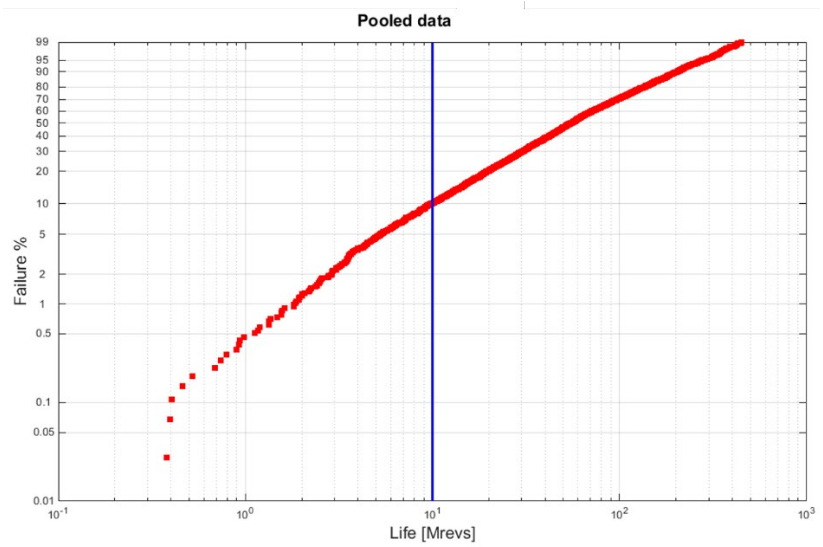


Figure 1 - Normalized pooled bearing data reproduced from [9]

Appendix B offers an overview of the extrapolation method for the Weibull distributions and two values of the β parameters (1.5 and 1.1, corresponding the 2 classical assumed values for bearing lives).

3. Quality of the estimation of high reliability levels

Here some details are given regarding the precision of the estimations and the effect of the parameter settings. In the first subsection the Maximum Likelihood method is taken into account and quantitative evaluation on the robustness of target parameters is provided. In the second subsection the effect of deviations in the distribution parameter is quantified.

3.1. Precision of the estimations

Any parameter estimation comes with a confidence interval. This depicts 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. This means that for 90% of such confidence intervals, the true value of L_{10} will lie inside. For example, with $L_{10,5} = 100$ Mrevs and $L_{10,95} = 800$ Mrevs, it means that the true L_{10} has then 90% chances to lay between 100 and 800. The precision of such intervals is measured via the ratio between the upper and lower bounds. For example, with $L_{10,5} = 100$ Mrevs and $L_{10,95} = 800$ Mrevs, we get a precision of 8 (800/100) while with $L_{10,5} = 300$ Mrevs and $L_{10,95} = 600$ Mrevs, the precision improved to a factor 2 (600/300). Although 5 and 95 are classical confidence levels, 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 precision depends on many factors, mostly the following ones:

i. Fit between data and model

ii. Sample size

The first factor is related to the discussion about the extrapolation (see Appendix B) since all the models have a scope of applicability. Weibull models are historically targeting L_{10} estimation. Moreover, the failure modes accounted in the early life might be different from the one encountered around the L_{10} (see Section 3.2). Therefore, the fit between the data and the model can be poor. The sample size is also a key issue in the high reliability precision. The width of confidence interval is directly related to this size. For Gaussian models, the precision increases at a speed of order $\sqrt{n}$ where n is the sample size. It means that the confidence intervals width decays following $1/\sqrt{n}$. For Weibull models, the situation is different and there is no formula for that width. Nevertheless, Monte Carlo simulations have been run to evaluate this precision speed increase on tests of n bearings up to ($0.2 \times n$) failures. The result is that the interval width (ratio between the upper and lower bounds of $[L_{10,5}, L_{10,95}]$ confidence intervals) decays towards 1 in a less predictable way. Namely, the logarithm of the ratios decays to 0 as $1/n^{1/3}$. The latter decay has been obtained by extensive Monte Carlo simulation similar to the one in Figure 2 with increasing values for the sample size n leading to the above mentioned decay. This being valid also for L_5 and L_1. This decay being independent on β. Figure 2 is showing the impact of the sample size onto the width of the confidence interval. The data used to build Figure 2 are randomly generated (Monte Carlo simulations). For the sake of clarity, samples have been scaled at different decades to avoid superimposed intervals, only width of intervals have to be seen from Figure 2.

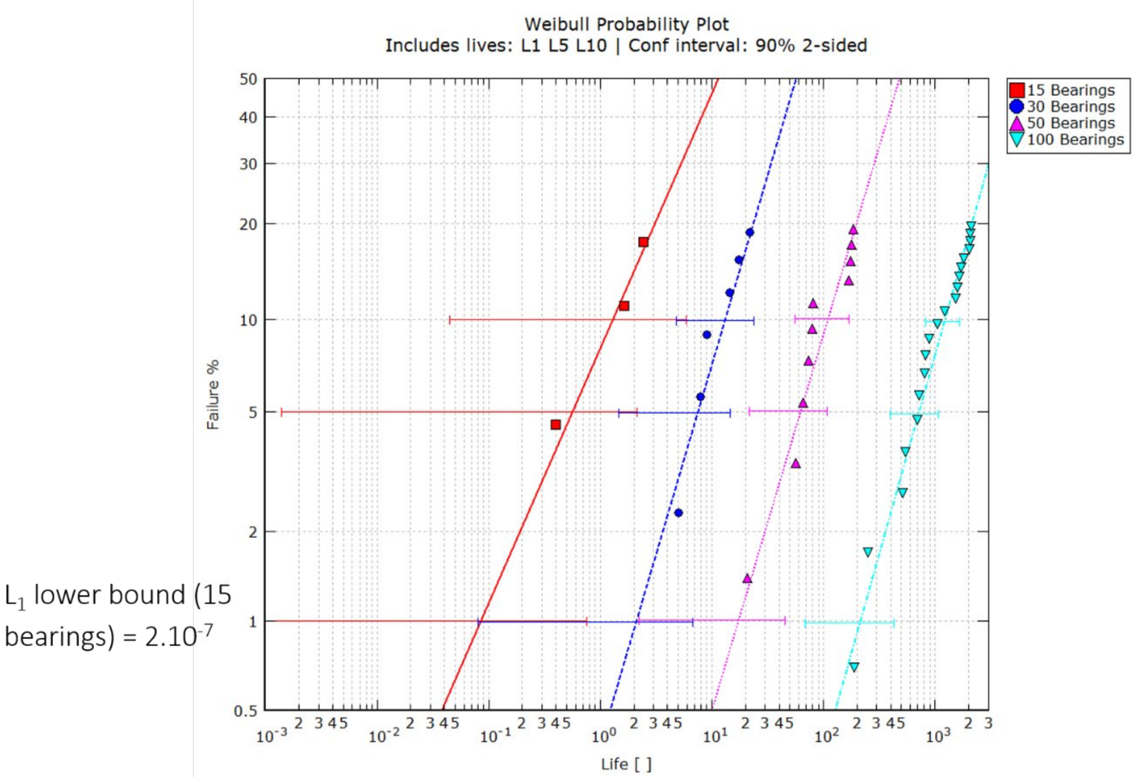


Figure 2 - Width confidence interval on L1, L5 and L10 and their evolutions with sample size (all samples censored at the time of the X^{th} failure where X = 20% of the sample size)

Figure 2 clearly shows the high risk of a too high reliability level estimation, like the L_1 parameter which keeps an uncertainty of factor 100 at a sample size typical for a life test (30 bearings).

From Figure 2 and extended Monte Carlo simulations, minimum sample sizes can be derived for life percentiles L_1, L_5 and L_{10}. The Monte Carlo simulations were done, for each sample size, with 10,000 runs each. The large number of runs ensures stability of the results. They were using parametric (Weibull) random generator at the same Beta value (1.1, typical of bearing life). The input L10 value was taken as 1 since it is only a scale factor. These minimum sizes are divided in two sets, a strict minimum size and a recommended one offering better robustness. The recommended sizes correspond to a high probability for the confidence interval to have a ratio less than 10 between its upper bound and lower bound. This ratio 10 is coming from long time experience in life testing where performance comparison needed at least one decade to be conclusive. As for the minimum number of failures, the ratio of 20% of the sample size taken in Figure 2 must be kept to achieve enough failures and then a good fit with the Weibull statistical distribution. This value (20%) is chosen to ensure failures below and above the target life percentile L_{10} corresponding to 10% failure.

Reliability level	Min. Sample Size	Min. No of Failures	Recom. Sample Size	Recom. No of Failures
L_1	100	20	200	40
L_5	40	8	50	10
L_{10}	20	4	30	6

Table 1 – Rules on minimum sample size and number of failures for different target reliability levels

The results in Table 1 show clearly that in order to have a good accuracy in reliability levels of L_1, a substantially higher number of failures is required, which means that a much greater number of bearings need to be tested in comparison to more conventional reliability levels like L_{10}. In practice this does not have economic sense, thus it is not common practice.

3.2. Treatment of experimental data

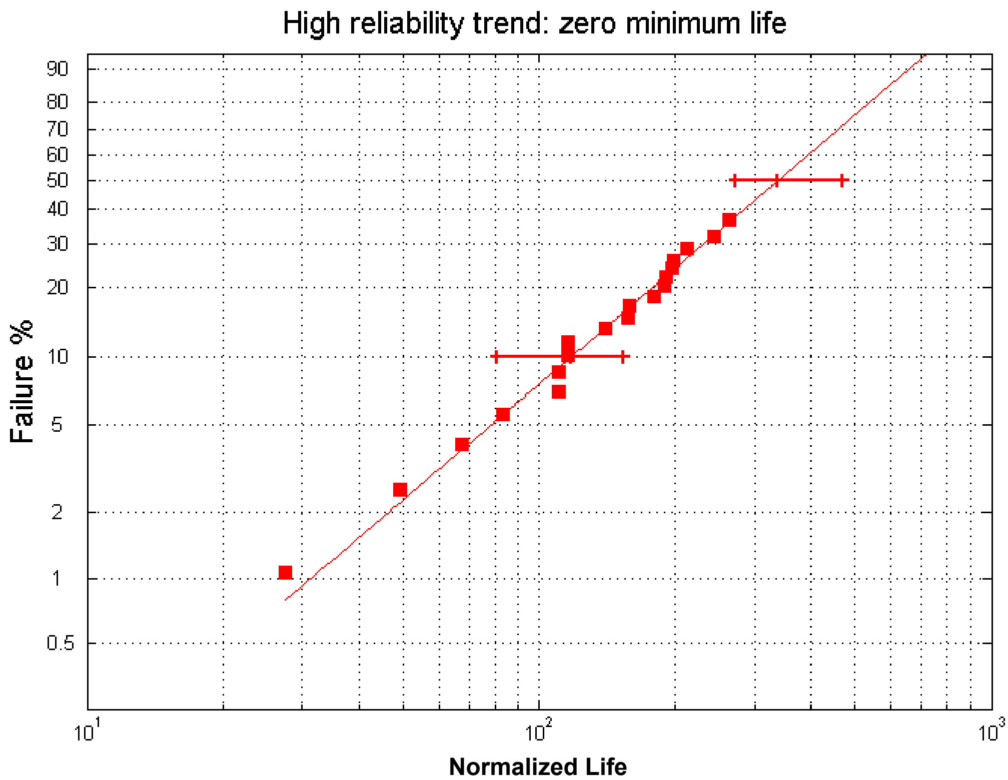
As seen in Section 3.1, the precision and accuracy of the statistical estimation of the life parameters (whatever choice has been made for a life model) is very sensitive to the sample size. Therefore, it is valuable to extend the sample size as much as possible. When dealing with test data, a solution is to pool these data. Such a pooling needs to be done carefully to guarantee homogeneity among the individual data sets. The pooling should then stick to one of the following two cases:

i. Pool data coming from tests of the same product of same size under similar conditions.

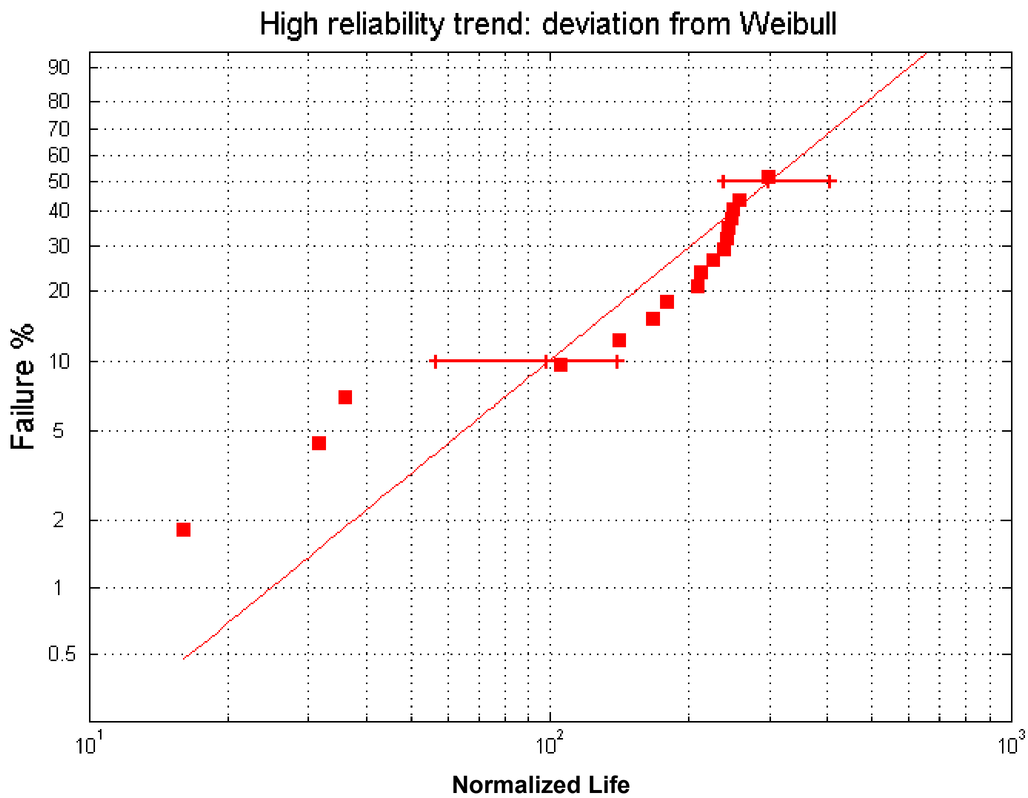
ii. Pool data coming from tests of the same product under scalable sizes and conditions.

By "scalable", the authors mean that the lives can be compared via a multiplication factor (same β in the case of 2-parameters Weibull data). In such a case, a physical model (or a prior knowledge) can help to develop this multiplication factor. For example, rolling bearings under the same contact pressure but different geometries or bearings with the same geometry and same contact pressure but different sizes and loads. This can typically be the case for a size effect within a unique size range (e.g. two different medium sizes). Then, one size and condition are chosen as reference and the data coming from other sizes and / or conditions are rescaled according to a multiplication factor. Then, the data are pooled together with the reference. Such a pooling is of great interest for the quality of the life parameter estimation but it relies on the identical β assumption and the multiplication factor chosen. For instance, a difference of 10% in β between two pooled test samples may lead to an error (lack of accuracy) of 5% on L_{10}, 10% on L_5 and 20% on L_1 once these reliability levels are estimated from the pooled data.

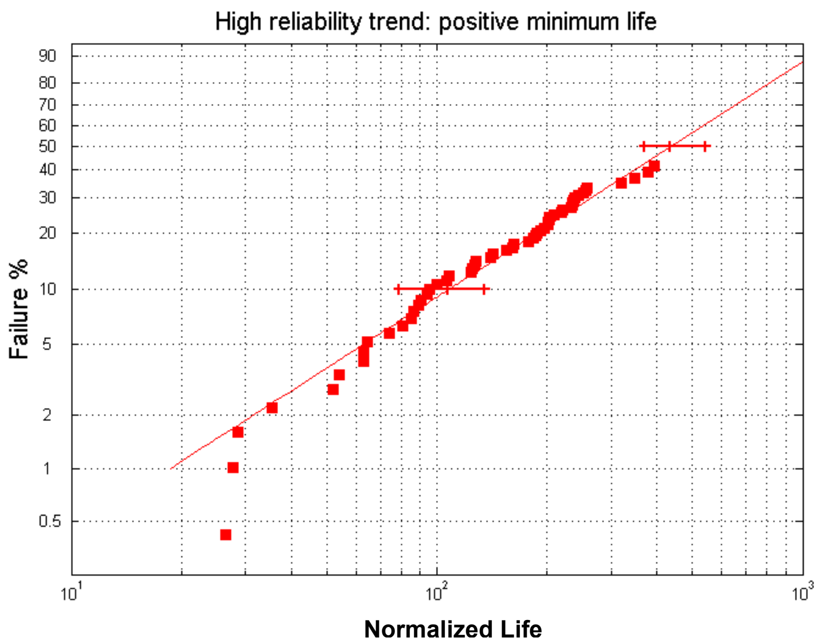
Pooling test data is common practice in bearing endurance testing, extensive pooling of test data has been done in the past [9] and the authors have performed new pooling here (From large amount of in-house endurance tests on CRB's, SRB's and TRB's). Figure 3 shows three examples of pooled experimental data (for the 3 different bearing types) leading to very different behaviors for the early failures.



(a) Case 1



(b) Case 2



(c) Case 3

Figure 3 – Various trends at high reliability with endurance test data performed in-house with three different bearing types.

The 3 cases of Figure 3 (a), (b) and (c) correspond to extensive pooling of actual bearing data for 3 different bearing types, they cover around 100 samples allowing for high reliability level estimation even for L_1.

Nevertheless, the first failures show 3 different trends towards these high reliability levels (L_1). This behavior shows the risk of assuming a theoretical Weibull distribution (2 or 3 parameters) when the reality might be more complex. The potential deviation between the assumed statistical distribution and the actual one will lead to severe error in the estimation with high reliability levels. Such error for L_1, is illustrated in Figure 3 where the plot (c) fits with a 2-parameter Weibull, the plot (b) fits with a 3-parameter Weibull and the plot (a) has no known theoretical fit.

Figure 3 also illustrates clearly the high risk of having any estimation of reliability levels beyond L_1. Indeed, the discrepancies observed at the L_1 levels can only increase when moving further into higher reliability levels.

4. Good Practice Principles

Next, based on the results shown in this paper some good practice recommendations are given in order to statistically assess bearing life from endurance testing results for medium and high reliability levels.

Limitations

- If the test target concerns the L_{10} or higher reliability levels, then it is recommended to stop the test after maximum 50% of the tested item has failed or has been suspended (for sudden death tests, an item is a full group of individuals): above the L_{50} the distribution of failures may deviate from any of the Weibull models and late failures will then affect the accuracy of the estimation.
- Without experimental evidence, the comparison between the failure modes at stake for the high reliability levels and the ones at stake for classical levels (like L_{10}) may not be valid or weakly linked. This can be illustrated by the classical examples of human life where the reason to die around the L_{10} (~ 60 years) is very different from the reason to die around L_1 (~15 years). To that respect, the Weibull 2-parameters offers the most conservative approach among the Weibull distributions.

Size of the data set

- The strict minimum sample size for testing is 20 for the L_{10} estimation, 40 for the L_5 estimation and 100 for the L_1 estimation. At these levels the variability of the results from the simulation is still high. The

recommended minimum size to avoid these variability leads to the following recommended minimum size: 30 for the L_{10} estimation, 50 for the L_5 estimation and 200 for the L_1 estimation.
- Ensure to reach at least 20% failures among the sample set in the test (to ensure a proper fit with a Weibull distribution).
- Pooling data sets is often necessary but it brings extra noise from potential variation in β introduced in the different tests. Only tests with close operating conditions should then be pooled to limit the risk of having different β.

Statistical model
- Use the Exponential model when a strong prior knowledge is giving reliable guess for the β value. Always use the lowest available assumption for β.
- Use the Weibull 3-parameters model for L_1 estimation in cases of very large sample sizes (at least 200 tested items, still with 20% of the sample to failure).
- Use the Weibull 2-parameters model in any other cases. Especially when no test data can give information on the early failure at stake for high reliability levels, the Weibull 2-parameters distribution offers the safest (more conservative) approach for the same β value.

Extrapolation

In the case where no robust statistical analysis can be achieved (too high reliability expectations with respect to the limited data available), extrapolations can be derived from estimated lower reliability levels (L_{10} for instance):

- Although theoretically any level can be calculated, the statistical robustness of the L_1 level is proven to be very hazardous. Therefore, no reliability level beyond L_1 can be recommended statistically. In particular L_0 must stay a pure theoretical parameter since no 100% reliability can never be guaranteed.
- Extrapolation from lower levels than L_{10} (like L_{20} or L_{50}) is to be avoided to limit the risk of artificially linking uncorrelated failure modes.
- When extrapolating the lower bound of the confidence interval the lowest assumed β value ($\beta = 1.1$ by default) is to be used.
- Estimations of reliability level higher than L_5, always have associated liability risk.

System life

The system life is a special case where high reliability levels are needed not much for the individual components but rather to achieve usable moderate reliability levels for the entire system.

In addition to the above recommendations, a dedicated analysis of the whole system (like a FMEA – Failure Mode and Effects Analysis) can help enlightening the dependencies between the sub-systems. Otherwise, extensive tests are needed, however sometimes only extrapolations are feasible. As for the extrapolation, a less conservative approach for the β value can be used ($\beta = 1.5$ typically). This less conservative extrapolation should only be used for the system life and should not be used for high reliability level for an individual sub-system.

Benefit of high reliability level estimations

Although most of the content of this article is aiming at giving practical limitations to the use of life extrapolation estimation of high reliability levels can also be beneficial once used properly.

- When extensive field data is available (getting then information on potential early failure modes), high reliability estimations can be obtained safely from this field data directly
- When field data allows to exclude the type of deviation illustrated in Figure 3a, assuming a Weibull 2-parameters with a low beta (1.1) appears to be a safe approach so that extrapolation can be possible from test data and the corresponding L_{10} estimation.
- Once established, such high reliability level estimations can serve tracking deviation in an application. Indeed, any early failure occurring before a high reliability value should serve as warning sign that deviations in the application (operating conditions, mounting…) are taking place.
- High reliability levels are also useful and even strongly necessary to derive estimation of system life, as explained in Appendix A.

5. Discussion

Estimation of high reliability levels is very challenging via standard testing. Although different statistical distributions are available their performance is always limited to a minimum number of data. In this paper, quantitative statements are given for this minimum number of data together with a comparison among the main distributions. From this comparison, the standard Weibull 2 parameters appears to be the most robust unless extensive testing data are available.

Apart from the choice of the statistical distribution, a key limitation for the use of high reliability levels is the risk to derive estimates based on tests which produce different failure modes from the ones in control of those high reliability levels. There, the Weibull 2-parameters offers also the most conservative approach.

A short description of the main statistical aspects of bearing life estimation from endurance tests has been given. The objective is to assess whether or not (and why) extrapolation to high reliability levels in bearing life estimation can be dangerous. From the analysis presented here, it is clear that the accuracy of the estimation decreases with the increase of the reliability level when using a fixed number of tested samples. Very high reliability levels require very high number of tested samples, which becomes economically prohibiting. Good practices have been revisited in order to minimize the risks to promote severe mismatches between the estimation of high reliability levels and true values. A potential solution to safety use extrapolation methods, like the one promoted in the ISO 281 standard [11] is to apply it together with the extrapolation factors described in Table 2 (Appendix B).

The potential weakness of the extrapolation comes from the need to assume a beta (β) value and also the assumption that the Weibull fit (2 or 3 parameters with a fixed ratio α between L_0 and L_{10}) stays valid even towards high reliability levels. The sensitivity to the fixed values α and β weakens the accuracy of the extrapolations. The high reliability levels do not come directly from actual data but appear as a function of the L_{10} parameter. Typically, a wrong assumption on β has an important effect onto the result. For instance, an error of 10% on β, in the case of a 2-parameters Weibull, leads to an error up to more than 20% on the L_1 parameter. The precise percentage depends on the β value and varies from 24% to 11% when β varies from 1 to 2. Likewise, an error of 20% on the true β value leads to an error on L_1 from 48% to 22% when β varies from 1 to 2. It needs to be added that such extrapolation should never be done starting from lower reliability level than the L_{10}, for instance from the L_{20} or the L_{50}. Then, the sensitivity to the β slope becomes inapplicable.

A second key issue with such extrapolations is the validity of the statistical model towards the tail (L_0). Even if experimental data fit well with one of the chosen statistical distributions around L_{10}, whether this model fits with the reality down to the L_5, L_1 and further or not is a completely different issue. These high reliability levels correspond to early failures that could derive from a different physical mechanism than the later failures.

A practical and illustrative example can be given about such extrapolation for L_1 (assumed to be equal to 0.25 x L_{10} in [11]). Consider the pooled experimental data illustrated in Figure 3. Even if the beta slope is matching well with the assumption made in [11] (β~1.5), the left hand side plot (a) (reproduced in Figure 4) shows a clear discrepancy between the extrapolation rule $L_1 = 0.25$ x L_{10} and the behavior of the early failures. The value 0.25 x L_{10}. corresponds more to the L_4, which could create a severe liability issue with a failure rate 4 times bigger than expected. This example illustrates also the risk of having early failures off respect of the main Weibull distribution.

An adding argument against the undifferentiated use of the extrapolation rule from [11] is related to the beta parameter. Although [11] assumes a 3-parameters Weibull and a value β=1.5, it is proven not to be the case under all operating conditions. From [16], a 2-parameter Weibull with β~1.1 is better suited. From Table 3 the corresponding extrapolation factor should be $L_1 = 0.12$ x L_{10}. In this case the value 0.25 x L_{10}. corresponds more to the L_2, which could create a liability issue with a failure rate 2 times larger than expected.

Those two concrete examples illustrate the industrial risk taken when using carelessly the high reliability extrapolation rule from [11]. If instead of 1.5, a Beta of 1.1 would be considered, a more conservative extrapolation could be derived as shown in Appendix B.

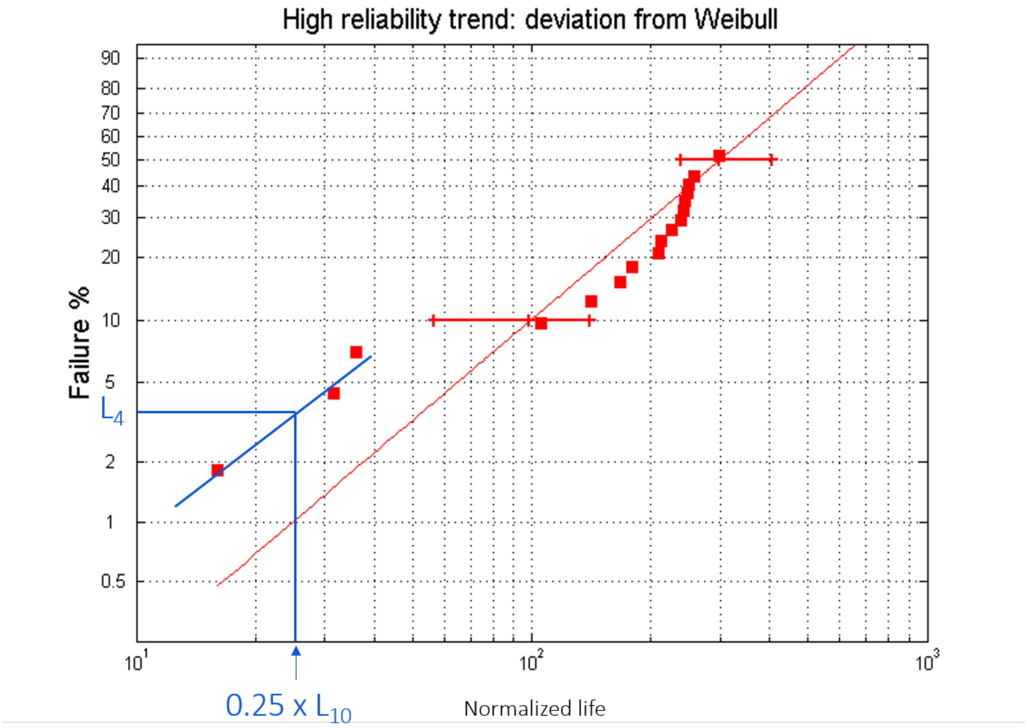


Figure 4 - High reliability deviation with endurance test data performed in-house on CRB.

6. Conclusion

In general, extrapolation from more moderate reliability levels (like L_{10}) is possible but can potentially lead to liability issues. Such extrapolation needs to apply the lowest assumed β value (β = 1.1 by default, β = 1.5 for system life).

This is not in contradiction with extensive studies proving a better fit with the Weibull 3 parameters distribution (and therefore higher values for L_1), but it highlights the need to have a conservative approach based on limitations of the estimation techniques when confronted to realistic test sample sizes. Indeed, the Weibull 3-parameters requires too large sample sizes to be practically used with confidence and may lead then to an overestimation of the high reliability levels as mentioned in Table 2 and Table 3.

The main conclusions from this investigation are:

- Extrapolation or estimation of reliability levels higher than L_1 [following discrepancies already observed at L_1 level - Figure 3] are hazardous and leading to severe industrial risks.
- Robust estimation of high reliability levels from test or field data requires a minimum number of data (200 for L_1, 50 for L_5) with 20% of the sample sets corresponding to failures [see Table 1]
- Extrapolation rules from the L_{10} estimation or L_{10} calculation from an established life model must use a 2-parameter Weibull assumption with a low Beta (1.1) [see Table 3, Appendix B]

Acknowledgment

The author wishes to thank Mr. Bernie van Leeuwen, Director SKF Research and Technology Center for his kind permission to publish this article.

Bibliography

[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 distributrion function of wide applicability," *J. Appl. Mech.,* vol. 18, no. 3, pp. 293-297, 1951.

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

[10] G. Morales-Espejel, "A model for gear life with surface and subsurface survival: Tribological effects," *Wear,* Vols. 404-405, pp. 133-142, 2018.

[11] ISO281, "Rolling Bearings, Dynamic Load Ratings and Rating Life (explanatory notes)," *International Standard,* 2007.

[12] IEC, " IEC 61400-4 -Wind Turbines - Part 4: Design Requirements For Wind Turbine Gearboxes," International Electromechanical Commission (IEC), 2019.

[13] B. Snare, "How Reliable are Bearings?," *Ball Bearing Journal,* vol. 162, pp. 3-5, 1970.

[14] G. Bergling, "The Operational Reliability of Rolling Bearings," *Ball Bearing Journal,* vol. 188, pp. 1-10, 1976.

[15] H. Takata, S. Suzuki and E. Maeda, "Experimental study of the life adjustment factor for reliability of rolling element bearings," in *Proceedings of the JSLE International Tribology Conference, 1985-07-08/10*, Tokyo, Japan.

[16] E. Zaretsky, "Rolling Bearing Life Prediction, theory and applications," *NASA/TP,* 2013.

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

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

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

[20] W. Harper, T. Eschenbach and T. James, "Concerns about maximum likelihood estimation for the 3-parameters weibull distribution: case study of statistical software," *The American Statistician,* vol. 65, no. 1, pp. 44-54, 2011.

A Appendix - Bearing life Statistical models

The Weibull statistical distribution is often used to model the randomness of physical phenomenon like fatigue of materials or mechanical product life. It has been introduced in the setting of material strength by Waloddi Weibull [2] and extended to a wide range of experimental data [8]. The Weibull distribution is widely used together with its special case, the exponential distribution. The Weibull distribution itself possesses two main forms, one with 2 parameters and one with 3 parameters. The 2 parameters Weibull distribution turns into an exponential distribution when the shape parameter β equals to 1.

In addition, a new Weibull-based distribution has been recently introduced [17] which allows having a non-zero minimum life (life reached with 100% probability) that could be statistically estimated using the maximum likelihood method.

The purpose of the current section is to present the 3 standard statistical distributions, their scope of applications and their restrictions. In all 3 definitions, L denotes the random variable standing for the Life duration. The distributions are given with their two most common expressions, the mathematical form with η (or λ) as a scale parameter, and the 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 population will not survive. Equivalently, L_p is the time that $(100 - p)$% of a large population will survive.

A.1 Exponential

Weibull 1-parameter corresponds to the exponential distribution:

$$P(L > x) = exp\ (-\lambda x)\ with\ \lambda > 0 \quad \text{(A1)}$$

This is a special case of the Weibull 2 parameters distribution with β=1, but it can serve to study the case of a Weibull 2-parameters on which the slope parameter is fixed at a known value.

A.2 Weibull 2 parameters

Weibull 2-parameters Weibull distribution:

$$P(L > x) = exp\left(-\left(\frac{x}{\eta}\right)^{\beta}\right) = 0.9^{\left(\frac{x}{L_{10}}\right)^{\beta}} \text{ with } \eta, \beta, L_{10} > 0\ ; \quad \text{(A2)}$$

The Weibull 2-parameters is the most used distribution for bearing life. It combines high flexibility due to its two parameters while keeping a simple expression. This simplicity makes it possible to develop statistical estimation techniques with proven accuracy and precision [18, 19, 17]. Past [9] and recent (Section 4.3) extensive pooling of test data gives further evidence for the matching between experimental life data and the Weibull distribution, at least in the main life span.

The main drawback of the 2-parameters Weibull model is the absence of minimum life (minimum life is a time that all items will survive with 100% probability). Affecting mainly 2 cases:

• When high reliability level is required for a critical application, having a zero minimum life may lead to a too conservative estimation for reliability levels strictly higher than L_{10}.

• In a mechanical system with several bearings, the whole system life is severely affected by the absence of minimum life. Considering a system with 10 identical bearings in series (the system fails as soon as one bearing fails), a weakest link approach makes the L_{10} life of the system close to the L_1 of an individual bearing. Therefore a too conservative estimate of this L_1 has a strong consequence of the L_{10} estimation of the system. The system life is calculated as follow:

$$L_{10}\ (System\ of\ N\ identical\ bearings) = \frac{L_{10}(individual\ bearings)}{N^{1/\beta}} \quad (A3)$$

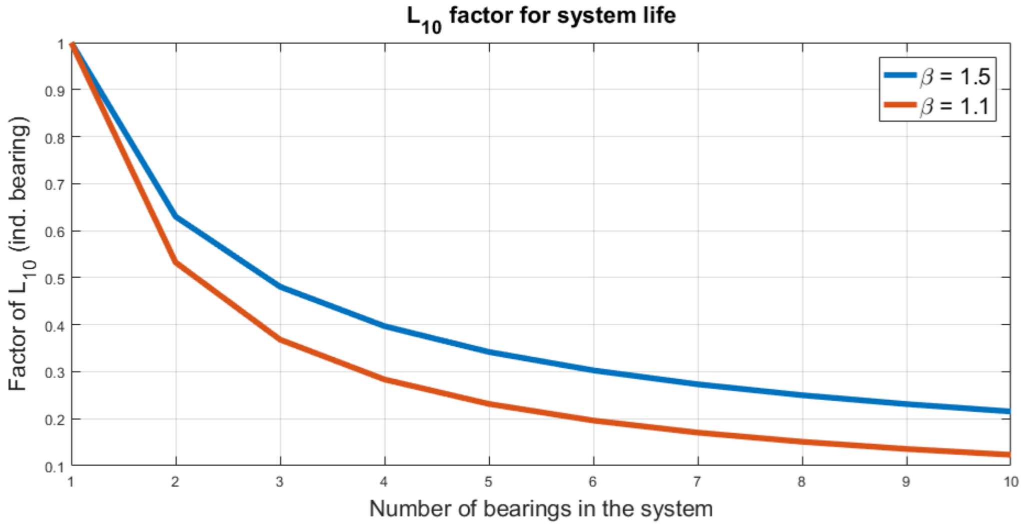


Figure 5: System life (L10 of the system) as a factor of the L10 value of an individual bearing and its dependency on the number of bearings in the system (using equation Eq A.3)

Figure 5 depicts the behavior of equation A3 for two different values of β and with increasing number of bearings (n) in the system.

The individual bearing life level corresponding to the L_{10} (Syst.) becomes

$$100 \times \left(1 - 0.9^{\left(\frac{L_{10}(Syst.)}{L_{10}(Ind.)}\right)^{\beta}}\right) = 100 \times \left(1 - 0.9^{1/n}\right) \quad \text{(A4)}$$

which neither depends only on n, β nor L_{10}. The exponential decay is illustrated in Figure 6.

Please notice that this figure shows the theoretical usefulness of obtaining knowledge on the L_1 or even higher levels for an individual bearing in order to estimate the L_{10} of a system. Another practical solution is to study in details the system and its interdependency between bearings. If some dependency can be proven, this will have a significant impact on the system reliability because a failure root cause will then be counted only once. The drawback of such analysis is that it must be made case by case through a FMEA (Failure Mode and Effect Analysis).

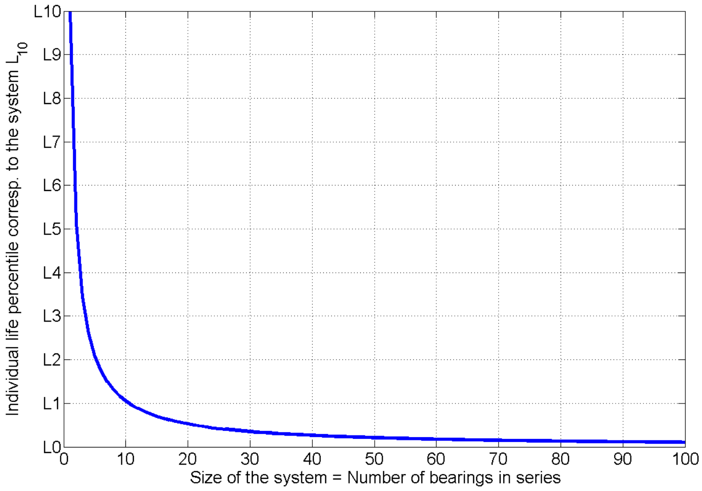


Figure 6. Individual life percentile corresponding to the system L10 life, as calculated from Eq. A.4

When prior knowledge allows to assume a fixed β , the Weibull 2-parameters can be turned into an exponential distribution (in order to use the same estimation techniques). Knowing β, the random variable L^{β} follows an exponential distribution with scale parameter:

$$\lambda = \frac{1}{N^{\beta}} = \frac{\log(0.9)}{L_{10}^{\beta}} \quad (A5)$$

A.3 Weibull 3 parameters

Weibull 3-parameters distribution:

$$P(L > x) = exp\left(-\left(\frac{x-L_0}{\eta}\right)^{\beta}\right) = 0.9^{\left(\frac{x-L_0}{L_{10}-L_0}\right)^{\beta}} \text{ with } \eta, \beta, L_{10} > L_0 \geq 0 . \quad (A6)$$

The Weibull 3-parameters is similar to the Weibull 2-parameters. The extra parameter L_0 offers a great flexibility for high reliability levels and therefore a better fit can be obtained between extensive experimental data and Weibull 3-parameters curves. The main drawback of this distribution is the absence of known bias correction techniques for the life parameters and the lack of robustness of the commercial Maximum Likelihood techniques [20]. This point stays valid even if the ratio L_{10}/L_0 is fixed (having then two unknown parameters). In practice, the L_0 parameter is estimated via curve fitting (or equivalent methods) and once L_0 is fixed, data can be treated as shifted 2-parameters Weibull for which standard unbiased estimation can be performed. Such an approach relies strongly on the accuracy of the L_0 estimation.

B Appendix. Extrapolation methods for the estimation of high reliability levels

The most usual life percentile that is accurately estimated is the L_{10}. A way to obtain higher reliability estimation is via extrapolation of the L_{10} one. This is done by fixing all the model parameters once L_{10} is known. This is the method used in the ISO 281 [11, 7]

The extrapolation factor is defined as the ratio between the target reliability level and the reference one (usually L_{10}). Table 2 and Table 3 present a matrix with the extrapolation factors (obtained from Eq. 7) for the Weibull distributions for a specific choice of parameters: the scale factor L_{10} is fixed at 1, the slope β fixed at 1.5 (Table 2) and 1.1 (Table 3) and $\alpha = 0.05$.

In terms of reliability levels, the list from Table 2 and Table 3 is taken from ISO 281 [11]. Although theoretically any level can be calculated, the statistical robustness of the L_1 level will already be proven to be very hazardous (Sections 3). Therefore, no reliability level beyond L_1 can be recommended statistically. In particular L_0 must stay a pure theoretical parameter since no 100% reliability can be guaranteed.

From a practical point of view, if no complete statistical assessment can be given (lack of data, different early failure mode, etc), it stays possible to extrapolate calculated L_{10} from life model or confidence intervals using an assumption for β. As proven in Sections 3, their sensitivity to the chosen value for the fixed parameters is even higher. Therefore such extrapolation is not a safe process and whenever needed the most conservative assumption for β should be taken (low $\beta = 1.1$).

Reliability level	Weibull 2	Weibull 3
L_{10}	1	1
L_5	0.62	0.64
L_2	0.33	0.37
L_1	0.21	0.25
$L_{0.1}$	0.045	0.093
$L_{0.05}$	0.028	0.077
L_0	0	0.05

Table 2. Extrapolation factors towards several reliability levels using β=1.5

Reliability level	Weibull 2	Weibull 3
L10	1	1
L5	0.52	0.54
L2	0.22	0.26
L1	0.12	0.16
L0.1	0.015	0.064
L0.05	0.0077	0.057
L0	0	0.05

Table 3. Extrapolation factors towards several reliability levels using β=1.1

Exact calculation of the cumulative failure rate; Study of the four parameter Rosemann's reliability model; Suggestion of a New four-parameter reliability model.

L. Houpert[a]* and J. Clarke[b]

[a]Luc Houpert Consulting, 1 rue de Fleurie, 68920 Wettolsheim, France

[b] *Smart Manufacturing Technology (SMT), Wilford House, 1 Clifton Lane, Nottingham, NG11 7AT, United Kingdom*

*Corresponding author: luc.houpert@orange.fr

Abstract

A detailed study of Rosemann's reliability model has been conducted. This model is very flexible and uses four parameters: η (or L_{10}) and β (as in a standard two-parameter Weibull model), but also L_0 and an exponent c. When c is infinite or very large (100 for example), Rosemann's model behaves as a three-parameter Weibull model, L_0 being then a minimum life. When c=1, Rosemann's model corresponds to the two-parameter model but using c > 1 (c=2, 3 or 10 for example) allows the life to be smaller than L_0, denying therefore the existence of a minimum life. When defining randomly a number F_i, i=1 to N, varying in a uniform manner between 0 and 1, and when sorting the N values of F_i in an ascending order, one can calculate analytically or numerically the probability P of having F_i smaller than a given value F, and vice-versa, so that median value of F_i (corresponding to P = 0.5) can be obtained, as well as the values corresponding to P = 0.05 or 0.95 used for defining the lower and upper bounds of the 90 % variation range of F_i. When assigning F to the cumulative failure probability of a life, the randomly generated values of F_i can be used for simulating an experimental database and calculating the life corresponding to a given set of inputs (η, β, L_0 and c), but also understanding its 90 % variation range.

Several curve-fitting techniques (Method 1 and 2) have been developed and tested for extracting the four unknowns. Using a few examples, it has been found that the individual accuracy on L_0 and c can be poor while the final match between curve-fitted and experimental life is satisfactory when using the set (L_0,c). In this case, the model cannot then be extrapolated to very low F values. This is due to some couplings observed between L_0 and c when trying to match a set of data observed within the confidence range, set of data matched using either a too large value of L_0 compensated by a too small value of c, or vice-versa.

The latter statement has been confirmed by conducting 10,000 Monte Carlo simulations and corresponding curve-fittings for defining the median value and 90 % confidence intervals of the ratios $\frac{L_{10}}{L_{10\_cf}}, \frac{\beta}{\beta_{cf}}, \frac{L_0}{L_{0\_cf}}$ & $\frac{c}{c_{cf}}$, the last two being of particular interest.

If the median ratio is often close to 1, its confidence interval on $\frac{L_0}{L_{0\_cf}}$ & $\frac{c}{c_{cf}}$ can be large when N is small ($N \leq$ 100 for example), mainly because any values of L_0 can be accepted when the curve-fitted value of c is equal or close to 1, illustrating some redundancy among Rosemann's four parameters.

It is therefore concluded that although powerful and very flexible, Rosemann's model is not easy to use in practical situations when dealing with a reduced number of bearing failures (small N number).

An alternative a "New" curve-fitting technique and model (also using four parameters) will be suggested, their advantages being that only two simple linear curve-fitting could be conducted.

Introduction

In the context of his cooperation with the IMKT department of Leibniz University, Dr. Houpert was asked by Prof. Poll to offer some comments about a paper issued by Prof. Rosemann [1]. His paper suggests a more flexible and powerful reliability model using 4 parameters (η, β, L_0 and c described later) instead of the standard 2 (η and β) or 3 (η, β and L_0) parameter Weibull model used for example by Houpert [2] and Kotzalas [3] respectively.

Rosemann's reliability model has been further studied and the objectives of this paper are to share some results obtained, starting with a short description of Rosemann's 4 parameter model and its behavior when varying the third and fourth parameters L_0 and c especially.

When generating N random values of the cumulative failure density F ($0<F<1$) and sorting these N values of F_i in an ascending order (i=1 to N), one can calculate N values of failed bearing life t_{exp_i}, (t_{exp_i} being defined as a function of F_i and Rosemann's 4 parameters), simulating hence an endurance database corresponding to a given set of N values of t_{exp_i} defined with Rosemann's 4 input parameters.

An interesting study of F_i has first been conducted for calculating the cumulative density $P_i(F)$ as a function of F, hence the probability P_i of having the i^{th} value F_i smaller than a given F value. Novel analytical relationships of $P_i(F)$ will be given for the first 10 and last 10 values of P_i for example. As a novelty (to the authors at least), $P_i(F)$ as well as its inverse value $F_i(P)$ will also be calculated numerically using the incomplete *beta* and *inverse beta* function respectively for any of the N values of F_i so that the median values of F_i (also called median rank and corresponding to $P = 0.5$) will be compared to approximated values suggested in the literature. The median value of F_i will be used for defining the median values of t_{exp_i}. Similarly, the values of F_i corresponding to P=0.05 or 0.95 can be calculated and used for defining the 90% range of F_i, hence also the 90 % range of t_{exp_i}, quite useful information to share for understanding possible bearing life scatters as a function of N and the order number i.

The next challenge was to define appropriate curve-fitting techniques for defining the 4 Rosemann parameters and two possible approaches (Method 1 and 2) will be described using a few examples.

Knowing the confidence intervals associated to each of the four Rosemann parameters is only possible by conducting Monte Carlo simulations, conducting for example 10,000 times such a curve-fitting exercise, a task conducted by Dr. Clarke from SMT.

At the end of this paper, the authors are finally able to fully describe the pros and cons of Rosemann's model with some comment about its usefulness in practical cases.

Miscellaneous reliability models

Bearing life is usually described using a Weibull model in which the cumulative failure probability F is described as a function of the time t and two or three parameters.

F is therefore the probability of observing a bearing failure at time t.

The two parameter Weibull (unknowns η and β) distribution reads, see Houpert detailed study conducted in [2]:

$$F = 1 - \exp\left[-\left(\frac{t}{\eta} \right)^{\beta} \right] \qquad (1)$$

where η is called the characteristic life and β is the Weibull slope.

When fixing F to 0.1 (or the survival probability to 0.9), one defines the life L_{10} used for defining η:

$$\eta = \frac{L_{10}}{\left[-\ln(0.9)\right]^{\frac{1}{\beta}}} \qquad hence: F = 1 - \exp\left[\ln(0.9) * \left(\frac{t}{L_{10}} \right)^{\beta} \right] \qquad (2)$$

The three parameter Weibull (unknowns: η, β and L_0) assumes the existence of a minimum life L_0 that is always exceeded even when considering very low values of F. Its cumulative distribution reads:

$$F = 1 - \exp\left[-\left(\frac{t - L_0}{\eta} \right)^{\beta} \right] \quad with \quad \eta = \frac{L_{10} - L_0}{\left(-\ln(0.9)\right)^{\frac{1}{\beta}}}$$

$$F = 1 - \exp\left[\ln(0.9) * \left(\frac{t - L_0}{L_{10} - L_0} \right)^{\beta} \right] = 1 - \exp\left[\ln(0.9) * \left(\frac{\frac{t}{L_{10}} - \frac{L_0}{L_{10}}}{1 - \frac{L_0}{L_{10}}} \right)^{\beta} \right] \qquad (3)$$

Defining L_0 is challenging and requires analyzing a large database including very low values of F, hence a large number N of failed bearings. Kotzalas suggested L_0/L_{10} = 0.221 in [3] while a more conservative value equal to 0.05 is suggested in ISO document [4] by the ISO bearing life working committee. Obviously, L_0/L_{10} can be a function of the bearing quality, number of bearing in the database and probably bearing operating conditions, see Tallian [5], Snare [6] and Takata [7]. For the sake of simplicity, one will adopt in this paper a fixed ratio L_0/L_{10} (L_0/L_{10} = 0.2 for example) although it is difficult to justify L_0 to be simply proportional to L_{10} irrespective of the operating conditions and steel quality.

Rosemann [1] disputes the existence of L_0 and any physical discontinuities, recognizing however larger bearing life at low F values relative to the ones obtained using a two-parameter model. He suggested a more flexible four parameter model. Rosemann's general four parameter Weibull (unknowns: η, β, L_0 and c) cumulative distribution reads:

$$F = 1 - \exp\left[-\left(\frac{\left(t^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\eta} \right)^{\beta} \right] \quad with \quad \eta = \frac{\left(L_{10}^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\left(-\ln(0.9)\right)^{\frac{1}{\beta}}}$$

$$F = 1 - \exp\left[\ln(0.9) * \left(\frac{\left(t^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\left(L_{10}^c + L_0^c\right)^{\frac{1}{c}} - L_0} \right)^{\beta} \right] = 1 - \exp\left[\ln(0.9) * \left(\frac{\left(\left(\frac{t}{L_{10}}\right)^c + \left(\frac{L_0}{L_{10}}\right)^c \right)^{\frac{1}{c}} - \frac{L_0}{L_{10}}}{\left(1 + \left(\frac{L_0}{L_{10}}\right)^c \right)^{\frac{1}{c}} - \frac{L_0}{L_{10}}} \right)^{\beta} \right] \qquad (4)$$

Behavior of Rosemann's model

Rosemann's model is indeed very flexible since it covers the two parameter Weibull model when *c*=1 and the three parameter models when *c* is very large (theoretically *c*=∞; in practice: *c* > 100 for example), but also all possible trends between these two extremes cases when 1< *c* <100, see Fig. 1 obtained using a Weibull plot, a scan of *F* from 1E-6 to 0.95 and L_0/L_{10}=0.2 with *c* = 1, 2, 3 and 10. The 3 parameter Weibull curve is also shown.

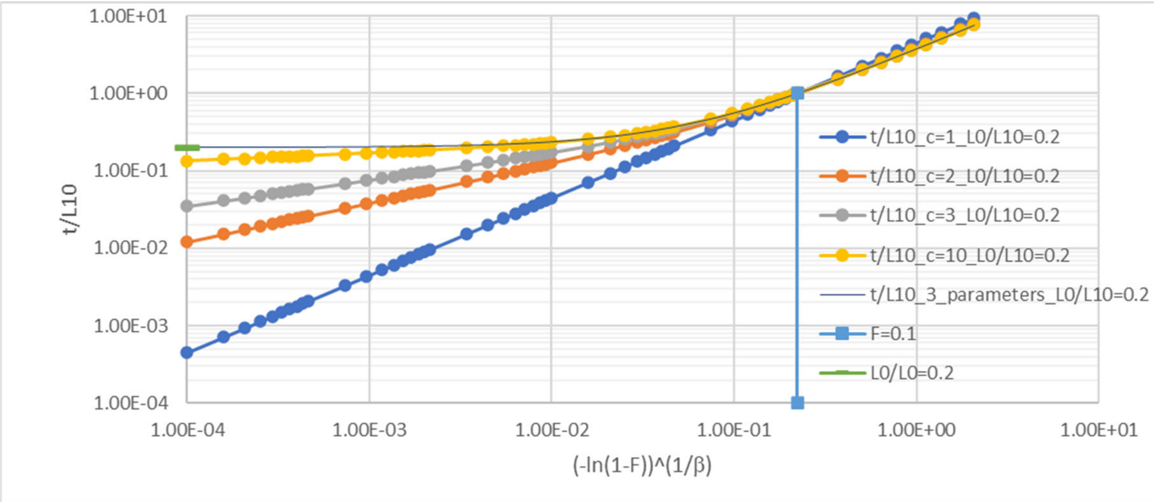


Fig. 1: Behavior of the four parameter Rosemann's model

The linear behavior is indeed observed when *c* = 1, while non-linear curves are observed when c >1, reaching asymptotically the 3 parameter Weibull curve when *c* is very large.

The ratio t/L_{10} is also called the reliability factor a_1 plotted next while reversing the *x* and *y* axis.

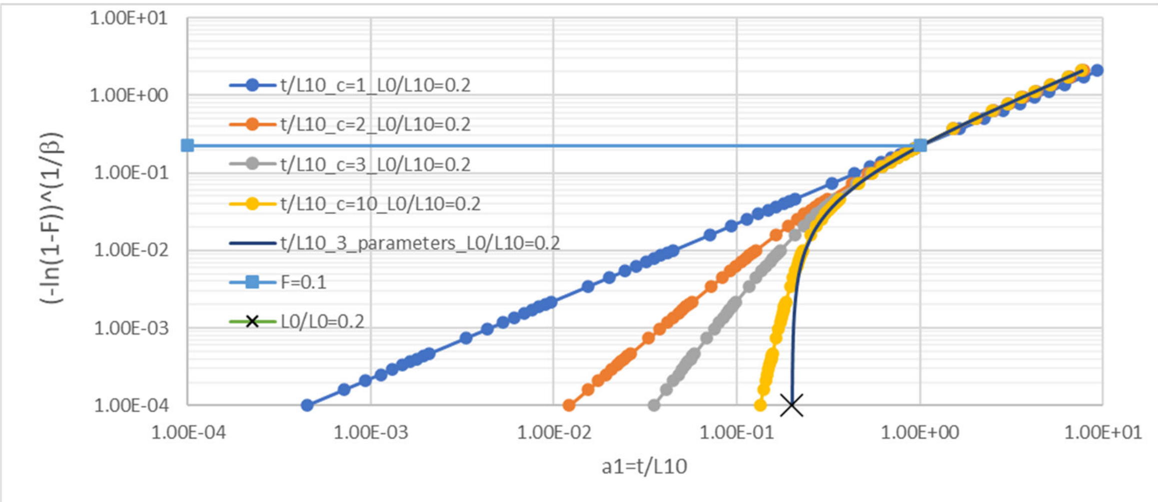


Fig. 2: Rosemann's reliability factor.

Study of F_i sorted in ascending order

Simulating randomly a set of *N* values of life *t* starts with the generation of *N* values of *F* (0 <F <1) to sort in an ascending order. One can calculate the density *f* and cumulative distribution *P* of each i^{th} number F_i.

For the sake of writing simplicity, it has been decided to attach next the index *i* (representing the i^{th} value) to the cumulative probability *P* (hence not on *F* as done previously).

When generating *N* numbers of *F* (0< F <1) and sorting them in an ascending order, one can calculate the density *f* and cumulative distribution P_i of each i^{th} number *F*. The density distribution *f(F)* corresponding to order i^{th} value of *F* is:

$$f(F)=\frac{N!}{(N-i)!.(i-1)!}.F^{i-1}.(1-F)^{N-i} \qquad (5)$$

The cumulative density P_i (probability that the i^{th} sorted random value is smaller or equal to F) is:

$$P_i=\frac{N!}{(N-i)!.(i-1)!}.\int_0^F x^{i-1}.(1-x)^{N-i}.dx=A_i*I_i$$
$$with \quad A_i=\frac{N!}{(N-i)!.(i-1)!} \quad , \quad I_i=\int_0^F x^{b_i}.(1-x)^{c_i}.dx \quad , \quad b_i=i-1 \quad \& \quad c_i=N-i \qquad (6)$$

Using analytical integration and integration by part approaches, a set of analytical polynomial relationships have been developed in appendix 1 for a few first and last values of I, see Eq. (39), (48), (51) and (52).

The next Figure shows the calculated values of P_i corresponding to the first 10 values when $N = 1000$:

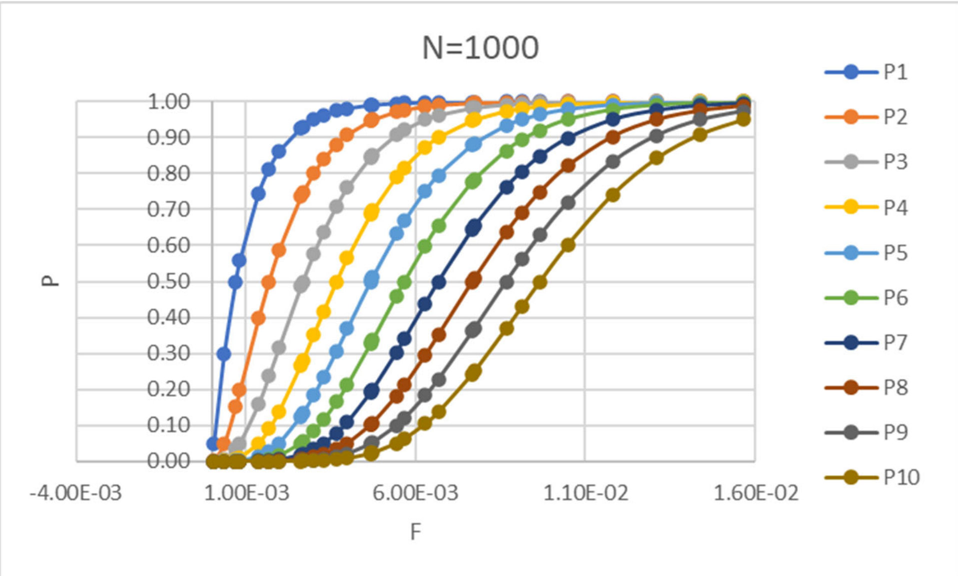


Fig. 3: First 10 values of P_i corresponding to $N = 1000$.

Let us now come back to the index i attached to the i^{th} value of F.

All previously defined analytical relationships can be used for solving numerically $F_i=F_{i\_P=PTargeted}$ corresponding to a targeted value of P, for example $P = 0.5$ when defining the median value of F_i, also called median rank, but also its lower and upper bounds using $P = 0.05$ or $P = 0.95$, hence defining the 90 % range of F_i.

There is however no need of conducting this tedious exercise because another exact approach is described next (known by reliability specialists but re-discovered by the authors with the help of Dr. Sicard [8]).

The exact approach is numerical and easy to program in Excel for example, requiring to simply use the *incomplete beta* and *inverse beta* functions.

The integral $\int_0^F x^a.(1-x)^b.dx$ corresponds to the *incomplete beta* function, itself calculated using the standard *beta* function and a ratio of *gamma* functions, such ratio being calculated numerically using the exponential of a sum of *gammalog* functions.

It should also be pointed out that the *gamma* function corresponds to a factorial number when using integers. As a consequence, the ratio of factorials (in front of the integral when calculating P) cancels out with the ratio of *gamma* functions, so that the final relationships (easy to program In Excel for example) read:

$$P = \frac{N!}{(N-i)!.(i-1)!}.\int_0^{F_i} x^{i-1}.(1-x)^{N-i}.dx \quad = \quad Beta\left(F_i, i, N-i+1\right) \quad (7)$$
$$F_i = InvBeta\left(P, i, N-i+1\right)$$

Any targeted value of P can now be used (P=0.05 or 0.5 or 0.95 for example) for calculating all N values of F_i with their medium lower and upper bounds, see next Figure.

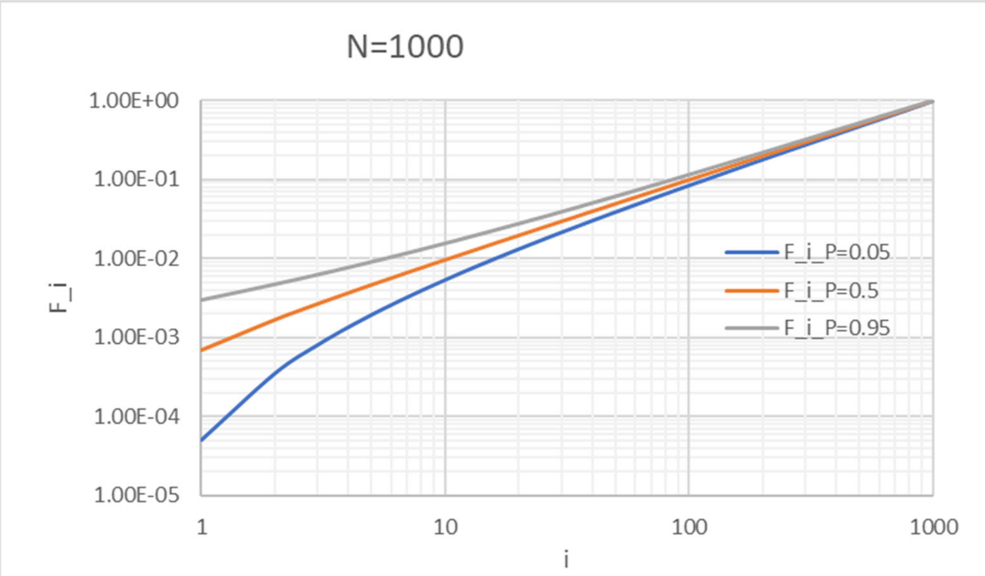


Fig. 4: Exact median, 5% lower and 95% upper bounds of F_i (for N=1000)

The median value can now be compared to some suggested approximations.

Miscellaneous simplified relationships for calculating the median value F_{median} have been provided in [2], the first one having been called in [2] ‘exact’ while it is understood today that the first one corresponded in reality to Johnson’s approximated relationship, [9] and [10]. A set of approximated relationships for F_{median} can now be tested:

$$Johnson1: F_{i\_P=0.5} \approx 1-2^{-\frac{1}{N}} + \frac{i-1}{N-1}.\left\{2^{\left(1-\frac{1}{N}\right)}-1\right\} \quad (used\ always\ while\ it\ should\ be\ when\ N \le N_{switch}\ with\ N_{switch}=20)$$

$$Johnson2: F_{i\_P=0.5} \approx \frac{i-0.30685-0.3863*\left(\frac{i-1}{N-1}\right)}{N} \quad when\ N > N_{switch} \qquad (8)$$

$$other\ approx: F_{i\_P=0.5} \approx \frac{i-0.305}{N+0.39}$$

$$Benard: F_{i\_P=0.5} \approx \frac{i-0.3}{N+0.4}$$

Johnson’s initial suggestion is to use *Johnson1* relationship when $N < N_{switch} = 20$, but a slightly improved accuracy has been observed (with N=1000) when not using N_{switch}. The accuracy is defined as *abs(error)* with *error* = $(F-F_{exact})/F_{exact}$. Results are shown in the next Table:

N--->	1000							
				max error--->	8.79E-03	8.93E-03	9.87E-03	1.28E-02
i	F_P=0.05	F_P=0.5	F_P=0.95		abs(error_Johnson1)	abs(error_Johnson2)	abs(error_other)	abs(error_Benard)
1	5.1291979E-05	6.9290701E-04	2.9912495E-03		4.6941E-16	3.5068E-04	2.6296E-03	9.8326E-03
2	3.5547613E-04	1.6777779E-03	4.7349936E-03		8.7872E-03	8.9317E-03	9.8710E-03	1.2840E-02
3	8.1817540E-04	2.6731593E-03	6.2822845E-03		7.0985E-03	7.1890E-03	7.7773E-03	9.6370E-03
4	1.3674363E-03	3.6708271E-03	7.7352447E-03		5.6993E-03	5.7651E-03	6.1927E-03	7.5442E-03
5	1.9721531E-03	4.6693423E-03	9.1299526E-03		4.7158E-03	4.7675E-03	5.1029E-03	6.1633E-03
6	2.6161408E-03	5.6682616E-03	1.0484077E-02		4.0073E-03	4.0497E-03	4.3255E-03	5.1972E-03
7	3.2897874E-03	6.6674044E-03	1.1807823E-02		3.4774E-03	3.5134E-03	3.7474E-03	4.4870E-03
8	3.9868512E-03	7.6666837E-03	1.3107715E-02		3.0678E-03	3.0991E-03	3.3021E-03	3.9440E-03
9	4.7030132E-03	8.6660523E-03	1.4388225E-02		2.7423E-03	2.7699E-03	2.9492E-03	3.5159E-03
10	5.4351401E-03	9.6654825E-03	1.5652574E-02		2.4778E-03	2.5024E-03	2.6629E-03	3.1699E-03
11	6.1808741E-03	1.0664957E-02	1.6903175E-02		2.2586E-03	2.2809E-03	2.4260E-03	2.8846E-03
12	6.9383880E-03	1.1664465E-02	1.8141892E-02		2.0742E-03	2.0946E-03	2.2269E-03	2.6454E-03
13	7.7062327E-03	1.2663997E-02	1.9370199E-02		1.9169E-03	1.9356E-03	2.0573E-03	2.4419E-03
14	8.4832353E-03	1.3663550E-02	2.0589287E-02		1.7811E-03	1.7985E-03	1.9110E-03	2.2668E-03
15	9.2684310E-03	1.4663118E-02	2.1800132E-02		1.6629E-03	1.6790E-03	1.7836E-03	2.1145E-03
16	1.0061015E-02	1.5662698E-02	2.3003549E-02		1.5589E-03	1.5739E-03	1.6717E-03	1.9808E-03
17	1.0860306E-02	1.6662289E-02	2.4200226E-02		1.4667E-03	1.4808E-03	1.5726E-03	1.8625E-03
18	1.1665724E-02	1.7661889E-02	2.5390749E-02		1.3845E-03	1.3978E-03	1.4842E-03	1.7571E-03
19	1.2476769E-02	1.8661495E-02	2.6575623E-02		1.3107E-03	1.3233E-03	1.4048E-03	1.6627E-03
20	1.3293006E-02	1.9661108E-02	2.7755286E-02		1.2441E-03	1.2560E-03	1.3333E-03	1.5775E-03
25	1.7440156E-02	2.4659238E-02	3.3587428E-02		9.8951E-04	9.9889E-04	1.0598E-03	1.2525E-03
30	2.1674784E-02	2.9657436E-02	3.9331876E-02		8.1838E-04	8.2609E-04	8.7625E-04	1.0348E-03
35	2.5976079E-02	3.4655674E-02	4.5009536E-02		6.9548E-04	7.0202E-04	7.4448E-04	8.7870E-04
40	3.0330486E-02	3.9653936E-02	5.0634007E-02		6.0296E-04	6.0861E-04	6.4532E-04	7.6136E-04
50	3.9163536E-02	4.9650503E-02	6.1758579E-02		4.7292E-04	4.7733E-04	5.0602E-04	5.9668E-04
60	4.8124364E-02	5.9647103E-02	7.2755267E-02		3.8590E-04	3.8950E-04	4.1284E-04	4.8663E-04
70	5.7183468E-02	6.9643723E-02	8.3653620E-02		3.2359E-04	3.2660E-04	3.4614E-04	4.0791E-04
80	6.6321578E-02	7.9640355E-02	9.4472929E-02		2.7677E-04	2.7934E-04	2.9603E-04	3.4879E-04
90	7.5525276E-02	8.9636994E-02	1.0522663E-01		2.4031E-04	2.4254E-04	2.5702E-04	3.0278E-04
100	8.4784768E-02	9.9633640E-02	1.1592451E-01		2.1111E-04	2.1306E-04	2.2577E-04	2.6594E-04
200	1.7936803E-01	1.9960021E-01	2.2091464E-01		7.9357E-05	8.0090E-05	8.4849E-05	9.9892E-05
300	2.7612277E-01	2.9956685E-01	3.2373309E-01		3.5319E-05	3.5645E-05	3.7760E-05	4.4448E-05
400	3.7426088E-01	3.9953351E-01	4.2516812E-01		1.3280E-05	1.3402E-05	1.4197E-05	1.6711E-05
500	4.7351773E-01	4.9950017E-01	5.2548440E-01		5.2853E-08	5.3340E-08	5.6505E-08	6.6507E-08
500.5	4.7401666E-01	5.0000000E-01	5.2598334E-01		1.1102E-15	8.8818E-16	8.8818E-16	8.8818E-16
501	4.7451560E-01	5.0049983E-01	5.2648227E-01		5.2748E-08	5.3234E-08	5.6392E-08	6.6374E-08
600	5.7382341E-01	5.9946683E-01	6.2475185E-01		8.7628E-06	8.8435E-06	9.3682E-06	1.1027E-05
601	5.7483188E-01	6.0046649E-01	6.2573912E-01		8.8361E-06	8.9175E-06	9.4466E-06	1.1119E-05
700	6.7524639E-01	6.9943348E-01	7.2290202E-01		1.5052E-05	1.5191E-05	1.6092E-05	1.8942E-05
701	6.7626691E-01	7.0043315E-01	7.2387723E-01		1.5106E-05	1.5245E-05	1.6150E-05	1.9010E-05
800	7.7804857E-01	7.9940012E-01	8.1967303E-01		1.9749E-05	1.9931E-05	2.1116E-05	2.4859E-05
801	7.7908536E-01	8.0039979E-01	8.2063197E-01		1.9790E-05	1.9972E-05	2.1159E-05	2.4911E-05
900	8.8300847E-01	8.9936670E-01	9.1428652E-01		2.3330E-05	2.3546E-05	2.4951E-05	2.9389E-05
901	8.8407549E-01	9.0036636E-01	9.1521523E-01		2.3361E-05	2.3577E-05	2.4984E-05	2.9428E-05
911	8.9477337E-01	9.1036301E-01	9.2447472E-01		2.3662E-05	2.3881E-05	2.5307E-05	2.9812E-05
921	9.0552707E-01	9.2035965E-01	9.3367842E-01		2.3950E-05	2.4172E-05	2.5616E-05	3.0182E-05
925	9.0984627E-01	9.2435830E-01	9.3734219E-01		2.4061E-05	2.4284E-05	2.5736E-05	3.0325E-05
931	9.1634638E-01	9.3035628E-01	9.4281653E-01		2.4223E-05	2.4448E-05	2.5911E-05	3.0535E-05
941	9.2724473E-01	9.4035290E-01	9.5187564E-01		2.4478E-05	2.4706E-05	2.6187E-05	3.0867E-05
950	9.3713660E-01	9.4934984E-01	9.5994553E-01		2.4686E-05	2.4916E-05	2.6413E-05	3.1144E-05
951	9.3824142E-01	9.5034950E-01	9.6083646E-01		2.4707E-05	2.4938E-05	2.6437E-05	3.1173E-05
961	9.4936599E-01	9.6034606E-01	9.6966951E-01		2.4897E-05	2.5130E-05	2.6646E-05	3.1437E-05
966	9.5499046E-01	9.6534433E-01	9.7402392E-01		2.4968E-05	2.5202E-05	2.6727E-05	3.1545E-05
971	9.6066812E-01	9.7034256E-01	9.7832522E-01		2.5013E-05	2.5249E-05	2.6781E-05	3.1626E-05
975	9.6525748E-01	9.7434113E-01	9.8171911E-01		2.5021E-05	2.5258E-05	2.6797E-05	3.1663E-05
976	9.6641257E-01	9.7534076E-01	9.8255984E-01		2.5017E-05	2.5255E-05	2.6796E-05	3.1667E-05
981	9.7224471E-01	9.8033889E-01	9.8670699E-01		2.4952E-05	2.5190E-05	2.6740E-05	3.1637E-05
982	9.7342438E-01	9.8133850E-01	9.8752323E-01		2.4925E-05	2.5164E-05	2.6715E-05	3.1618E-05
983	9.7460925E-01	9.8233811E-01	9.8833428E-01		2.4893E-05	2.5132E-05	2.6684E-05	3.1592E-05
984	9.7579977E-01	9.8333771E-01	9.8913969E-01		2.4853E-05	2.5092E-05	2.6646E-05	3.1559E-05
985	9.7699645E-01	9.8433730E-01	9.8993899E-01		2.4804E-05	2.5044E-05	2.6600E-05	3.1518E-05
986	9.7819987E-01	9.8533688E-01	9.9073157E-01		2.4745E-05	2.4985E-05	2.6543E-05	3.1466E-05
987	9.7941071E-01	9.8633645E-01	9.9151676E-01		2.4674E-05	2.4914E-05	2.6473E-05	3.1402E-05
988	9.8062980E-01	9.8733600E-01	9.9229377E-01		2.4587E-05	2.4827E-05	2.6388E-05	3.1321E-05
989	9.8185811E-01	9.8833554E-01	9.9306161E-01		2.4480E-05	2.4720E-05	2.6283E-05	3.1221E-05
990	9.8309682E-01	9.8933504E-01	9.9381913E-01		2.4348E-05	2.4588E-05	2.6152E-05	3.1096E-05
991	9.8434743E-01	9.9033452E-01	9.9456486E-01		2.4182E-05	2.4423E-05	2.5989E-05	3.0938E-05
992	9.8561178E-01	9.9133395E-01	9.9529699E-01		2.3973E-05	2.4214E-05	2.5781E-05	3.0735E-05
993	9.8689228E-01	9.9233332E-01	9.9601315E-01		2.3702E-05	2.3943E-05	2.5512E-05	3.0471E-05
994	9.8819218E-01	9.9333260E-01	9.9671021E-01		2.3341E-05	2.3583E-05	2.5153E-05	3.0117E-05
995	9.8951592E-01	9.9433174E-01	9.9738386E-01		2.2844E-05	2.3086E-05	2.4658E-05	2.9627E-05
996	9.9087005E-01	9.9533066E-01	9.9802785E-01		2.2123E-05	2.2365E-05	2.3939E-05	2.8913E-05
997	9.9226476E-01	9.9632917E-01	9.9863256E-01		2.0998E-05	2.1241E-05	2.2816E-05	2.7796E-05
998	9.9371772E-01	9.9732684E-01	9.9918182E-01		1.9026E-05	1.9269E-05	2.0846E-05	2.5830E-05
999	9.9526501E-01	9.9832222E-01	9.9964452E-01		1.4768E-05	1.5011E-05	1.6589E-05	2.1579E-05
1000	9.9700875E-01	9.9930709E-01	9.9994871E-01		6.6660E-16	2.4316E-07	1.8233E-06	6.8178E-06

Table 1: Exact median, 5% lower and 95% upper bounds of F_i (for N=1000), median values compared to suggested approximations (for N=1000)

Johnson1 approximation is quite accurate with a maximum error equal to 0.00879 when i = 2 and N=1000.

While testing the case N=10, the approximation called *other approximation* was found slightly more accurate, see next Table.

N--->	10							
				max error--->	5.76E-03	3.51E-02	5.39E-03	7.39E-03
i	F_P=0.05	F_P=0.5	F_P=0.95		abs(error_Johnson1)	abs(error_Johnson2)	abs(error_other)	abs(error_Benard)
1	5.1161969E-03	6.6967008E-02	2.5886555E-01		2.0723E-16	3.5062E-02	1.1314E-03	5.0873E-03
2	3.6771438E-02	1.6226273E-01	3.9416330E-01		5.7551E-03	1.7010E-02	5.3919E-03	7.3881E-03
3	8.7264434E-02	2.5857472E-01	5.0690130E-01		3.2926E-03	8.3374E-03	3.1298E-03	4.0246E-03
4	1.5002824E-01	3.5509997E-01	6.0662422E-01		1.5649E-03	3.7690E-03	1.4938E-03	1.8847E-03
5	2.2244110E-01	4.5169416E-01	6.9646279E-01		4.2300E-04	1.0006E-03	4.0436E-04	5.0680E-04
6	3.0353721E-01	5.4830584E-01	7.7755890E-01		3.4847E-04	8.2428E-04	3.3311E-04	4.1751E-04
7	3.9337578E-01	6.4490003E-01	8.4997176E-01		8.6169E-04	2.0753E-03	8.2252E-04	1.0378E-03
8	4.9309870E-01	7.4142528E-01	9.1273557E-01		1.1483E-03	2.9077E-03	1.0915E-03	1.4036E-03
9	6.0583670E-01	8.3773727E-01	9.6322856E-01		1.1147E-03	3.2946E-03	1.0444E-03	1.4310E-03
10	7.4113445E-01	9.3303299E-01	9.9488380E-01		0.0000E+00	2.5165E-03	8.1205E-05	3.6514E-04

Table 2: Exact median, 5% lower and 95% upper bounds of F_i (for N=1000), median values compared to suggested approximations (for N=10)

Generating random databases (simulating experimental databases) of t_{exp_i}

Having defined randomly N values of i, sorted F_i in an ascending order and understood its median value, 5% lower and 95% upper bounds, it is now possible to calculate t_{exp_i} using any set of (η or L_{10}, β, L_0 and c) and F_i. The life t_{exp_i} can then plotted versus $-ln(1-F_{median})$ in a standard Weibull plot (using $ln(t_{exp\_i})$ and $ln(-ln(1-F_{median}))$ and compared to t_i calculated with the median, lower and upper bounds of F_i, see next example and Figure obtained with 3 examples of randomly simulated experimental values of t_{exp} defined with: L_{10} =1, β = 1, L_0=0.2 , c=2 and N=1000. Zooms showing the results obtained at low and larger F_{median} values are also given.

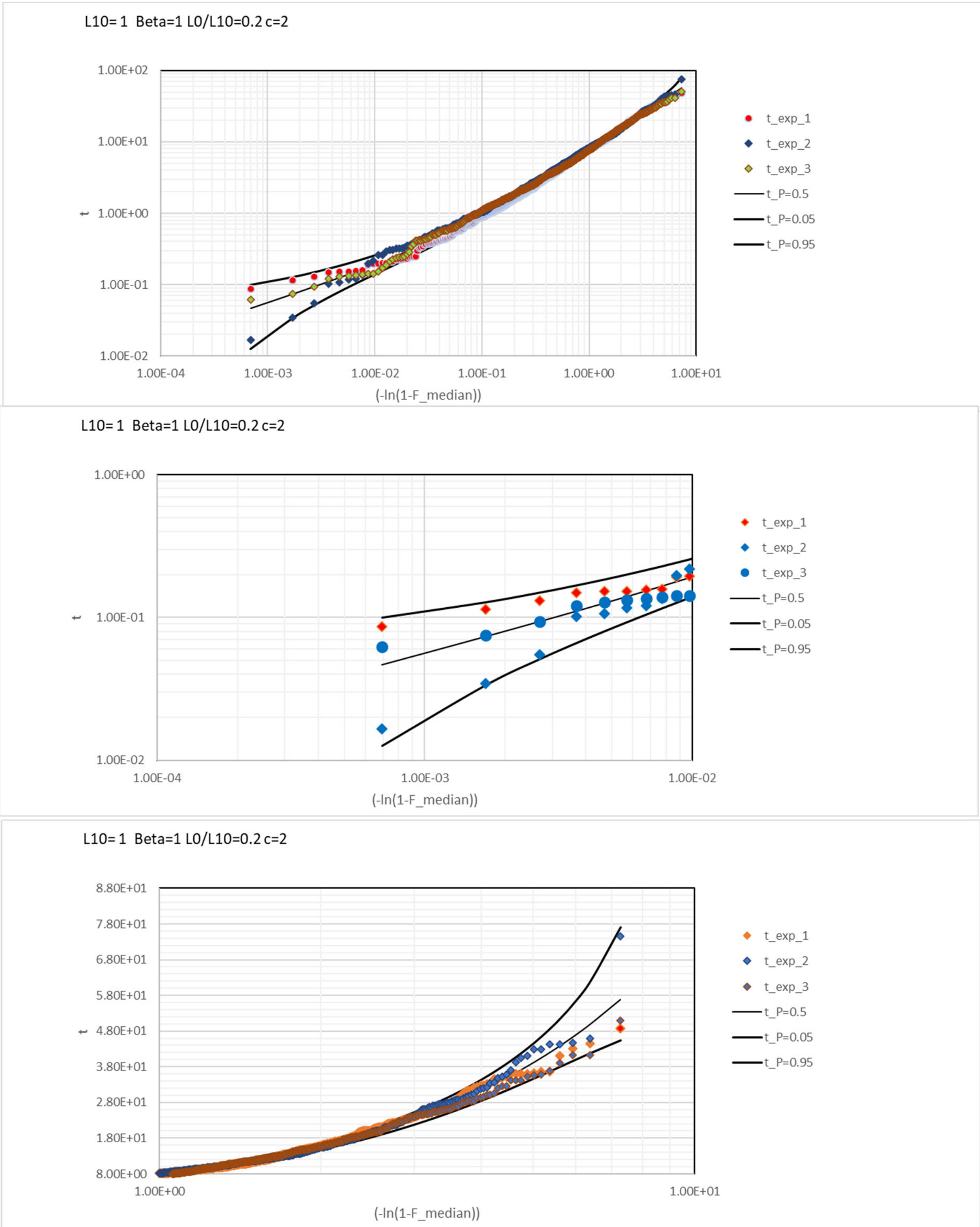


Fig. 5: Example of one random simulation of an experimental database (for c=2 & N=1000)

At low failure rate, the 90 % range, hence scatter of experimental points, can be quite large when c=2. The 90% range and scatter decrease substantially (at low F values only) as c increases, see next example obtained with c=10.

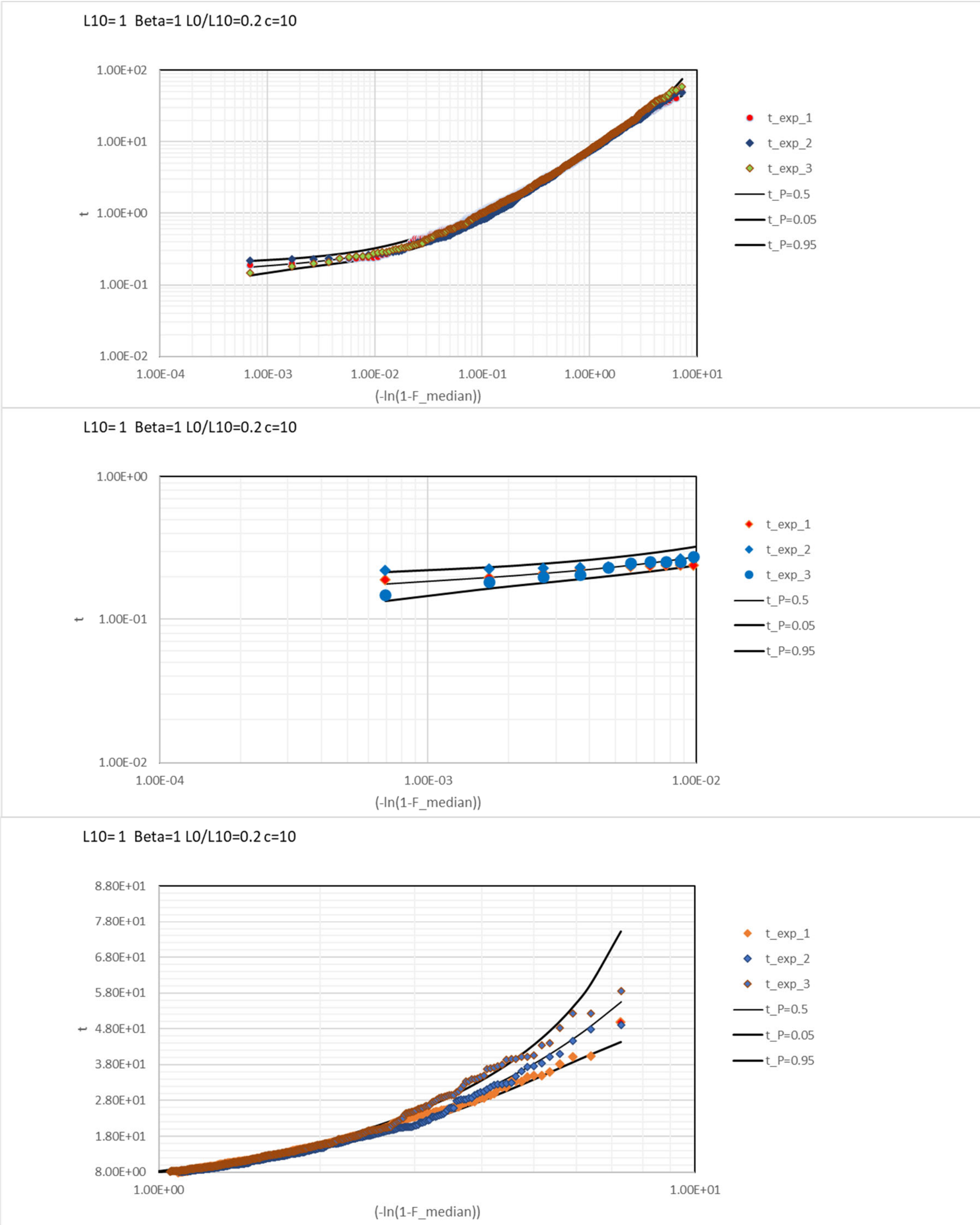


Fig. 6: Example of one random simulation of an experimental database (for c=10 & N=1000)

It can therefore already be anticipated that when trying to curve-fit an experimental database for extracting the 4 Rosemann parameters, it is very likely that the accuracy on L_0 and c might be poor when c is small since the 90 % range of t_{exp} is large.

Before trying to develop curve-fitting technics for extracting the 4 Rosemann parameters and their confidence intervals (using Monte Carlo simulations), one can already study the effect of N and c on the 90 % range, hence likely t_{exp} scatter.

The *inverse beta* function can be used for easily calculating the life and life range corresponding to F<0.1 using N = 10000, then 1000 and 100 and several c values.

The range is calculated using any boundary values: *P_lower* = 0.05 and *P_upper* = 0.95 for example. Also shown next (on the second *y* axis) is the ratio *R=t_P_upper/t_P_lower*.

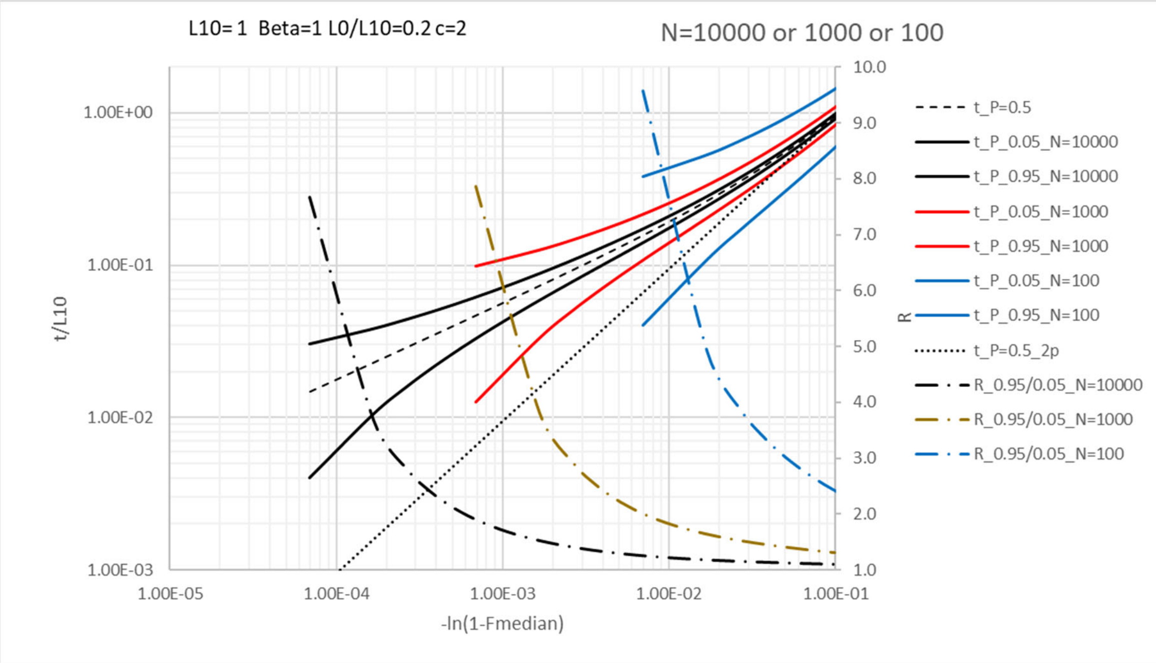


Fig. 7: Calculated 90% t_{exp} range for c = 2.

At low F_{median} value (or low $-ln(1-F_{median})$), the ratio R can be quite large and illustrates the most likely difficulty of extracting accurate values of L_0 and c

For example: R = 7.86 when F_{median}=6.93E-4

But this ratio drops to about 1.9 at F_{median}=6.93E-4 when N = 10000.

As anticipated this ratio R drops significantly as the exponent c increases, see next examples obtained with c = 10 and 300.

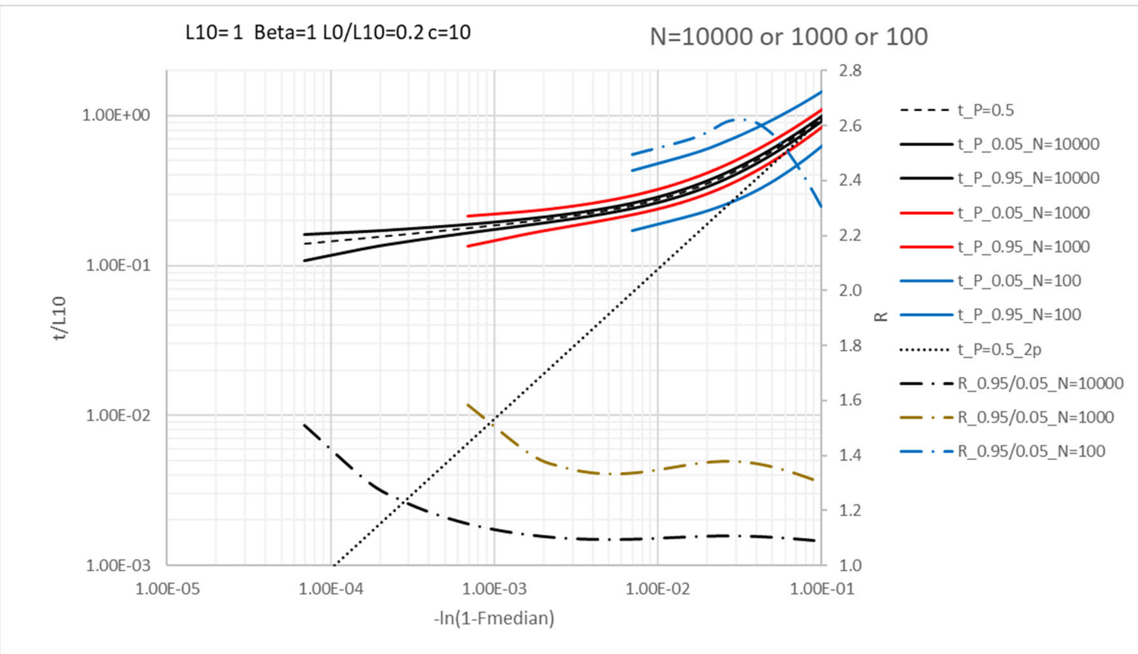


Fig. 8: Calculated 90% t_{exp} range for c = 10.

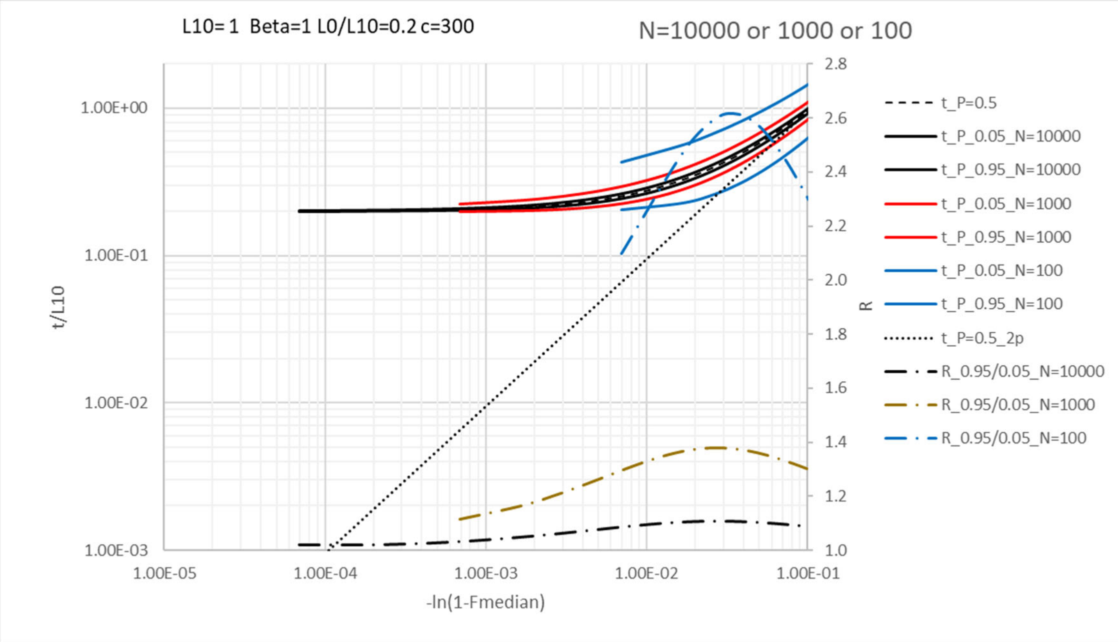


Fig. 9: Calculated 90% t_{exp} range for c = 300.

The latter example (c = 300) simulates almost a 3 parameter Weibull model leading to an accurate estimate of L_0 when N = 1000, but not very accurate when N = 100, the ratio R remaining large and of the order of 2.2

Curve-fitting technics of an experimental database

Rosemann's model can also be written:

$$t = \left[\left\{ \eta.\left(-\ln(1-F)\right)^{\frac{1}{\beta}} + L_0 \right\}^{c} - L_0^{c} \right]^{\frac{1}{c}} \qquad (9)$$

η, β, L_0 and c being the four unknowns to define by curve-fitting.

Note that one will also use later the following relationships:

$$\ln\left[\eta.\left(-\ln(1-F)\right)^{\frac{1}{\beta}}\right] = \ln(\eta) + \frac{1}{\beta}.\ln\left(-\ln\left(1-F\right)\right) \ \ hence: \qquad (10)$$

$$\eta.\left(-\ln(1-F)\right)^{\frac{1}{\beta}} = \exp\left[\frac{1}{\beta}.\ln\left(-\ln\left(1-F\right)\right) + \ln(\eta)\right]$$

<u>ML approach:</u>

One possible approach consists of using the maximum likelihood approach (ML) developed in appendix 2 but not tested herein.

The ML approach consists of maximizing the product of the density function $f(t)$:

$$f(t) = \frac{dF}{dt} = \frac{\beta}{\eta}.\exp\left[-\left(\frac{\left(t^{c}+L_0^{c}\right)^{\frac{1}{c}} - L_0}{\eta}\right)^{\beta}\right].\left(\frac{\left(t^{c}+L_0^{c}\right)^{\frac{1}{c}} - L_0}{\eta}\right)^{\beta-1}.\left(t^{c}+L_0^{c}\right)^{\frac{1}{c}-1}.t^{c-1} \qquad (11)$$

$$\text{Product} = \prod_{i=1}^{N} f(t_i) = Max$$

The other standard approach consists of sorting the set of experimental life t_{exp_i} in ascending order, and to use the median rank F_{median_i} for estimating the corresponding cumulative failure probability. A non-linear curve-fitting between t_{exp_i} and F_{median_i} (or F_{median_i} versus t_{exp_i}) must then be conducted for obtaining t_{cf_i} to compare to t_{exp_i}.

It is recommended to use the log function, hence *ln(t),* for putting the same weight to small and large values of t_{exp_i} and ratio $\frac{t_{cf\_i}}{t_{\exp\_i}}$ since $\ln\left(\frac{t_{cf\_i}}{t_{\exp\_i}}\right) = \ln\left(t_{cf\_i}\right) - \ln\left(t_{\exp\_i}\right)$.

$$\ln\left(t_{cf\_i}\right) = \frac{1}{c}.\ln\left[\left\{\eta.\left(-\ln(1-F_{median\_i})\right)^{\frac{1}{\beta}} + L_0\right\}^c - L_0^c\right] \quad (12)$$

Method 1:

One therefore needs to conduct a non-linear curve-fitting of $Y = ln(t)$ versus $X=ln(-ln(1-F_{median}))$, minimizing for example the sum of the vertical distance between Y_{cf_i} and Y_{exp_i}, leading to the so-called (herein) Method 1 also studied in detail by Houpert in [2] with a 2 parameter Weibull model:

$$S^2 = \sum_{i=1,N} S_i^2 = \sum_{i=1,N} \left(Y_{cf\_i} - Y_{\exp\_i}\right)^2 = \min \quad (Method\, 1) \quad (13)$$

$$\begin{aligned}
&\ln\left(t_{cf_i}\right) = \ln\left\{\left[\left\{\exp\left(a.X_i + b\right) + L_0\right\}^c - L_0^c\right]^{\frac{1}{c}}\right\} \\
&with: \\
&X_i = \ln\left(-\ln\left(1-F_{median_i}\right)\right) \;\; and\;\; 4\;\; unknowns: \\
&a = \frac{1}{\beta} \quad b = \ln(\eta) \quad L_0 \;\; \& \;\; c
\end{aligned} \quad (14)$$

Hence:

$$Y_{cf\_i} = \ln\left(t_{cf\_i}\right) = \ln\left\{\left[\left\{\exp\left(a.X_i + b\right) + L_0\right\}^c - L_0^c\right]^{\frac{1}{c}}\right\} \;\; to \;\; compare \;\; to \;\; Y_{\exp\_i} = \ln\left(t_{\exp\_i}\right) \quad (15)$$

Details of Method 1 are given in appendix 3.

Method 2:

A second approach called herein Method 2, consists of curve-fitting X versus t_{exp} and to minimize the horizontal distance between X_{cf_i} and X_i defined as:

$$\begin{aligned}
&X_{cf_i} = a.\ln\left\{\left(t_{\exp_i}^c + L_0^c\right)^{\frac{1}{c}} - L_0\right\} + b \qquad X_i = \ln\left(-\ln\left(1-F_{median_i}\right)\right) \\
&with\; 4\; unknowns: \\
&a = \beta \quad b = -\beta.\ln(\eta) \quad L_0 \quad and \quad c
\end{aligned} \quad (16)$$

$$S^2 = \sum_{i=1,N} \left(X_{cf\_i} - X_i\right)^2 = \min \quad (Method\, 2) \quad (17)$$

Details about Method 2 are given in appendix 4.

Prior of showing the results obtained using a few examples, the robustness of the two approaches (Method 1 and 2) has been tested and confirmed, replacing the experiment values of t_{exp} by the exact values of t and confirming that the calculated set of unknowns (a,b,L_0 and c), initially estimated, does converge towards the exact set used for defining the exact values of t.

The curve-fitted values of a, b, L_0 and c will be compared next to the ones used as inputs for simulating our random set of t_{exp_i}.

New curve-fitting suggestion:

Also, following some results shown next, an alternative curve-fitting technique cited as "New" will be suggested at the end of this paper and fully tested in [14].

Results obtained

One can now simulate experimental cases via a set of random values of t_{exp} obtained using a random value of F (instead of the median value of F_{median}) with L_{10} =1, β = 1, L_0 = 0.2 while the exponent c will range from 2 to 100.

Beside some problems described next and observed when c_{cf} = 1, some numerical problems can be found in some seldom cases (especially when conducting 10000 or 100000 Monte-Carlo simulations) using method 1 or 2:

- The sum S^2 can decrease nicely during the first iterations and then start to increase.
- The suggested solution or convergence may also depend on the initial guess of 4 unknowns and accepted tolerance.

Following are a few examples of results obtained.

For avoiding showing dense Figures, the 90 % range of t_{exp} for the first 3 and last 3 points only is shown.

Also shown are the curve-fitted results obtained using the 2-parameter model and F_{median} > 0.05

Si2 represents the calculated value of $\sum_{i=1}^{N} S_i^2$.

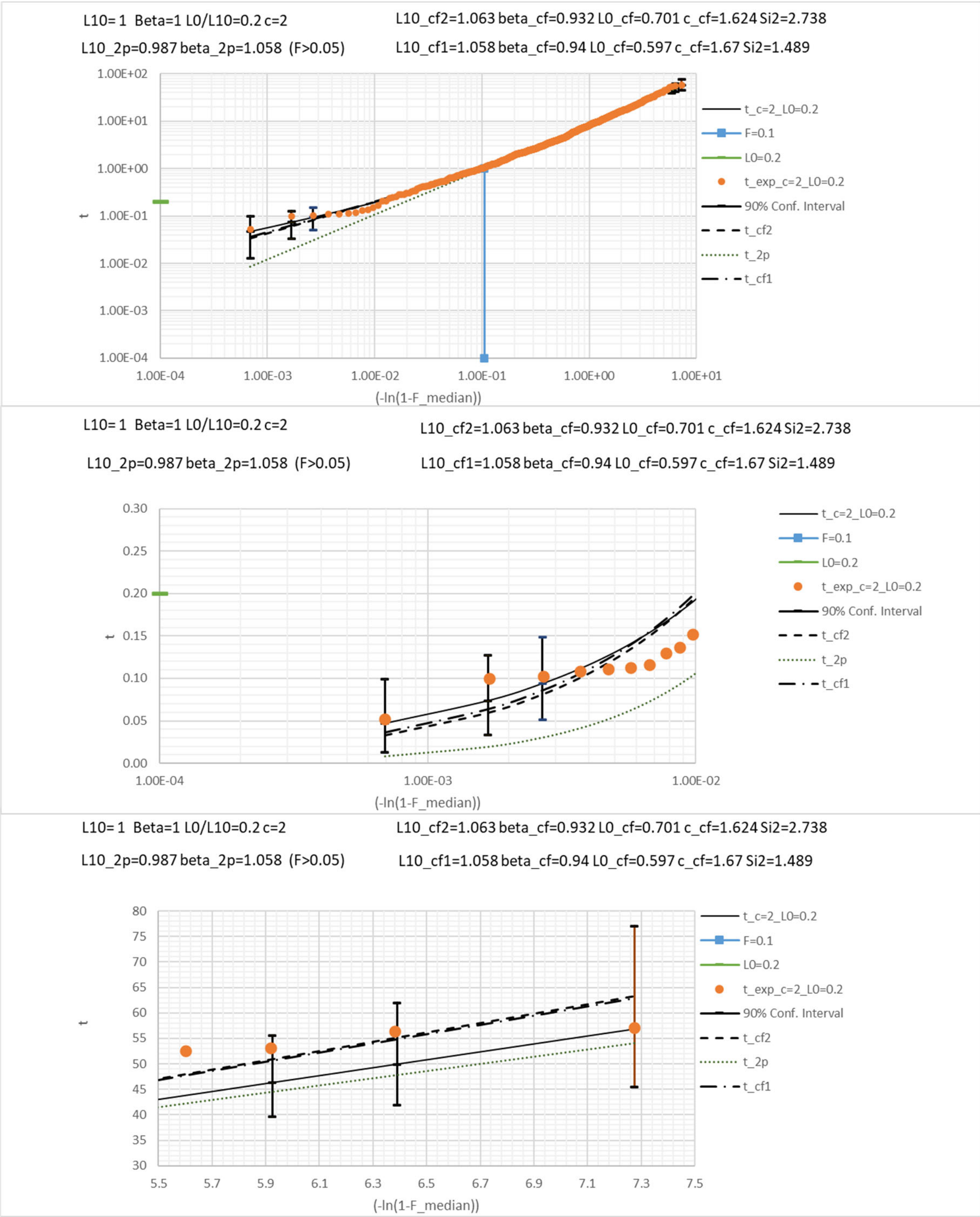


Fig. 10: Example of curve-fitted results obtained using c =2

Results obtained using Method 1 and 2 differ slightly.

In this first example, larger values of L_0 are found (with Method 2 for example: 0.701 instead of 0.2) compensated by smaller values of c (with Method 2 for example: 1.624 instead of 2). The curve-fitted curves do however pass successfully through the experimental points at low F_{median} values.

Let's show next some additional simulations using c = 2:

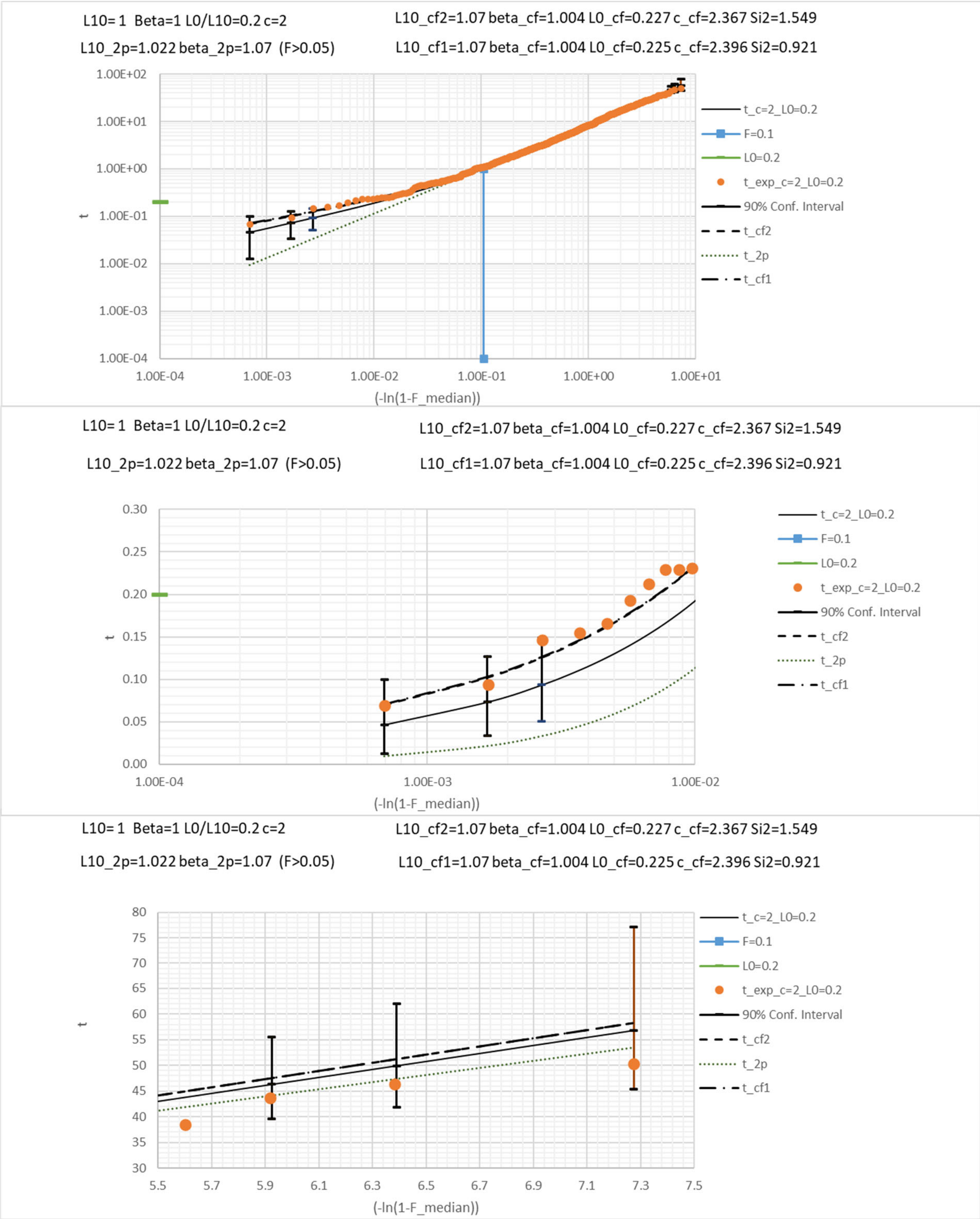


Fig. 11: Second example of curve-fitted results obtained using c =2

Here, the curve-fitted values of L_o and c are quite satisfactory.

When duplicating such an exercise 10,000 times, confidential intervals will be defined next. One can anticipate large confidence intervals when c = 2. Note also that defining confidence intervals applicable to each single unknown L_0 and c is certainly not appropriate since the accuracy of the final result is defined by the set (L_0, c).

A specific study conducted later will show that the same trend (concerning the first points) can be explained using either a low L_0 value compensated by a large c value or the opposite.

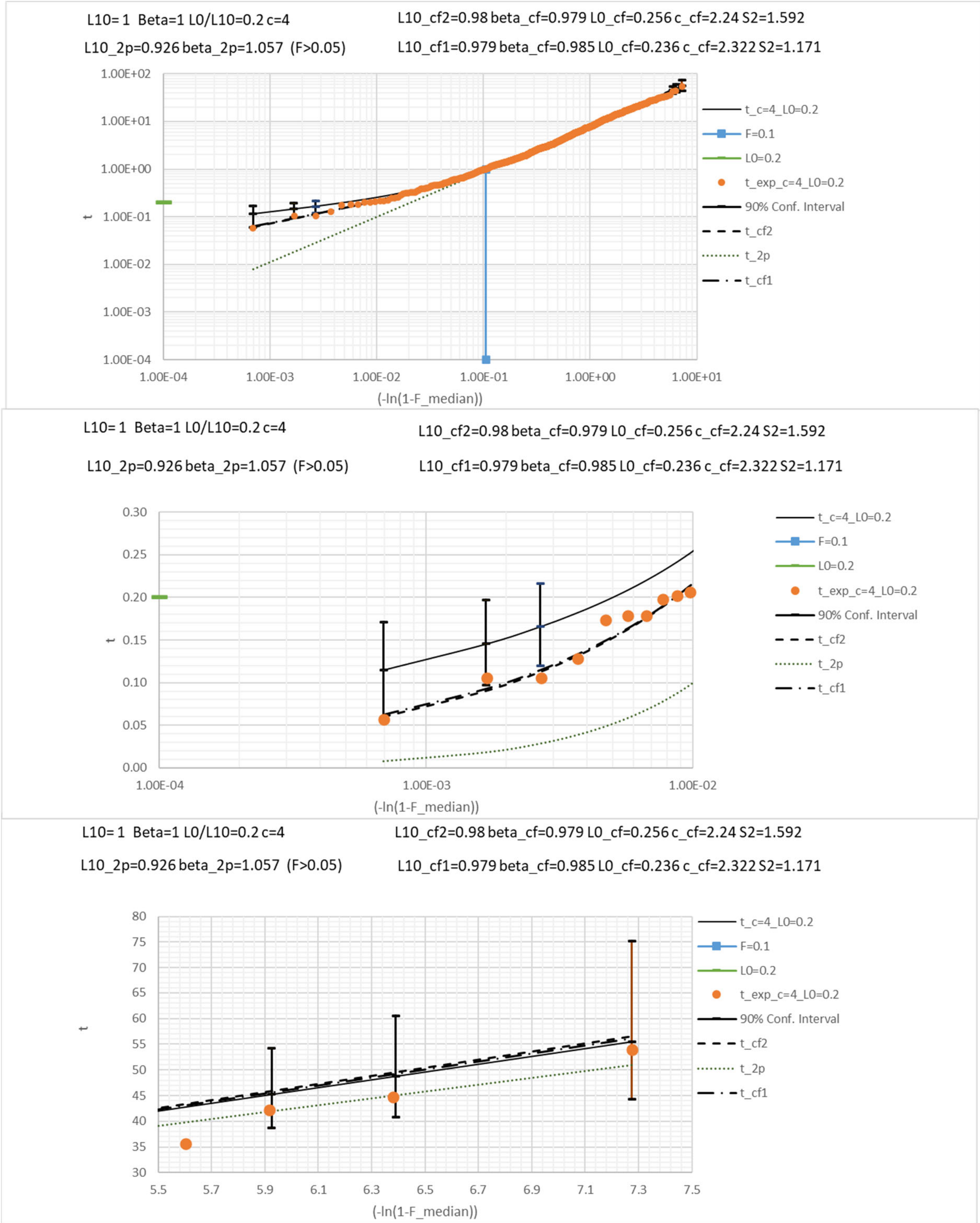


Fig. 12: Example of curve-fitted results obtained using c =4

Again, the individual values of L_0 and c are difficult to retrieve, but the final curve-fitted curves do match the experimental results. A smaller value of c can compensate a large value of L_0 when the random points are below the exact curve. The opposite applies when the random cases are above the exact curve.

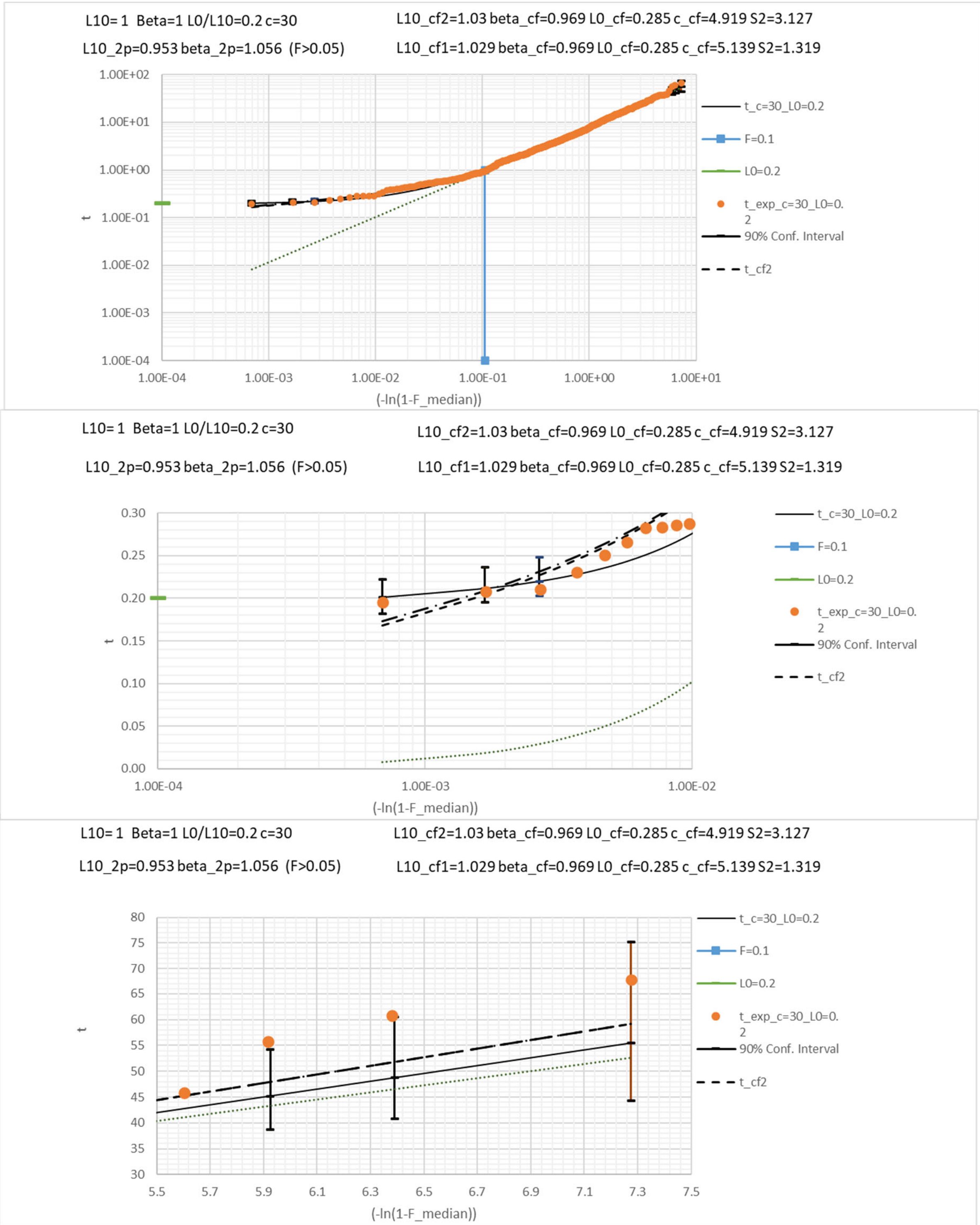


Fig. 13: Example of curve-fitted results obtained using c =30

Again, the estimate of L_0 and c is poor, but the final match at low F values is satisfactory.

As mentioned before, an alternative curve-fitting technique, simply called “New” will be described later for solving the latter problem.

Preliminary conclusions:

Estimating L_0 and c when c is small, of the order of 2 for example, is challenging since miscellaneous set of (L_0 and c) can fit at set of experiments results within the 90 % range of t_{exp} at low F values.

As a demonstration of the latter claim, a specific study has also been conducted next with L_{10}=1 and β= 1.

In the following, Method 2 is used for defining the exact value of F obtained when scanning on small values of t with miscellaneous value of L_0 and c. The slope β is fixed to 1 and the constant b is defined for retrieving F = 0.1 when t = 1. Reference case corresponds to L_0=0.2 and c=2.

$$\ln\left[-\ln(1-F)\right] = \beta.\ln\left[\left(t^c + L_0^c\right)^{\frac{1}{c}} - L_0\right] + b$$

$$b = \ln\left[-\ln(0.9)\right] - \beta.\ln\left[\left(1 + L_0^c\right)^{\frac{1}{c}} - L_0\right] \qquad \left(for\ \text{info}: b = -\beta.\ln(\eta)\right) \tag{18}$$

$$\ln\left[-\ln(1-F)\right] = \beta.\ln\left[\frac{\left(t^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\left(1 + L_0^c\right)^{\frac{1}{c}} - L_0}\right] + \ln\left[-\ln(0.9)\right] \tag{19}$$

One sees next that similar trends can be obtained using either very low values of L_0 compensated by very large values of c (*L_0=0.05 & c=10* for example), or the opposite (*L_0=1 & c=1.6* for example).

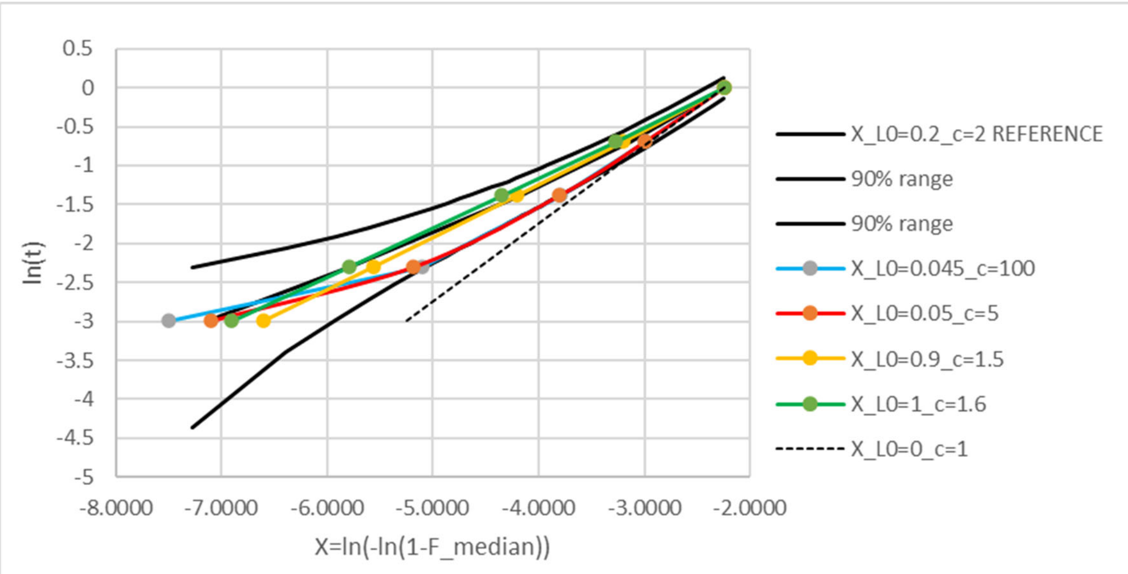


Fig. 14: Example of miscellaneous set of (L_o & c) values compatible the 90% range

As anticipated, the previously calculated life t using miscellaneous wrong but possible sets of (L_o,c) fits the 90% possible range, confirming the difficulty of correctly defining any single value of L_0 and c (when c=2 and N=1000) while the final curve-fitting can be acceptable.

The latter statement will be confirmed next by conducting Monte-Carlo simulations, duplicating for example 10,000 times a curve-fitting exercise for retrieving 10,000 time the ratios $\frac{L_{10}}{L_{10\_cf}}, \frac{\beta}{\beta_{cf}}, \frac{L_0}{L_{0\_cf}}$ & $\frac{c}{c_{cf}}$ to sort in ascending order and defining their median values and confidence intervals.

Also, because of the problems encountered for defining L_0 and c, an alternative "New" curve-fitting technique and model will be suggested at the end of this paper.

Monte-Carlo simulations ; confidence intervals

Monte-Carlo simulations have been used for conducting *NS* times (*NS*=10,000 in the following results) the curve-fitting of *N* (*N*=1000 in the following example) randomly generated values of *t_exp* (generated using a given set of (η, β, L_0 and c) or (L_{10}, β, L_0 and c) inputs, for example:

L_{10}=1 β=*1*, L_0 =0.2 and c=*2* in the following example.

The ratio $\left(\frac{L_{10}}{L_{10\_cf}}\right)^{\beta_{cf}}$, $\frac{\beta}{\beta_{cf}}$, $\frac{L_0}{L_{0\_cf}}$ & $\frac{c}{c_{cf}}$ can be sorted in ascending order and plotted versus their median rank *P*, see next Figure obtained using Method 1.

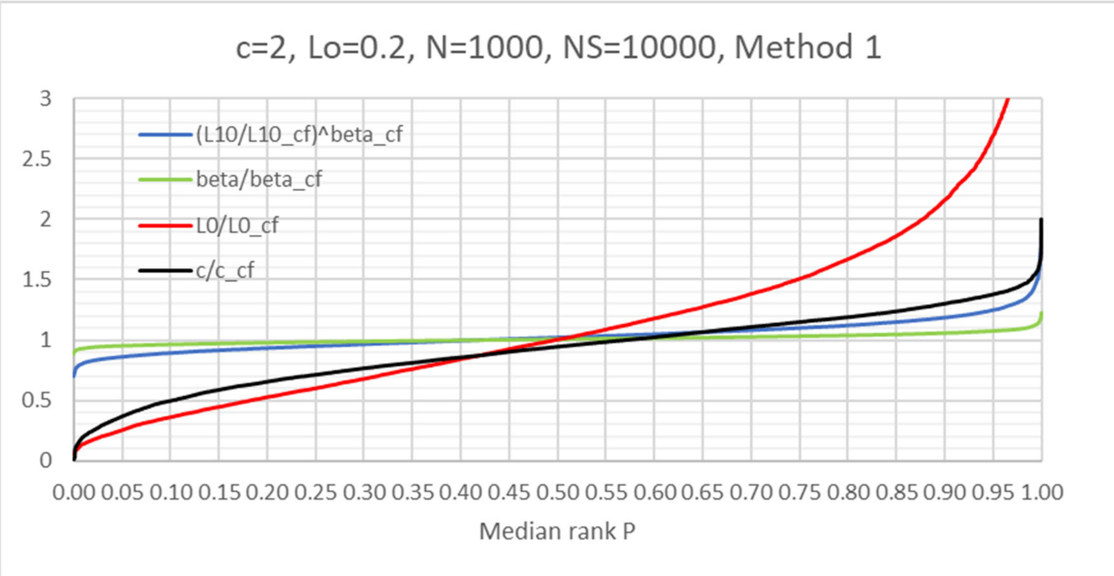


Fig. 15: Results obtained using a Monte-Carlo simulation

When fixing the median rank *P* to 0.05, 0.5 and 0.95, one can define the median 0.5 values of these ratio, as well as their 90 % confidence intervals with their lower 0.05 and upper 0.95 bounds. Of interest is also the ratio $value_{0.95}/value_{0.05}$ that one would like small and close to 1, see next table:

	Method1			
c=2	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05
(L10/L10_cf)^beta_cf	0.861	1.022	1.250	**1.451**
beta/beta_cf	0.953	1.008	1.076	**1.129**
L0/L0_cf	0.255	1.005	2.695	**10.581**
c/c_cf	0.369	0.945	1.378	**3.733**

Table 3: Example of confidence intervals obtained using method 1, N=1000, NS=10000 and L_{10}=1 β=*1*, L_0 =0.2 and c=*2*

The next table summarizes results obtained with c=2, 3 and 10:

	Method1				Method2			
c=2	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05
(L10/L10_cf)^beta_cf	0.861	1.022	1.250	**1.451**	0.847	0.982	1.112	**1.312**
beta/beta_cf	0.953	1.008	1.076	**1.129**	0.943	0.993	1.040	**1.103**
L0/L0_cf	0.255	1.005	2.695	**10.581**	0.714	1.031	1.838	**2.573**
c/c_cf	0.369	0.945	1.378	**3.733**	0.508	0.982	1.462	**2.877**
c=3	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05
(L10/L10_cf)^beta_cf	0.866	1.014	1.192	**1.376**	0.852	0.991	1.130	**1.326**
beta/beta_cf	0.954	1.005	1.062	**1.113**	0.944	0.996	1.046	**1.108**
L0/L0_cf	0.428	1.052	1.779	**4.157**	0.854	1.011	1.471	**1.721**
c/c_cf	0.344	0.916	1.485	**4.310**	0.473	1.018	1.558	**3.293**
c=10	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05
(L10/L10_cf)^beta_cf	0.894	1.034	1.170	**1.309**	0.857	0.998	1.132	**1.321**
beta/beta_cf	0.963	1.013	1.058	**1.099**	0.945	0.998	1.048	**1.109**
L0/L0_cf	0.547	0.963	1.241	**2.270**	0.922	1.008	1.184	**1.285**
c/c_cf	0.445	1.023	3.058	**6.865**	0.332	1.197	2.197	**6.618**

Table 4: Confidence intervals obtained using N=1000, NS=10000, L_{10}=1 $\beta=1$, L_0 =0.2 and c=2, 3 *or* 10

As anticipated, the lower and upper bound of the single ratio L_0/L_{0_cf} and c/c_{cf} can be quite far from 1, mainly because of the poor results obtained when c_{cf}= 1.

One can also notice that all these median ratios are slightly biased (close to 1 however because N is large). These median ratios can be used for defining correction factors and unbiased results, as shown by Houpert in [2] and Blachère in [11, 12]. Another reference (Houpert, [13]) can also be requested in which five approaches (including the MLE) are tested for defining and comparing unbiased ratios.

When using the ratio 0.95/0.05 as criterion, one sees that method 2 seems more accurate for defining L_0 and c.

Also interesting in the next Figure showing trends between L_o and c:

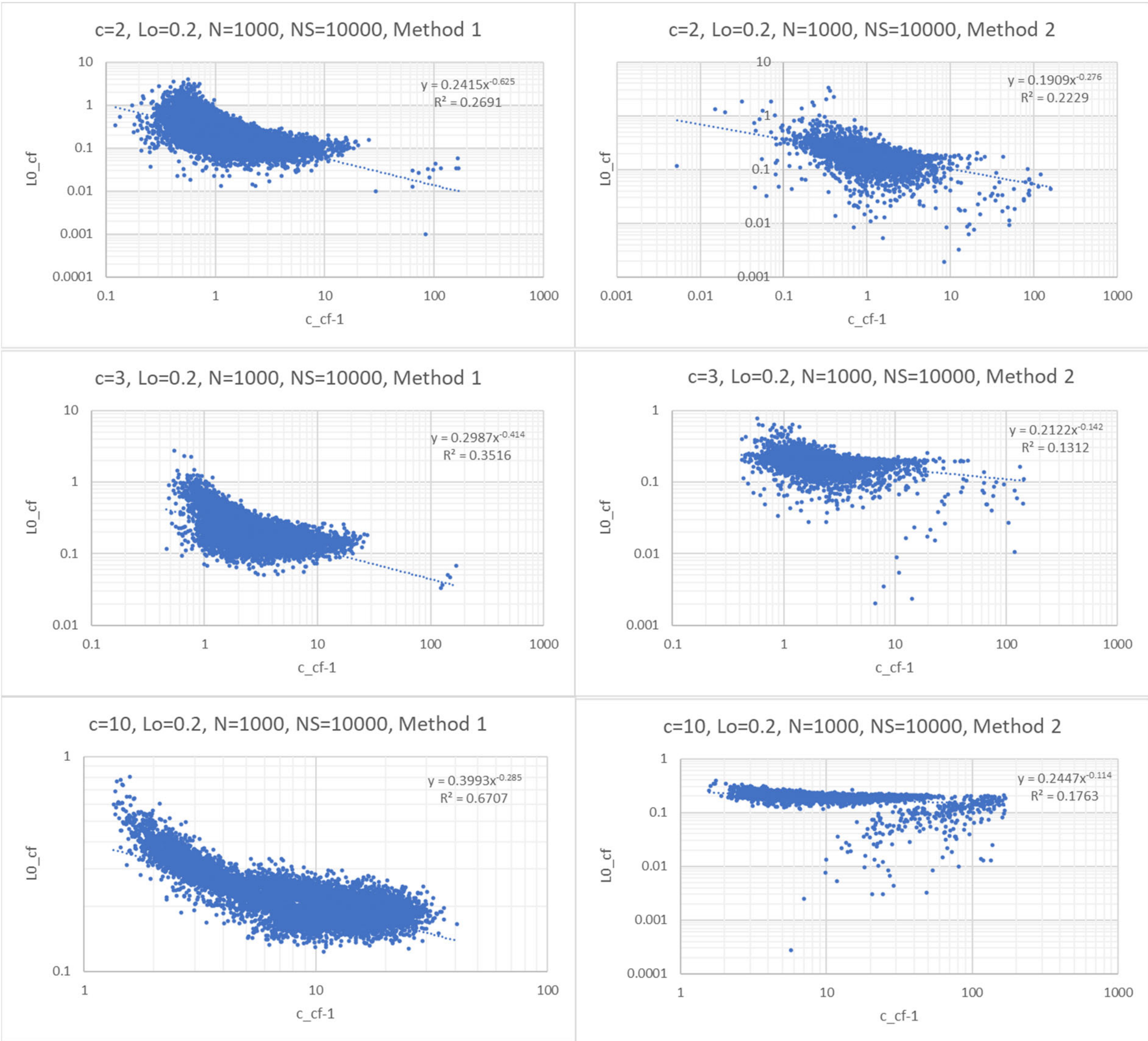


Fig. 16: Observed correlation between L_0 and c when using 10,000 simulations.

The inaccuracy or difficulty of defining L_0 and c is confirmed, as well as the suspected coupling between L_0 and c. Large values of L_{0_cf} are indeed observed when c_{cf} is small, for example with method 1 when c=2:

$$L_{o\_cf} \approx 0.2415*\left(c_{cf}-1\right)^{-0.625} \quad or \quad c_{cf} \approx 1+0.103*L_{0\_cf}^{-1.6} \quad (20)$$

For trying to understand why or how large values of L_0 can be obtained, one also plotted next the results corresponding to the maximum value of L_{0_cf} (L_{0_cf}=3.902) found using method 1. No especially abnormal values of t_{exp} at low F values are found, but surprisingly at large F values with t_{exp} values much larger than the exact 0.95 bounds of life t (leading to a large L_{10_cf} value too).

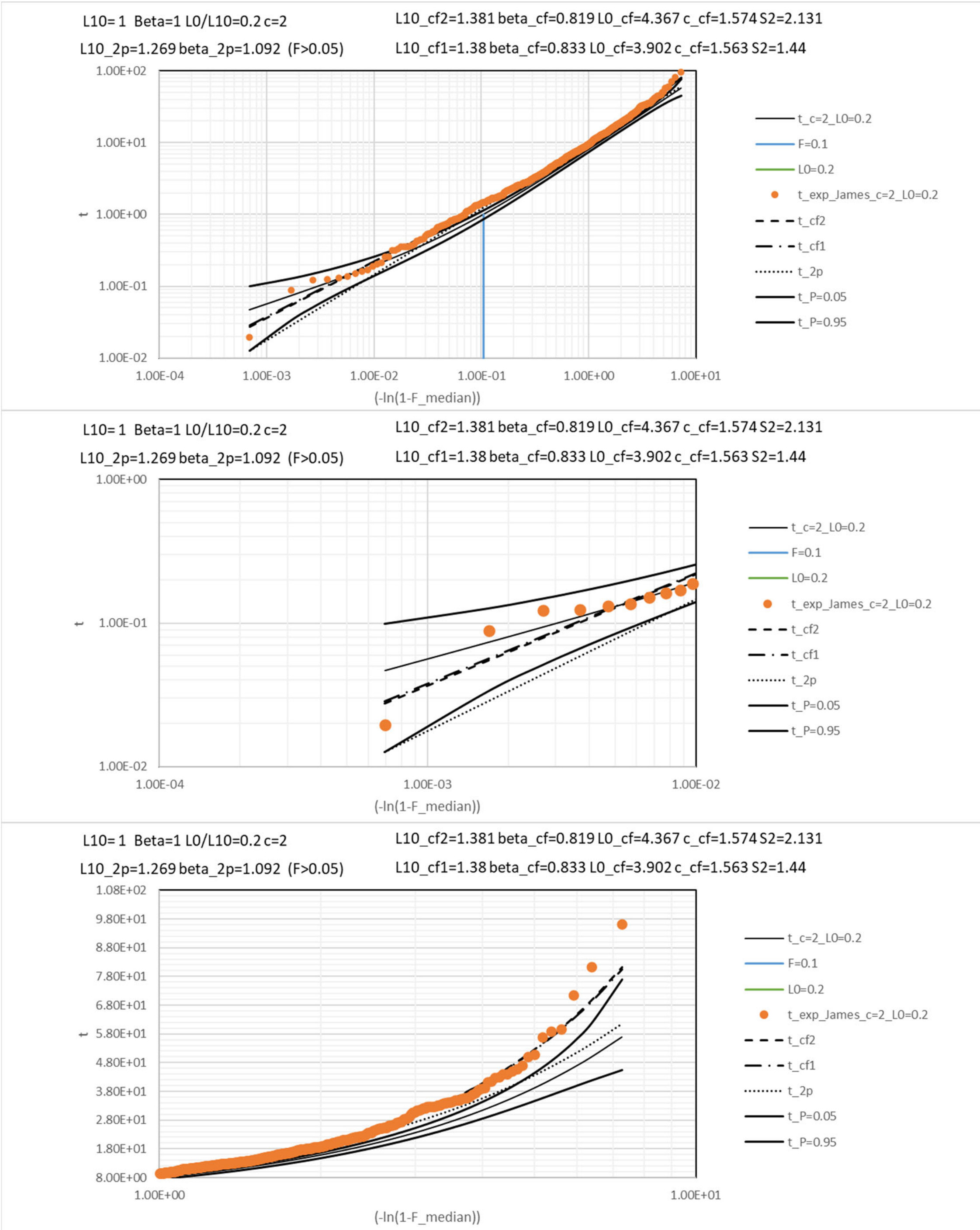


Fig. 17: Example corresponding to a large L_{o_cf} case

Last, for the sake of completeness, a Monte Carlo simulation has also been conducted using a more realistic value of N, N=100, with c=2, 3 and 10 and L_0=0.2, confirming even larger confidence intervals, see next Table and Figure.

Initial run, N=100, NS=10000								
	Method1				**Method2**			
c=2	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05
(L10/L10_cf)^beta_cf	0.700	1.031	2.074	**2.965**	0.604	0.937	1.374	**2.275**
beta/beta_cf	0.885	1.012	1.282	**1.448**	0.846	0.975	1.132	**1.338**
L0/L0_cf	0.067	1.182	5.19E+05	**7.75E+06**	0.117	1.095	1.79E+05	**1.54E+06**
c/c_cf	0.055	0.783	2.000	**36.660**	0.069	0.857	2.000	**29.115**
c=3	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05
(L10/L10_cf)^beta_cf	0.691	1.042	2.055	**2.975**	0.604	0.942	1.363	**2.255**
beta/beta_cf	0.884	1.013	1.274	**1.441**	0.845	0.974	1.129	**1.336**
L0/L0_cf	0.070	0.955	7.33E+04	**1.04E+06**	0.260	1.028	2.18E+04	**8.37E+04**
c/c_cf	0.065	0.606	3.000	**46.247**	0.067	0.662	3.000	44.764
c=10	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05
(L10/L10_cf)^beta_cf	0.670	1.042	1.779	**2.654**	0.610	0.960	1.390	**2.279**
beta/beta_cf	0.876	1.016	1.229	**1.403**	0.847	0.981	1.138	**1.343**
L0/L0_cf	0.138	0.918	3.600	**26.10**	0.600	1.059	10.471	**17.44**
c/c_cf	0.145	0.620	6.109	**42.269**	0.110	0.609	8.396	**76.103**
Second run, N=100, NS=10000								
	Method1				**Method2**			
c=2	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05
(L10/L10_cf)^beta_cf	0.697	1.030	2.047	**2.937**	0.603	0.937	1.371	**2.273**
beta/beta_cf	0.882	1.014	1.275	**1.446**	0.844	0.975	1.133	**1.342**
L0/L0_cf	0.063	1.159	5.28E+05	**8.40E+06**	0.115	1.091	1.65E+05	**1.44E+06**
c/c_cf	0.055	0.785	2.000	**36.148**	0.069	0.860	2.000	**29.124**
c=3	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05
(L10/L10_cf)^beta_cf	0.697	1.047	2.014	**2.890**	0.598	0.945	1.349	**2.256**
beta/beta_cf	0.883	1.016	1.276	**1.445**	0.846	0.975	1.130	**1.336**
L0/L0_cf	0.073	0.962	6.36E+04	**8.68E+05**	0.229	1.019	2.59E+04	**1.13E+05**
c/c_cf	0.064	0.584	3.000	**46.731**	0.067	0.710	3.000	44.958
c=10	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05	lower_0.05	Median	upper 0.95	Ratio 0.95/0.05
(L10/L10_cf)^beta_cf	0.667	1.046	1.784	**2.676**	0.613	0.955	1.386	**2.259**
beta/beta_cf	0.875	1.017	1.230	**1.406**	0.848	0.981	1.137	**1.341**
L0/L0_cf	0.139	0.925	3.993	**28.71**	0.607	1.053	9.704	**15.98**
c/c_cf	0.146	0.609	6.040	**41.309**	0.112	0.604	8.554	**76.413**
Third run, N=100, NS=100000								
					Method2			
c=2					lower_0.05	Median	upper 0.95	Ratio 0.95/0.05
(L10/L10_cf)^beta_cf					0.600	0.939	1.379	**2.296**
beta/beta_cf					0.846	0.976	1.134	**1.341**
L0/L0_cf					0.112	1.091	1.79E+05	**1.60E+06**
c/c_cf					0.067	0.854	2.000	**29.915**

Table 5: Confidence intervals obtained using N=100, NS=10000, L_{10}=1 β=*1*, L_0 =0.2 and c=2, 3 *or* 10

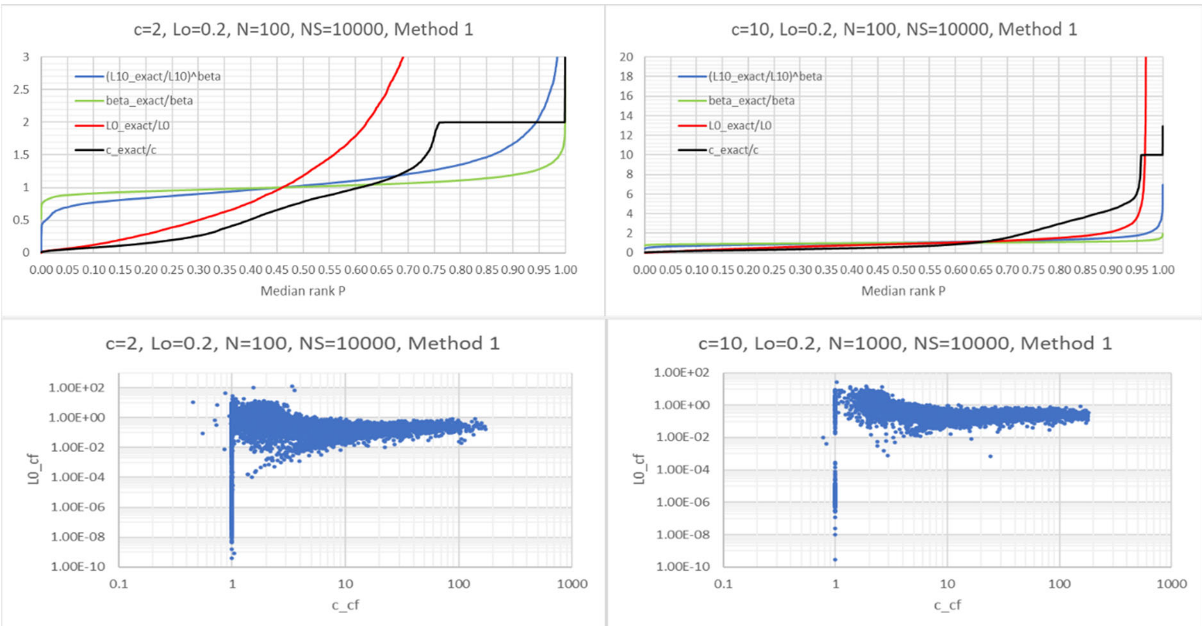


Fig. 18: Example of calculated results using N = 100 and c=2 or 10

For trying to estimate the accuracy of the numbers provided in Table 5, a second and third Monte-Carlo simulation has also been conducted using NS = 10000 and even NS= 100000 (for c=2, Method 2 only), showing minor variations of the median, lower and upper bound values. The latter run can be quite CPU time consuming. Seeking for a higher accuracy of the number provided is also difficult to justify because the numbers provided are function of the exponent c which is unknown. As an alternative and non-perfect solution to this problem, it can be suggested to conduct Monte-Carlo simulations using the experimental curve-fitted value of c before conducting next the 10,000 random simulations and defining the confidence interval on L_0/L_{0_cf} and $c/_{cf}$.

When using a reduced number of points (N =100) with c = 2, 3 or 10, the curve-fitted exponent c_{cf} is sometimes (quite often when c = 2) close to 1 meaning that any values of L_{0_cf} can be accepted since c=1 corresponds to a 2 parameter Weibull distribution in which the L_0 effect on t cancels out, see Eq. (9) for example. Also, when c_{cf}=1, all partial derivatives relative to L_0 are nil, meaning that the third equation to solve (f_3=0) is always satisfied, see Eq. (64) and (65) for example. As a results, the confidence interval on L_0 can be very large, illustrating some redundancy in Rosemann's model. Note also that the median ratio c/c_{cf} can be quite biased as a consequence when N = 100 or smaller.

These results are not very encouraging and confirm that Rosemann's model is difficult to use in practical situations when dealing with realistic endurance databases with N often smaller than 100.

The determination of L_0 for example seems quite inaccurate when using realistic N values (smaller than 100 for example) at any c values, even when c is large (equivalent to using a 3 Weibull model).

Defining Rosemann L_0 and c seems therefore very challenging when using realistic values of N (N < 100) because a few points only corresponding to low F values are available. Defining its confidence interval is even more challenging because c is unknown. As explained earlier, it can be suggested to conduct Monte-Carlo simulations using the experimental curve-fitted value of c before conducting next the 10,000 random simulations and defining the confidence interval on L_0/L_{0_cf}.

For overcoming these problems, an alternative curve-fitting and model, also using four parameters but simpler to use, is suggested in the next chapter.

Alternative New curve-fitting and New four-parameter model

New Curve-fitting:

The alternative curve-fitting is based on two linear models, the first one being simply a two-parameter model applied in the large F range, for example $F > F_{1min}$ with F_{1min} = 0.05:

$$\begin{aligned}
&For\ \ F\ >\ F_{1\min} = 0.05: \\
&Y_1 = b_1 + a_1 * X \\
&with\ \ Y_1 = \ln(t),\quad X = \ln\left(-\ln\left(1-F\right)\right) \\
&b_1 = \ln(\eta_1)\ \ \&\ \ \ a_1 = \frac{1}{\beta_1}
\end{aligned} \tag{21}$$

When sufficient points are available, a second two-parameter linear model can be tested in the low F range, for example $F < F_{2max}$ with F_{2max} = 0.01:

$$\begin{aligned}
&For\ \ F\ <\ F_{2\max} = 0.01: \\
&Y_2 = b_2 + a_2 * X \\
&with\ \ Y_2 = \ln(t),\quad X = \ln\left(-\ln\left(1-F\right)\right) \\
&b_2 = \ln(\eta_2)\ \ \&\ \ \ a_2 = \frac{1}{\beta_2}
\end{aligned} \tag{22}$$

Note that a large value of N is requested for having sufficient points to curve-fit below F_{2max}, for example N = 1000 for having (only) 10 points to curve-fit.

A slope $a_2 = 0$ corresponds to a three-parameter model; b_2 is then equal to $ln(L_0)$.

A case $a_2 = a_1$ and $b_2 = b_1$ corresponds to a two-parameter model.

The general case ($a_2 < a_1$) corresponds to a deny of a minimum life.

The latter two linear curves intersect at abscissa $X_{intersection}$ or $F_{intersection}$:

$$X_{\text{intersection}} = \frac{b_2 - b_1}{a_1 - a_2} \quad or \quad F_{\text{intersection}} = 1 - \exp\left[-\exp(X_{\text{intersection}})\right] \tag{23}$$

For ensuring a smooth transition with the latter two linear curves (considered as asymptotic values to reach when F is either very small or very large), one can suggest:

$$Y_{New} = Y_1 + \frac{Y_2 - Y_1}{1 + \left(\dfrac{F}{F_{Transition}}\right)^n} \tag{24}$$

where $F_{Transition}$ and n are theoretically two additional unknowns. Using a trial-and-error approach and Rosemann's values to benchmark against the suggested new curve-fitting, one can finally recommend the following relationship:

$$Y_{New} = Y_1 + \frac{Y_2 - Y_1}{1 + \left(\dfrac{F}{F_{Transition}}\right)^2} \quad with \quad F_{Transition} = F_{\text{intersection}} \tag{25}$$

Following are a few results obtained with the suggested new proposal when curve-fitting some results obtained with Rosemann's model, N=1000, L_{10} = 1, β = 1, L_0 = 0.2 and miscellaneous c exponents.

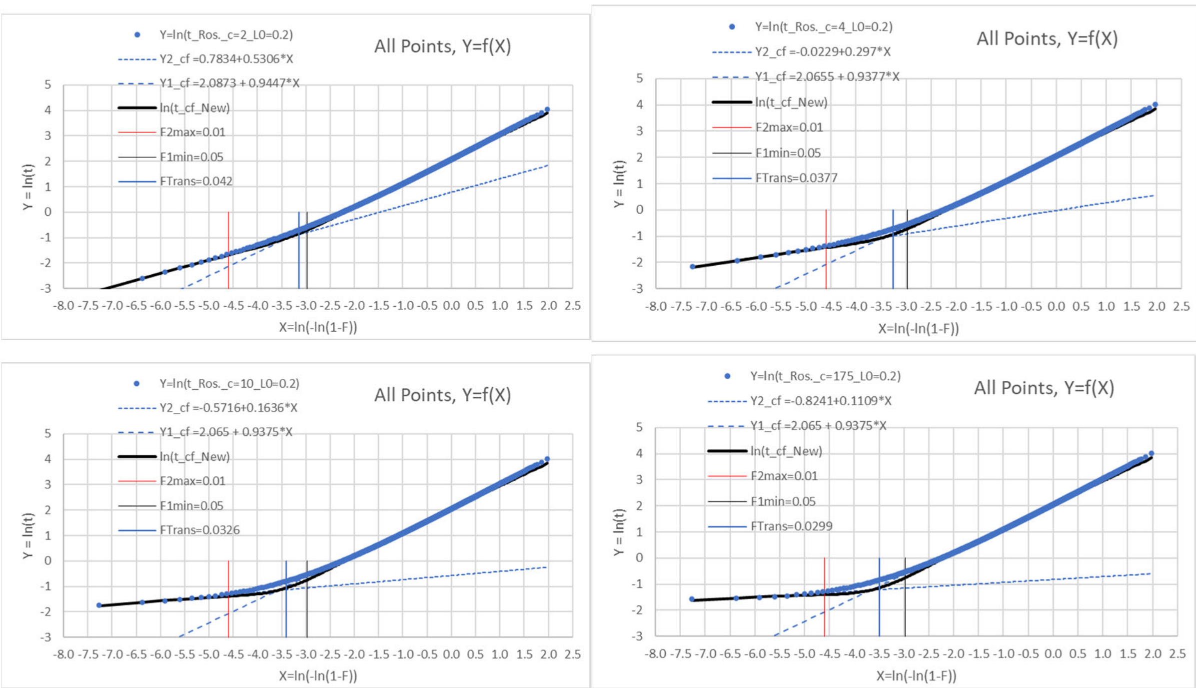


Fig. 19: New curve-fitted results obtained with N=1000, L_{10} = 1, β = 1, L_0 = 0.2 and c=2, 4, 10 and 175

As expected, there is an obvious link between c, a_2 and b_2 shown in the next Table. The value $L_{0.1}$ is used later and corresponds to the life when F = 0.001

c	b2	a2	L0.1	b1	a1	L10
2	0.7834	0.5306	0.0560	2.0873	0.9447	1
4	-0.0229	0.297	0.1256	2.0655	0.9377	1
10	-0.5716	0.1636	0.1824	2.0650	0.9375	1
175	-0.8241	0.1109	0.2039	2.0650	0.9375	1

Table 6: Rosemann versus New model correlation between c, a_2 and b_2 when L_{10}=1, a_1=1 and L_0 = 0.2

Using the linear trend observed between Y and X at low F values, two points calculated with Rosemann model at $X_{0.01}=ln(-ln(1-0.01)$ and $X_{0.001} = ln(-ln(1-0.001)$ can for example be used for approximating a_2 and b_2 as a function of c and L_0 mainly, but also a_1, b_1:

$$b_1 = \ln\left[\left(L_{10}^c + L_0^c\right)^{\frac{1}{c}} - L_0\right] - a_1.\ln\left[-\ln(0.9)\right] \text{ (exact Rosemann's model)}$$

$$Y = \frac{1}{c}.\ln\left\{\left[\exp(a_1.X + b_1) + L_0\right]^c - L_0^c\right\} \approx a_2.X + b_2 \textit{ when F is very small}$$

$$a_2 \approx \frac{1}{c}.\frac{1}{X_{0.01} - X_{0.001}}.\ln\left(\frac{\left\{\exp(a_1.X_{0.01} + b_1) + L_0\right\}^c - L_0^c}{\left\{\exp(a_1.X_{0.001} + b_1) + L_0\right\}^c - L_0^c}\right) \tag{26}$$

$$b_2 \approx \frac{1}{c}.\ln\left(\left\{\exp(a_1.X_{0.01} + b_1) + L_0\right\}^c - L_0^c\right) - a_2.X_{0.01}$$

The match between our new model and Rosemann's model is not perfect when $F_{2max} < F < F_{1min}$, but this is not our objective, our aim being to demonstrate that Rosemann's complex non-linear model behaves almost as two simple linear models (easy to curve-fit) with an appropriate smooth transition near $F_{Transition}$. Consequently, a new model exhibiting trends similar to Rosemann's ones will be introduced next.

When using a random distribution of F for generating an experimental database (based on Rosemann's model), it becomes almost impossible to distinguish the two models (Rosemann and New) with their three curve-fitted proposals (Method 1, Method 2 and New), see next example:

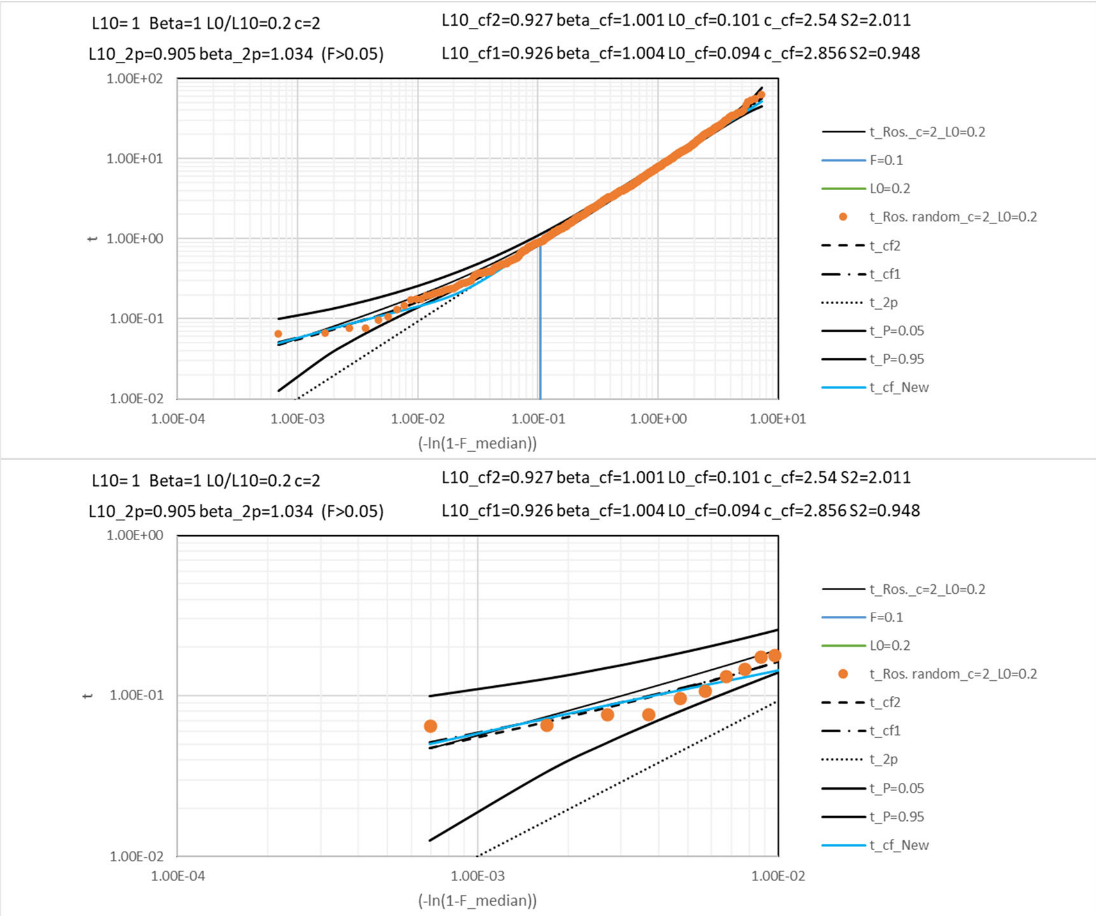


Fig. 20: Comparison between the two models (Rosemann and New) and three curve-fitted results

When conducting several random simulations, some rare abnormal cases can be found where $a_2 > a_1$, see for example the next Figure for which Method 1 would give L_0=0.0046 and c= 22.77, the final Y_cf1 curve-fitted curve matching almost a linear two-parameter curve.

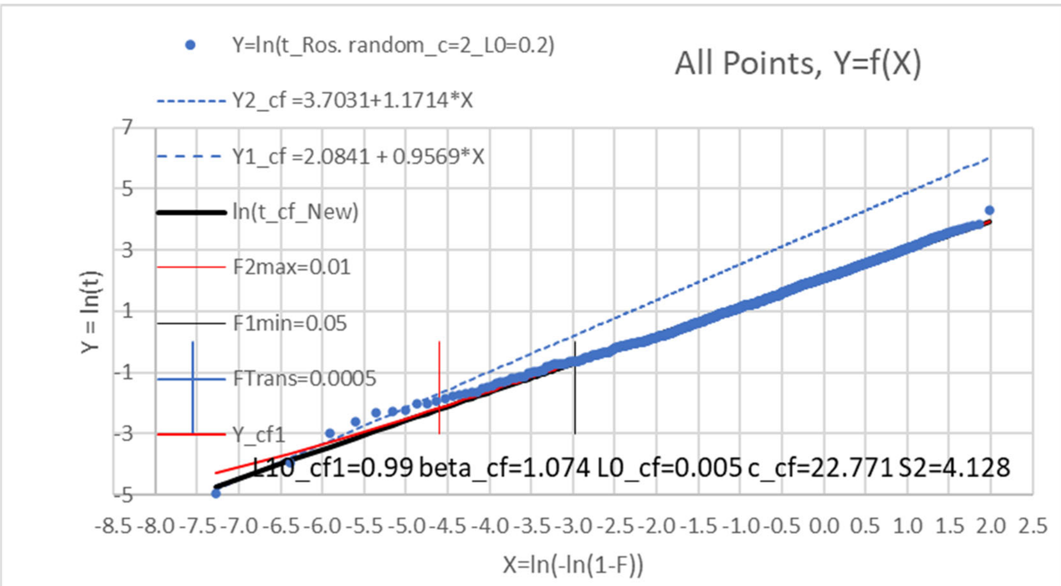


Fig. 21: Abnormal case corresponding to $a_2 > a_1$ when using the new model.

When $a_2 > a_1$, one may therefore simply suggest to reject the solution $a_2 > a_1$ and conduct a simple linear curve-fitting in the entire range of F.

Final new model suggested:

As an alternative model suggested, one can therefore suggest the following "New" model also using four parameters and the standard Y and X variables:

$$\begin{array}{l} L_{10} \quad corresponding\ to\ F = 0.1\ or\ 10\% \\ \beta_1 \quad corresponding\ to\ the\ \text{standard}\ Weibull\ slope\ (in\ the\ \text{large}\ F\ range) \\ L_{0.1} \quad corresponding\ to\ F = 0.001\ or\ 0.1\% \\ \beta_2 \quad corresponding\ to\ the\ \text{standard}\ Weibull\ slope\ (in\ the\ very\,low\ F\ range)\ with\ \beta_2 \geq \beta_1 \\ \\ X = \ln\left[-\ln(1-F)\right] \quad \& \quad Y = \ln(t) \end{array} \tag{27}$$

With the latter four inputs, one can define:

$$\begin{array}{llll} \eta_1 = \dfrac{L_{10}}{\left[-\ln(0.9)\right]^{\frac{1}{\beta_1}}} & b_1 = \ln(\eta_1) & a_1 = \dfrac{1}{\beta_1} & \\ \eta_2 = \dfrac{L_{0.1}}{\left[-\ln(0.999)\right]^{\frac{1}{\beta_2}}} & b_2 = \ln(\eta_2) & a_2 = \dfrac{1}{\beta_2} & with \quad a_2 \leq a_1 \end{array} \tag{28}$$

and finally:

$$\begin{array}{l} For\ \ F \ \geq\ F_{1\min} = 0.05: \quad Y_1 = b_1 + a_1 * X \\ For\ \ F \ \leq\ F_{2\max} = 0.01: \quad Y_2 = b_2 + a_2 * X \\ X_{\text{intersection}} = \dfrac{b_2 - b_1}{a_1 - a_2} \quad \& \quad F_{Transition} = 1 - \exp\left[-\exp(X_{\text{intersection}})\right] \\ \\ Y_{New} = Y_1 + \dfrac{Y_2 - Y_1}{1 + \left(\dfrac{F}{F_{Transition}}\right)^2} \\ t = \exp(Y_{New}) \end{array} \tag{29}$$

When defining $X_{0.01}$ and $X_{0.05}$ as the value of X=*-ln[-ln(1-F)]* calculated respectively with F =0.01 *and* 0.05, one can define realistic inputs for ensuring $X_{0.01} \leq X_{\text{intersection}} \leq X_{0.05}$ or a realistic coupling between the four parameters of the new suggested model. But the following conditions, if realistic, are not compulsory.

$$\begin{array}{lll} With\ \ a_2 \leq a_1: & b_1 + (a_1 - a_2).X_{0.01} \leq b_2 \leq b_1 + (a_1 - a_2).X_{0.05} & or \\ & L_{0.1\_\min} \leq L_{0.1} \leq L_{0.1\_\max} & \\ with: & L_{0.1\_\min} = \exp\left[b_1 + (a_1 - a_2).X_{0.01}\right] * \left[-\ln(0.999)\right]^{a_2} & \\ & L_{0.1\_\max} = \exp\left[b_1 + (a_1 - a_2).X_{0.05}\right] * \left[-\ln(0.999)\right]^{a_2} & \end{array} \tag{30}$$

The "New" model also allows describing any cases found between the standard two and three standard Weibull cases, with the possibility of denying the existence of minimum life L_0 via β_2 not infinite.

When sufficient points are available (N=1000 for example), two linear regressions can be suggested in the range $F > 0.05$ (with 950 points) and $F < 0.01$ (with 10 points), not using therefore the 40 points corresponding to $0.01 < F < 0.05$

It should of course also be possible to use all points (especially when N is not very large) for defining the four unknowns via a curve-fitting of four non-linear equations as done in this paper with Rosemann's model.

Appropriate Monte-Carlo simulations could also be conducted for defining the confidence intervals assigned to each of the four unknowns to define.

This New model should be fully tested and described (including confidence intervals) in a subsequent paper, Ref. [14].

In the latter study, 6000 relative lives (defined as L_i / $L_{15.91\_G}$) corresponding to 100 endurance tests. Each test is using 6 first in 4 lives L_i used for defining the $L_{15.91\_G}$ life of each group of 6. This approach allows the user to reach very failure rate F (also defined analytically via the inverse beta function) and obtain acceptable confidence intervals on $L_{0.1}$ for example, only slightly function of the ratio a_{2_cf}/a_{1_cf}.

Conclusions

Rosemann's 4 parameter reliability model has been studied in detail for better understanding the effect of the third and fourth parameters L_0 and c on the life. Rosemann's model is very flexible and able to describe a 2 parameter Weibull model when c = 1, or a 3 parameter Weibull model when c is infinite or very large (c=100 for example), the minimum life being then described by L_0. When c is larger than 1 (c=2, 3 or 10 for example) the existence of a minimum life is denied, the life t at low F value being smaller than L_0, a physical point that can be understood and accepted conceptually.

When generating N random values of the cumulative failure density F ($0<F<1$) and sorting these N values of F_i in an ascending order (i=1 to N), one can calculate N values of failed bearing life t_{exp_i}, (t_{exp_i} being defined as a function of F_i and Rosemann's 4 parameters), simulating hence an endurance database corresponding to a given set of N values of t_{exp_i} defined with 4 Rosemann's input parameters.

An interesting study of F_i has first been conducted for defining analytically or numerically its cumulative distribution $P(F)$. When fixing P to 0.05 or 0.5 or 0.95 for example, one can calculate the median estimate of F_i, as well as it 90 % confidence interval of F_i, hence also the median life t_{exp_i} and its confidence interval. The 'exact' median value of F_i has been obtained using the *inverse beta* function) and compared successfully to approximated values suggested in the literature.

The understanding to the 90% range of t_{exp_i} is useful for understanding why a large bearing life scatter can be obtained when N is small, at low c values especially.

Appropriate curve-fitting techniques for defining the 4 Rosemann parameters have been defined and tested (Method 1 and 2) using a few examples and have been used for anticipating large variations of the curve-fitted values L_{0_cf} and c_{cf} when c is small, the final accuracy and match to a simulated database being defined by the curve-fitted set (L_{0_cf},c_{cf}). A large value of L_{0_cf} can be compensated by a small value of c_{cf} and vice-versa.

The latter results have been confirmed by conducting Monte Carlo simulations for defining the median values and confidence intervals of the ratios $\left(\frac{L_{10}}{L_{10\_cf}}\right)^{\beta_{cf}}, \frac{\beta}{\beta_{cf}}, \frac{L_0}{L_{0\_cf}} \;\&\; \frac{c}{c_{cf}}$

Median values of these ratio are close to 1 (when N is large especially), but the 90% confidence intervals of $\frac{L_0}{L_{0\_cf}} \;\&\; \frac{c}{c_{cf}}$ can be large, at small N values especially.

Although Rosemann's model is attractive, flexible, and able to consider or deny the existence of a minimum life L_0, it's use in practical situation is difficult since the accuracy on the curve-fitted values L_{0_cf} and c_{cf} is poor, especially when $N \leq 100$, while the final accuracy using the set (L_{0_cf},c_{cf}) is satisfactory. Using the curve-fitted set (L_{0_cf} ,c_{cf}) therefore becomes risky when extrapolating the predicted life to small and untested values of F.

An alternative "New" curve-fitting technique and model (also using four parameters) have finally been suggested, with the advantages of having to conduct two simple linear curve-fittings for defining the four unknowns when enough points are available.

Acknowledgements

The authors would like to acknowledge the German Federal Ministry of Economics, Nordex SE and Thyssenkrupp Rothe Erde Germany GmbH for having sponsored this initial study (Study of Rosemann's model, analytical F values, miscellaneous curve-fitting techniques) in the frame of the HAPT2 project, as well as SMT for having sponsored the subsequent study (J. Clarke's Monte Carlo simulations and Houpert's writing of this paper).

Appendix 1: details about the analytical derivation of *P(F)*

For the sake of writing simplicity, it has been decided to attach next the index i (representing the i[th] value) to the cumulative probability P_i (hence not of F as done initially).

The cumulative density P_i (probability that the i^{th} sorted random value is smaller or equal to F) is:

$$P_i = \frac{N!}{(N-i)!.(i-1)!}.\int_0^F x^{i-1}.(1-x)^{N-i}.dx = A_i * I_i$$

$$with \quad A_i = \frac{N!}{(N-i)!.(i-1)!} \quad , \quad I_i = \int_0^F x^{b_i}.(1-x)^{c_i}.dx \quad , \quad b_i = i-1 \quad \& \quad c_i = N-i$$

$$i=1:$$

$$A_1 = N \quad , \quad b_i = 0 \quad \& \quad c_1 = N-1 \tag{31}$$

$$I_1 = \int_0^F (1-x)^{N-1}.dx = -\frac{1}{N}.\left|(1-x)^N\right|_0^F = -\frac{1}{N}.\left((1-F)^N - 1\right)$$

$$i=1: \quad P_1 = 1-(1-F)^N \qquad\qquad F = 1-(1-P_1)^{\frac{1}{N}} \tag{32}$$

$$i=2:$$

$$A_2 = N.(N-1) \quad , \quad b_2 = 1 \quad \& \quad c_2 = N-2$$

$$I_2 = \int_0^F x.(1-x)^{N-2}.dx = \int_0^F u.dv$$

$$with$$

$$u = x \qquad\qquad du = 1.dx$$

$$dv = (1-x)^{N-2}.dx \qquad\qquad v = -\frac{1}{N-1}.(1-x)^{N-1}$$

$$I_2 = \left|u.v\right|_0^F - \int_0^F v.du = -\frac{1}{N-1}.\left|x.(1-x)^{N-1}\right|_0^F + \frac{1}{N-1}.\int_0^F (1-x)^{N-1}.dx$$

$$I_2 = -\frac{1}{N-1}.F.(1-F)^{N-1} + \frac{1}{N-1}.I_1 \quad with \quad I_1 = -\frac{1}{N}.\left((1-F)^N - 1\right)$$

$$I_2 = -\frac{1}{N-1}.F.(1-F)^{N-1} - \frac{1}{N-1}.\frac{1}{N}.\left((1-F)^N - 1\right) \tag{33}$$

$$i=2: \quad P_2 = -N.F.(1-F)^{N-1} - (1-F)^N + 1 \tag{34}$$

$$i=3:$$
$$A_3=\frac{N.(N-1).(N-2)}{2} \quad , \quad b_3=2 \quad \& \quad c_3=N-3$$
$$I_3=\int_0^F x^2.(1-x)^{N-3}.dx=\int_0^F u.dv$$
$$with$$
$$u=x^2 \qquad du=2.x.dx$$
$$dv=(1-x)^{N-3}.dx \qquad v=-\frac{1}{N-2}.(1-x)^{N-2}$$
$$I_3=\left|u.v\right|_0^F-\int_0^F v.du \;=-\frac{1}{N-2}.\left|x^2.(1-x)^{N-2}\right|_0^F+\frac{2}{N-2}.\int_0^F x.(1-x)^{N-2}.dx$$
$$I_3=-\frac{1}{N-2}.F^2.(1-F)^{N-2}+\frac{2}{N-2}.I_2 \quad with \quad I_2=-\frac{1}{N-1}.F.(1-F)^{N-1}-\frac{1}{N-1}.\frac{1}{N}.\left((1-F)^N-1\right) \tag{35}$$
$$I_3=-\frac{1}{N-2}.F^2.(1-F)^{N-2}+\frac{2}{N-2}.\left\{-\frac{1}{N-1}.F.(1-F)^{N-1}-\frac{1}{N-1}.\frac{1}{N}.\left((1-F)^N-1\right)\right\}$$

$$i=3: \; P_3=-\frac{N.(N-1)}{2}.F^2.(1-F)^{N-2}-N.F.(1-F)^{N-1}-(1-F)^N+1 \tag{36}$$

$$i=4:$$
$$A_4=\frac{N.(N-1).(N-2).(N-3)}{2*3} \quad , \quad b_4=3 \quad \& \quad c_4=N-4$$
$$I_4=\int_0^F x^3.(1-x)^{N-4}.dx=\int_0^F u.dv$$
$$with$$
$$u=x^3 \qquad du=3.x^2.dx$$
$$dv=(1-x)^{N-4}.dx \qquad v=-\frac{1}{N-3}.(1-x)^{N-3}$$
$$I_4=\left|u.v\right|_0^F-\int_0^F v.du \;=-\frac{1}{N-3}.\left|x^3.(1-x)^{N-3}\right|_0^F+\frac{3}{N-3}.\int_0^F x^2.(1-x)^{N-3}.dx$$
$$I_4=-\frac{1}{N-3}.F^3.(1-F)^{N-3}+\frac{3}{N-3}.I_4 \quad with \quad I_4=-\frac{1}{N-2}.F^2.(1-F)^{N-2}+\frac{2}{N-2}.I_3 \tag{37}$$
$$I_4=-\frac{1}{N-3}.F^3.(1-F)^{N-3}+\frac{3}{N-3}.\left\{-\frac{1}{N-2}.F^2.(1-F)^{N-2}+\frac{2}{N-2}.I_3\right\}$$

While developing these calculations, a novel and useful recurrent algorithm has been found:

$$P_i = A_i.I_i \qquad with: A_i = \frac{N!}{(N-i)!.(i-1)!} \qquad I_i = Diagonal_i.\left|x^{i-1}.(1-x)^{N-i+1}\right|_0^F + factor_i.I_{i-1}$$
$$Diagonal_i = -\frac{1}{N-i+1} \quad \& \quad factor_i = \frac{i-1}{N-i+1} \tag{38}$$

Leading to the following final analytical result:

$$in\ general:$$
$$Starting\ with\ \ Coef_1 = 1\ \ \&\ \ P_1 = 1-(1-F)^N$$
$$Coef_i = Coef_{i-1}.\frac{N-i+2}{i-1}$$
$$P_i = P_{i-1} - Coef_i.F^{i-1}.(1-F)^{N-i+1} \qquad \left(also: P_i = P_{i-1} - \frac{N!}{(i-1)!.(N-i+1)!}.F^{i-1}.(1-F)^{N-i+1}\right)$$
$$hence:$$
$$P_1 = 1-(1-F)^N$$
$$P_2 = 1-(1-F)^N - N.F.(1-F)^{N-1}$$
$$P_3 = 1-(1-F)^N - N.F.(1-F)^{N-1} - \frac{N.(N-1)}{2}.F^2.(1-F)^{N-2}$$
$$P_4 = 1-(1-F)^N - N.F.(1-F)^{N-1} - \frac{N.(N-1)}{2}.F^2.(1-F)^{N-2} - \frac{N.(N-1).(N-2)}{2*3}.F^3.(1-F)^{N-3}$$
$$P_5 = 1-(1-F)^N - N.F.(1-F)^{N-1} - \frac{N.(N-1)}{2}.F^2.(1-F)^{N-2} - \frac{N.(N-1).(N-2)}{2*3}.F^3.(1-F)^{N-3}$$
$$- \frac{N.(N-1).(N-2).(N-3)}{2*3*4}.F^4.(1-F)^{N-4}$$
$$P_6 = 1-(1-F)^N - N.F.(1-F)^{N-1} - \frac{N.(N-1)}{2}.F^2.(1-F)^{N-2} - \frac{N.(N-1).(N-2)}{2*3}.F^3.(1-F)^{N-3} \tag{39}$$
$$- \frac{N.(N-1).(N-2).(N-3)}{2*3*4}.F^4.(1-F)^{N-4} - \frac{N.(N-1).(N-2).(N-3).(N-4)}{2*3*4*5}.F^5.(1-F)^{N-5}$$
$$P_7 = P_6 - \frac{N.(N-1).(N-2).(N-3).(N-4).(N-5)}{2*3*4*5*6}.F^6.(1-F)^{N-6}$$
$$P_8 = P_7 - \frac{N.(N-1).(N-2).(N-3).(N-4).(N-5).(N-6)}{2*3*4*5*6*7}.F^7.(1-F)^{N-7}$$

General analytical relationships have also been developed when decreasing i from N to 1 although large failure rate results are usually of little interest.

$$i = N:$$
$$A_N = N \quad , \quad b_N = N-1 \quad \& \quad c_N = 0 \tag{40}$$
$$I_N = \int_0^F x^{N-1}.dx = \frac{1}{N}.\left|x^N\right|_0^F = \frac{1}{N}.F^N$$

$$P_N = F^N \qquad\qquad F = P_N^{\frac{1}{N}} \tag{41}$$

$i = N-1:$

$$A_{N-1} = N.(N-1) \quad , \; b_{N-1} = N-2 \quad \& \quad c_{N-1} = 1 \tag{42}$$

$$I_{N-1} = \int_0^F x^{N-2}.(1-x).dx = \frac{1}{N-1}.\left|x^{N-1}\right|_0^F - \frac{1}{N}.\left|x^N\right|_0^F$$

$$P_{N-1} = N.F^{N-1} - (N-1).F^N \tag{43}$$

$i = N-2:$

$$A_{N-2} = \frac{N.(N-1).(N-2)}{2} \; , \; b_{N-2} = N-3 \quad \& \quad c_{N-2} = 2$$

$$I_{N-2} = \int_0^F x^{N-3}.(1-x)^2.dx = \int_0^F \left(x^{N-3} - 2.x^{N-2} + x^{N-1}\right).dx = \tag{44}$$

$$I_{N-2} = \frac{1}{N-2}.\left|x^{N-2}\right|_0^F - \frac{2}{N-1}.\left|x^{N-1}\right|_0^F + \frac{1}{N}.\left|x^N\right|_0^F$$

$$P_{N-2} = \frac{N.(N-1)}{2}.F^{N-2} - N.(N-2).F^{N-1} + \frac{(N-1).(N-2)}{2}.F^N \tag{45}$$

$i = N-3:$

$$A_{N-3} = \frac{N.(N-1).(N-2).(N-3)}{2*3} \; , \; b_{N-3} = N-4 \quad \& \quad c_N = 3$$

$$I_{N-3} = \int_0^F x^{N-4}.(1-x)^3.dx = \int_0^{F_{N-2}} \left(x^{N-4} - 3.x^{N-3} + 3.x^{N-2} - x^{N-1}\right).dx = \tag{46}$$

$$I_{N-3} = \frac{1}{N-3}.\left|x^{N-3}\right|_0^F - \frac{3}{N-2}.\left|x^{N-2}\right|_0^F + \frac{3}{N-1}.\left|x^{N-1}\right|_0^F - \frac{1}{N}.\left|x^N\right|_0^F$$

$$P_{N-3} = \frac{N.(N-1).(N-2)}{2*3}.F^{N-3} - \frac{N.(N-1).(N-3)}{2}.F^{N-2} + \frac{N.(N-2).(N-3)}{2}.F^{N-1} - \frac{(N-1).(N-2).(N-3)}{2*3}.F^N \tag{47}$$

So:

$$
\begin{aligned}
&P_N = F^N \\
&P_{N-1} = N.F^{N-1} - (N-1).F^N \\
&P_{N-2} = \frac{N.(N-1)}{2}.F^{N-2} - N.(N-2).F^{N-1} + \frac{(N-1).(N-2)}{2}.F^N \\
&P_{N-3} = \frac{N.(N-1).(N-2)}{2*3}.F^{N-3} - \frac{N.(N-1).(N-3)}{2}.F^{N-2} + \frac{N.(N-2).(N-3)}{2}.F^{N-1} - \frac{(N-1).(N-2).(N-3)}{2*3}.F^N
\end{aligned}
\tag{48}
$$

As before, a recurrent numerical algorithm can also be suggested. The only interest of trying to develop an analytical recurrent algorithm is that when calculating $P_{i=950}$ for example, one does not need to calculate the previous 949 values of P, but only 49 values in a decreasing order, starting with i =1000 in our example.

$$
\begin{aligned}
&P_i = A_i.I_i \quad with \\
&A_i = \frac{N!}{(N-i)!.(i-1)!} \qquad and \qquad I_i = \int_0^F x^{i-1}.(1-x)^{N-i} \\
&Starting\ with\ i = N: \\
&A_N = N \qquad\qquad I_N = \int_0^F x^{N-1}.dx = \frac{1}{N}.\left|x^N\right|_0^F = \frac{1}{N}.F^N
\end{aligned}
\tag{49}
$$

Recurrent algorithm:

$$
A_i = \frac{N!}{(N-i)!.(i-1)!} = A_{i+1}.\frac{i}{N-i}
$$

$$
I_i = \int_0^F x^{i-1}.(1-x)^{N-i}.dx = \int u.dv
$$

$$
u = (1-x)^{N-i} \qquad du = -\left(N-i\right).(1-x)^{N-i-1}.dx
$$

$$
dv = x^{i-1}.dx \qquad v = \frac{1}{i}.x^i
$$

$$
I_i = \left|u.v\right|_0^F - \int_0^F v.du = \frac{1}{i}.F^i.(1-F)^{N-i} + \frac{(N-i)}{i}.\underbrace{\int_0^F x^i.(1-x)^{N-i-1}.dx}_{I_{i+1}}
$$

$$
I_i = \frac{1}{i}.F^i.(1-F)^{N-i} + \frac{(N-i)}{i}.I_{i+1} \tag{50}
$$

So, one can finally use:

$$
\begin{aligned}
&P_i = A_i.I_i \\
&with: \\
&A_i = A_{i+1}.\frac{i}{N-i} \quad \& \quad I_i = \frac{1}{i}.F^i.(1-F)^{N-i} + \frac{(N-i)}{i}.I_{i+1} \\
&starting\ with \\
&A_N = N \quad \& \quad I_N = \frac{1}{N}.F^N
\end{aligned} \tag{51}
$$

Note that one can also write:

$$
\begin{aligned}
&P_i = \frac{A_{i+1}}{N-i}.F^i.(1-F)^{N-i} + P_{i+1} = \frac{A_i}{i}.F^i.(1-F)^{N-i} + P_{i+1} \\
&with: \\
&A_i = A_{i+1}.\frac{i}{N-i} \\
&starting\ with \\
&A_N = N \quad \& \quad P_N = F^N
\end{aligned} \tag{52}
$$

It can also be demonstrated that:

$$
\begin{aligned}
&For\ i < \frac{N}{2}: \\
&F_{N-i+1\_P=0.95} = 1 - F_{i\_P=0.05} \\
&F_{N-i+1\_P=0.5} = 1 - F_{i\_P=0.5} \\
&F_{N-i+1\_P=0.05} = 1 - F_{i\_P=0.95}
\end{aligned} \tag{53}
$$

Appendix 2:
Explanations about the Maximum Likelihood (ML) approach using Rosemann's model

$$F = 1 - \exp\left[-\left(\frac{\left(t^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\eta}\right)^{\beta}\right]$$

$$f(t) = \frac{dF}{dt} = \frac{\beta}{\eta}.\exp\left[-\left(\frac{\left(t^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\eta}\right)^{\beta}\right].\left(\frac{\left(t^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\eta}\right)^{\beta-1}.\left(t^c + L_0^c\right)^{\frac{1}{c}-1}.t^{c-1}$$

$$\text{Product} = \prod_{i=1}^{N} f(t_i)$$

$$\ln\left[\text{Product}\right] = N.\ln(\beta) - N.\ln(\eta) - \sum_{i=1,N}\left(\frac{\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\eta}\right)^{\beta} + (\beta-1).\sum_{i=1,N}\ln\left(\frac{\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\eta}\right)$$
$$+\left(\frac{1}{c}-1\right).\sum_{i=1,N}\ln\left(t_i^c + L_0^c\right) + (c-1).\sum_{i=1,N}\ln(t_i)$$
$$= Max$$

$$\ln\left[\text{Product}\right] = N.\ln(\beta) - N.\beta.\ln(\eta) - \sum_{i=1,N}\left(\frac{\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\eta}\right)^{\beta} - (\beta-1).\sum_{i=1,N}\ln\left(\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0\right)$$
$$+\left(\frac{1}{c}-1\right).\sum_{i=1,N}\ln\left(t_i^c + L_0^c\right) + (c-1).\sum_{i=1,N}\ln(t_i) \quad (54)$$
$$= Max$$

$$N.\ln(\beta) - N.\beta.\ln(\eta) - \sum_{i=1,N}\left(\frac{\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\eta}\right)^{\beta} - (\beta-1).\sum_{i=1,N}\ln\left(\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0\right) + \left(\frac{1}{c}-1\right).\sum_{i=1,N}\ln\left(t_i^c + L_0^c\right) + (c-1).\sum_{i=1,N}\ln(t_i) = Max \quad (55)$$

The ML approach consists therefore to solve:

$$F(\beta,\eta,L_0,c) =$$
$$N.\ln(\beta) - N.\beta.\ln(\eta) - \frac{1}{\eta^{\beta}}.\sum_{i=1,N}\left(\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0\right)^{\beta} - (\beta-1).\sum_{i=1,N}\ln\left(\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0\right) + \left(\frac{1}{c}-1\right).\sum_{i=1,N}\ln\left(t_i^c + L_0^c\right) + (c-1).\sum_{i=1,N}\ln(t_i) \quad (56)$$
$$= Max$$

via a set of 4 non-linear equations:

$$f_1(\beta,\eta,L_0,c) = \frac{dF}{d\beta} = N.\frac{1}{\beta} - N.\ln(\eta) - \frac{1}{\eta^{\beta}}.\left\{\sum_{i=1,N}\left(\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0\right)^{\beta}.\ln\left(\frac{\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0}{\eta}\right)\right\} - \sum_{i=1,N}\ln\left(\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0\right) = 0 \quad (57)$$

$$f_2(\beta,\eta,L_0,c) = \frac{dF}{d\eta} = -N.\beta.\frac{1}{\eta} + \beta.\frac{1}{\eta^{\beta+1}}.\sum_{i=1,N}\left(\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0\right)^{\beta} = 0 \quad (58)$$

$$f_3(\beta,\eta,L_0,c) = \frac{dF}{dL_0} - \frac{\beta}{\eta^{\beta}}.\sum_{i=1,N}\left(\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0\right)^{\beta-1}.\left[\left(t_i^c + L_0^c\right)^{\frac{1}{c}-1}.L_0^{c-1} - 1\right] - (\beta-1).\sum_{i=1,N}\left(\frac{\left(t_i^c + L_0^c\right)^{\frac{1}{c}-1}L_0^{c-1} - 1}{\left(t_i^c + L_0^c\right)^{\frac{1}{c}} - L_0}\right) + (1-c).\sum_{i=1,N}\frac{L_0^{c-1}}{t_i^c + L_0^c} = 0 \quad (59)$$

The fourth derivation requires to use:

$$\frac{d\left[\left(t_{\exp\_i}^{c}+L_{o}^{c}\right)^{\frac{1}{c}}\right]}{dc}=D_{i}=\frac{\left(t_{\exp\_i}^{c}+L_{o}^{c}\right)^{\frac{1}{c}-1}.\left[-\left(t_{\exp\_i}^{c}+L_{o}^{c}\right).\ln\left(t_{\exp\_i}^{c}+L_{o}^{c}\right)+L_{0}^{c}.c.\ln(L_{0})+t_{\exp\_i}^{c}.c.\ln(t_{\exp\_i})\right]}{c^{2}} \quad (60)$$

$$f_{4}(\beta,\eta,L_{0},c)=\frac{dF}{dc}-\frac{\beta}{\eta^{\beta}}.\sum_{i=1,N}\left\{\left(\left(t_{i}^{c}+L_{0}^{c}\right)^{\frac{1}{c}}-L_{0}\right)^{\beta-1}.D_{i}\right\}-\left(\beta-1\right).\sum_{i=1,N}\frac{D_{i}}{\left(t_{i}^{c}+L_{0}^{c}\right)^{\frac{1}{c}}-L_{0}}.+\left(1-c\right).\sum_{i=1,N}\frac{t_{i}^{c1}}{t_{i}^{c}+L_{0}^{c}}+c.\sum_{i=1,N}\ln(t_{i}) \quad (61)$$

Appendix 3: Method 1

$$Y_{cf\_i} = \ln\left(t_{cf\_i}\right) = \frac{1}{c}.\ln\left[\left\{\exp\left(a.X_i + b\right) + L_0\right\}^c - L_0^c\right] \quad to \quad compare \ to \ Y_{\exp\_i} = \ln\left(t_{\exp\_i}\right)$$
$$via:$$
$$S_i = Y_{cf\_i} - Y_{\exp\_i}$$
$$with: \tag{62}$$
$$X_i = \ln\left(-\ln\left(1 - F_{median\_i}\right)\right) \ and \ 4 \ unknowns: a = \frac{1}{\beta} \qquad b = \ln(\eta) \quad L_0 \ \ \& \ \ c$$
$$Y_{cf\_i} = \ln\left(t_{cf\_i}\right) = \frac{1}{c}.\ln\left[Z_i\right] \ with \ Z_i = \left\{\exp\left(a.X_i + b\right) + L_0\right\}^c - L_0^c$$

Minimizing S^2 can be done by solving 4 non-linear equations:

$$\begin{cases} f_1(a,b,L_0,c) = \dfrac{dS^2}{da} = 2.\sum\limits_{i=1,N} S_i.\dfrac{dS_i}{da} = 2.\sum\limits_{i=1,N} S_i.\dfrac{dY_{cf\_i}}{da} = 0 \\ f_2(a,b,L_0,c) = \dfrac{dS^2}{db} = 2.\sum\limits_{i=1,N} S_i.\dfrac{dS_i}{db} = 2.\sum\limits_{i=1,N} S_i.\dfrac{dY_{cf\_i}}{db} = 0 \\ f_3(a,b,L_0,c) = \dfrac{dS^2}{dL_0} = 2.\sum\limits_{i=1,N} S_i.\dfrac{dS_i}{dL_0} = 2.\sum\limits_{i=1,N} S_i.\dfrac{dY_{cf\_i}}{dL_0} = 0 \\ f_4(a,b,L_0,c) = \dfrac{dS^2}{dc} = 2.\sum\limits_{i=1,N} S_i.\dfrac{dS_i}{dc} = 2.\sum\limits_{i=1,N} S_i.\dfrac{dY_{cf\_i}}{dc} = 0 \end{cases} \tag{63}$$

The factor 2 can of course be eliminated.

One can calculate numerically (using finite differences) or analytically (recommended approach because more accurate) the derivative of Y_{cf} versus any unknow:

$$\frac{dS_i}{d(unknown)} = \frac{dY_{cf\_i}}{d(unknown)} = \frac{1}{c}.\frac{1}{Z_i}.\frac{dZ_i}{d(unknown)} \quad when \ \ unknown = a, b \, or \ L_0$$
$$but: \frac{dS_i}{dc} = \frac{dY_{cf\_i}}{dc} = -\frac{1}{c^2}.\ln(Z_i) + \frac{1}{c}.\frac{1}{Z_i}.\frac{dZ_i}{dc} \tag{64}$$

When selecting the analytical approach, one therefore needs to define analytically the four following derivatives:

$$
\begin{aligned}
&\mathit{Method}\ 1:\\
&\frac{dZ_i}{da} = c.\left[\exp(a.X_i+b)+L_0\right]^{c-1}.\exp(a.X_i+b).X_i\\
&\frac{dZ_i}{db} = c.\left[\exp(a.X_i+b)+L_0\right]^{c-1}.\exp(a.X_i+b)\\
&\frac{dZ_i}{dL_0} = c.\left[\exp(a.X_i+b)+L_0\right]^{c-1} - c.L_0^{c-1}\\
&\frac{dZ_i}{dc} = \left[\exp(a.X_i+b)+L_0\right]^{c}.\ln\left[\exp(a.X_i+b)+L_0\right] - L_0^c.\ln\left(L_0\right)\\
&\text{with as final objective:}\\
&\begin{cases}
f_1 = \dfrac{1}{c}.\displaystyle\sum_{i=1,N} S_i.\dfrac{1}{Z_i}.\dfrac{dZ_i}{da} = 0\\
f_2 = \dfrac{1}{c}.\displaystyle\sum_{i=1,N} S_i.\dfrac{1}{Z_i}.\dfrac{dZ_i}{db} = 0\\
f_3 = \dfrac{1}{c}.\displaystyle\sum_{i=1,N} S_i.\dfrac{1}{Z_i}.\dfrac{dZ_i}{dL_0} = 0\\
f_4 = \dfrac{1}{c}.\displaystyle\sum_{i=1,N} S_i.\left(-\dfrac{1}{c}.\ln(Z_i)+\dfrac{1}{Z_i}.\dfrac{dZ_i}{dc}\right) = 0
\end{cases}\\
&\mathit{and}:\\
&Z_i = \left\{\exp\left(a.X_i+b\right)+L_0\right\}^c - L_0^c\\
&X_i = \ln\left(-\ln\left(1-F_{median\_i}\right)\right)\\
&S_i = \frac{1}{c}.\ln(Z_i) - \ln(t_{\exp\_i})
\end{aligned}
\tag{65}
$$

Solving any of the previously defined equations $f_i(a,b,L_0,c) = 0$ (for i=1 to 4) is the next step.

This can be done by using a first order Taylor approximation and writing:

$$
\begin{aligned}
&f_i(a+\Delta a, b+\Delta b, L_0+\Delta L_0, c+\Delta c) = f_i(a,b,L_0,c) + \frac{df_i}{da}.\Delta a + \frac{df_i}{db}.\Delta b + \frac{df_i}{dL_0}.\Delta L_0 + \frac{df_i}{dc}.\Delta c = 0\\
&\mathit{or}:\\
&\frac{df_i}{da}.\Delta a + \frac{df_i}{db}.\Delta b + \frac{df_i}{dL_0}.\Delta L_0 + \frac{df_i}{dc}.\Delta c = -f_i(a,b,L_0,c)
\end{aligned}
\tag{66}
$$

Meaning that one needs to solve in an iterative manner a linear set of 4 equations with 4 unknowns: Δa, Δb, ΔL_0 and Δc

One now needs to define the partial derivative matrix:

$$\begin{bmatrix} \frac{df_1}{da} & \frac{df_1}{db} & \frac{df_1}{dL_0} & \frac{df_1}{dc} \\ \frac{df_2}{da} & \frac{df_2}{db} & \frac{df_2}{dL_0} & \frac{df_2}{dc} \\ \frac{df_3}{da} & \frac{df_3}{db} & \frac{df_3}{dL_0} & \frac{df_3}{dc} \\ \frac{df_4}{da} & \frac{df_4}{db} & \frac{df_4}{dL_0} & \frac{df_4}{dc} \end{bmatrix} \tag{67}$$

These partial derivatives could perhaps be calculated again analytically but for the sake of simplicity, a numerical approach has been preferred, the four terms of any given line "*i*" of the partial derivative matrix being defined using finite differences:

$$\begin{aligned} \frac{df_i}{da} &= \frac{f_i(a+da,b,L_0,c)-f_i(a-da,b,L_0,c)}{2.da} \\ \frac{df_i}{db} &= \frac{f_i(a,b+db,L_0,c)-f_i(a,b-db,L_0,c)}{2.db} \\ \frac{df_i}{dL_0} &= \frac{f_i(a,b,L_0+dL_0,c)-f_i(a,b,L_0-dL_0,c)}{2.dL_0} \\ \frac{df_i}{dc} &= \frac{f_i(a,b,L_0,c+dc)-f_i(a,b,L_0,c+dc)}{2.dc} \end{aligned} \tag{68}$$

The increments *da, db* and *dc* have fixed to 0.01 while dL_0 was fixed to 0.001

Appendix 4: Method 2

The approach is simpler (relative to Method 1) to develop and program:

$$S_i = a.\ln\left\{\left(t_{\exp\_i}^c + L_0^c\right)^{\frac{1}{c}} - L_0\right\} + b - X_i$$
$$with\ X_i = \ln\left(-\ln\left(1 - F_{median\_i}\right)\right) \tag{69}$$
$$with\ 4\ unknowns:$$
$$a = \beta \quad b = -\beta.\ln(\eta) \quad L_0 \quad and \quad c$$

$$\frac{dS_i}{da} = \ln\left[\left(t_{\exp\_i}^c + L_0^c\right)^{\frac{1}{c}} - L_0\right]$$
$$\frac{dS_i}{db} = 1 \tag{70}$$
$$\frac{dS_i}{dL_0} = \frac{a}{\left(t_{\exp\_i}^c + L_0^c\right)^{\frac{1}{c}} - L_0}.\left[\left(t_{\exp\_i}^c + L_0^c\right)^{\frac{1}{c}-1}.L_0^{c-1} - 1\right]$$

$$\frac{dS_i}{dc} = \frac{a}{\left(t_{\exp\_i}^c + L_0^c\right)^{\frac{1}{c}} - L_0}.\frac{d\left[\left(t_{\exp\_i}^c + L_0^c\right)^{\frac{1}{c}}\right]}{dc} \tag{71}$$

The latter derivative can be obtained using an online application:

$$\frac{\partial}{\partial x}\left(\sqrt[x]{\mathrm{L0}^x + t^x}\right) = \frac{\left(\mathrm{L0}^x + t^x\right)^{1/x-1}\left(-\left(\mathrm{L0}^x + t^x\right)\log\left(\mathrm{L0}^x + t^x\right) + x\,\mathrm{L0}^x\log(\mathrm{L0}) + x\,t^x\log(t)\right)}{x^2} \tag{72}$$

hence:

$$\frac{d\left[\left(t_{\exp\_i}^c + L_o^c\right)^{\frac{1}{c}}\right]}{dc} = \frac{\left(t_{\exp\_i}^c + L_o^c\right)^{\frac{1}{c}-1}.\left[-\left(t_{\exp\_i}^c + L_o^c\right).\ln\left(t_{\exp\_i}^c + L_o^c\right) + L_0^c.c.\ln(L_0) + t_{\exp\_i}^c.c.\ln(t_{\exp\_i})\right]}{c^2} \tag{73}$$

leading to:

$$\frac{dS_i}{dc} = \frac{a}{c^2}.\frac{\left(t_{\exp\_i}^c + L_o^c\right)^{\frac{1}{c}-1}.\left[-\left(t_{\exp\_i}^c + L_o^c\right).\ln\left(t_{\exp\_i}^c + L_o^c\right) + L_0^c.\ln(L_0^c) + t_{\exp\_i}^c.\ln(t_{\exp\_i}^c)\right]}{\left(t_{\exp\_i}^c + L_0^c\right)^{\frac{1}{c}} - L_0} \tag{74}$$

The set of 4 non-linear equations to solve is describe by the following, (i = 1 to 4):

$$\begin{cases} f_1 = \sum_{i=1,N} S_i.\frac{dS_i}{da} = 0 \\ f_2 = \sum_{i=1,N} S_i.\frac{dS_i}{db} = 0 \\ f_3 = \sum_{i=1,N} S_i.\frac{dS_i}{dL_0} = 0 \\ f_4 = \sum_{i=1,N} S_i.\frac{dS_i}{dc} = 0 \end{cases}$$

$$f_i(a+\Delta a, b+\Delta b, L_0+\Delta L_0, c+\Delta c) = f_i(a,b,L_0,c) + \frac{df_i}{da}.\Delta a + \frac{df_i}{db}.\Delta b + \frac{df_i}{dL_0}.\Delta L_0 + \frac{df_i}{dc}.\Delta c = 0$$

$$or:$$

$$\frac{df_i}{da}.\Delta a + \frac{df_i}{db}.\Delta b + \frac{df_i}{dL_0}.\Delta L_0 + \frac{df_i}{dc}.\Delta c = -f_i(a,b,L_0,c) \qquad (75)$$

with:

$$\frac{df_i}{da} = \frac{f_i(a+da,b,L_0,c) - f_i(a-da,b,L_0,c)}{2.da}$$
$$\frac{df_i}{db} = \frac{f_i(a,b+db,L_0,c) - f_i(a,b-db,L_0,c)}{2.db}$$
$$\frac{df_i}{dL_0} = \frac{f_i(a,b,L_0+dL_0,c) - f_i(a,b,L_0-dL_0,c)}{2.dL_0} \qquad (76)$$
$$\frac{df_i}{dc} = \frac{f_i(a,b,L_0,c+dc) - f_i(a,b,L_0,c+dc)}{2.dc}$$

The increments *da, db* and *dc* have fixed to 0.01 while dL_0 was fixed to 0.001

References:

1. H. Rosemann, 'The Weibull Distribution and the Problem of Guaranteed Minimum Lifetimes', Bearing World Journal (2021), https://doi.org/10.15488/11469

2. L. Houpert, 'An engineering approach to confidence intervals & endurance test strategies', STLE Tribology Trans., Vol 46 (2003), 2, 248-259.

3. M.N. Kotzalas, 'Statistical Distribution of Tapered Roller Bearing Fatigue Lives at High Levels of Reliability', ASME Journal of Tribology, October 2005, Vol. 127
4. L. Houpert, 'Reliability factor a_1 written for ISO-TR-1281', 2005

5. T. Tallian, 1962, 'Weibull Distribution of Rolling Contact Fatigue Life and Deviations Therefrom,' ASLE Trans., 5, pp. 183–196.

6. B. Snare, 1970, 'How Reliable are Bearings?,' Ball Bearing Journal, 162, pp. 3–5.

7. H. Takata, S. Suzuki,and E. Maeda, July 8–10, 1985, "Experimental Study of the Life Adjustment Factor for Reliability of Rolling Element Bearings," Proc. JSLE Int. Trib. Conf., pp. 603–608.

8. R. Sicard, Private communication, 2021

9. L. Johnson, L., 1964, 'The Statistical Treatment of Fatigue Experiments', Elsevier, New York.

10. L. Johnson, 1964, 'Theory and Technique of Variation Research', Elsevier, New York

11. S. Blachère, 2015, "A new Bias Correction Technique for Weibull Parametric Estimation", Quality Engineering Applications and Research, paper 2015- 0377

12. S. Blachère,, A. Gabelli, 2016, "Monte Carlo comparison of Weibull two and three parameters in the context of the statistical analysis of rolling bearings fatigue testing", ASTM.

13. L. Houpert, 2003, "Five methods for defining statistical confidence intervals", available under the permission of the Timken Company

14. L. Houpert, J. Clarke, 2022, ""A new four parameter reliability model applied to a first-in-N testing strategy using a large database of relative lives", submitted to the 2022 Bearing World Journal

A new four parameter reliability model applied to a first-in-N testing strategy using a large database of relative lives

L. Houpert[a]* and J. Clarke[b]

[a] Luc Houpert Consulting, 1 rue de Fleurie, 68920 Wettolsheim, France
[b] Smart Manufacturing Technology (SMT), Wilford House, 1 Clifton Lane, Nottingham, NG11 7AT, United Kingdom
*Corresponding author: luc.houpert@orange.fr

Abstract

A new reliability model is suggested in which the failure rate F is calculated as a function of the life L and four parameters (a_1 and L_{10}) and (a_2 and $L_{0.1}$) corresponding to two asymptotic linear models used at large and low F values respectively (with $a_2 \leq a_1$), with a smooth non-linear transition between these two straight lines. $L_{0.1}$ is the life corresponding to F=0.001 while L_{10} is the standard life corresponding to F = 0.1

An appropriate non-linear curve-fitting technique is suggested for retrieving the four parameters which are satisfactorily compared to the results obtained using two simple linear curve-fittings in the range $F > 0.05$ and $F < 0.01$.

The median value of F, as well as its 90 % variation range can be calculated exactly using the *inverse beta* function and the numbers N and NR corresponding to a first-in-N testing strategy and NR test rigs (or failed bearings); N=4 & NR = 6 for example. A standard two-parameter Weibull analysis of the 6 estimates of the $L_{15.91}$ lives can be done for estimating the $L_{15.91\_G}$ life of the group (of 6 failed bearings) as well as 6 values of the relative life $L/L_{15.91\_G}$ used later. Only failed or (failed + suspended) items can be considered in this exercise. It is demonstrated analytically that the same Weibull slope and life are obtained using both approaches, provided the L_{50} life of the failed-only items is used as best estimate of $L_{15.91}$.

Using the relative lives, a large database of 600 relative lives can be obtained by gathering 100 endurance tests (using N=4 and NR=6) or numerically simulated life (using random values of F sorted in ascending orders).

These 600 relative lives can be analyzed using the previously described non-linear or linear curve-fitting techniques, so that curve-fitted values of a_{2_cf} and $L_{0.1\_cf}/L_{10\_cf}$ can be obtained and compared to the exact values of a_2 and $L_{0.1}/L_{10}$. Using L_{10} as a reference, the value of $L_{15.91\_G}$ can also be estimated and used for estimating $L_{0.1\_cf}$ to compare to the exact value $L_{0.1}$ using again a non-linear and linear curve-fitting approach.

1000 Monte Carlo simulations of this exercise can be done for defining the median value and 90% confidence intervals of the ratio a_2/a_{2_cf} and $L_{0.1}/L_{0.1\_cf}$. Almost similar results are obtained using the non-linear or simple linear curve-fittings, so that two simple linear curve-fittings are finally suggested in the appropriate range of F ($F > 0.05$ and $F < 0.01$). These ratios are slightly biased but close to 1, confirming that relative lives can be used for retrieving the four parameters of our model. The biased ratios can be corrected by introducing a correction factor, curve fitted as a function of the ratio a_{2_cf}/a_{1_cf} and leading to an excellent estimate of a_2 and $L_{0.1}$ with reasonable confidence intervals.

Keywords: Reliability, Weibull models, analysis of large endurance database

Objectives

A detailed analysis of Rosemann's four parameter model [1] has been conducted in [2] by Houpert and Clarke.

In the latter, the cumulative failure probability F of the i^{th} failure is calculated exactly with all N tested bearings failing.

Rosemann's model is quite general and powerful, allowing the user to refute the existence of the third Weibull parameter L_0. Rosemann's model is however difficult to curve-fit and can exhibit some redundancies when the exponent c is small and close to 1.

It has however been observed in [2] that Rosemann's model behaves almost linearly at very low and very large values of F, so that a new model, easier to curve-fit and duplicating quite well Rosemann's trends, has been suggested in [2].

Defining the third and fourth parameters of these models (Rosemann or the new model suggested herein) requires however to use a very large number N of failed bearings in order to have access at low F results.

The latter problem can somewhat be avoided by using first in N failures strategy and relative lives.

It is indeed common to adopt a first-in-N testing strategy using NR test rigs, and the first objective of this paper is to show how the cumulative failure probability F of the i^{th} failure (out of NR) can be calculated exactly for defining its median value for example, but also lower and upper bound of its 90 % variation range. The median value can then be used for matching experimental life results to curve-fitted ones using any reliability life model, Rosemann's model for example or the new model suggested in [2].

The new model will be studied herein using two possible curve-fitting techniques for defining the four parameters of our model: a non-linear curve-fitting using a Newton-Raphson approach or a simplified approach using linear curve-fittings.

When using the first in *4* testing strategy (N=4) and 6 test rigs (NR=6), 6 estimates of the $L_{15.91}$ life can be used for defining (using a linear curve-fitting usually) the $L_{15.91\_G}$ of the group, so that 6 relative lives can be defined by dividing the 6 lives by $L_{15.91\_G}$. 100 endurance tests then lead to 600 points to analyze, the first 23 ones corresponding to very low failure rates.

Monte-Carlo simulations can be used for simulating the creation of a large database using relative lives, and the third objective is to demonstrate that the relative lives also follow the initial four parameters reliability model.

Finally, a second loop of Monte-Carlo simulations can be conducted for defining the confidence intervals assigned to the four unknown parameters.

Standard testing strategy using N failed bearings; two approaches for calculating F_{i_median}

A standard testing strategy consists of using N failed bearings and analyze $Y_i=ln(t_i)$ versus $X_i = ln(-ln(1-F_{i\_median}))$ using a 2 or 3 or 4 parameter reliability model where F_{i_median} is the cumulative median failure probability for the i^{th} bearing to fail before time t_i (sorted in ascending order).

The calculation of F_{i_median} has been described [2] using two approaches explained here below.

When generating N values of F (0< F <1) and sorting them in an ascending order, one can calculate the density f and cumulative distribution P_i of each i^{th} number F. The density distribution $f(F)$ corresponding to order i^{th} value of F is:

$$f(F) = \frac{N!}{(N-i)!.(i-1)!}.F^{i-1}.(1-F)^{N-i} \qquad (1)$$

The cumulative density P_i (probability that the i^{th} sorted random value is smaller or equal to F) is:

$$\begin{aligned} &P_i = \frac{N!}{(N-i)!.(i-1)!}.\int_0^F x^{i-1}.(1-x)^{N-i}.dx = A_i * I_i \\ &with \quad A_i = \frac{N!}{(N-i)!.(i-1)!} \quad , \quad I_i = \int_0^F x^{b_i}.(1-x)^{c_i}.dx \quad , \quad b_i = i-1 \quad \& \quad c_i = N-i \end{aligned} \qquad (2)$$

Two means of calculating F as a function of P have been developed in [2] when failing all N bearings: an analytical approach, somewhat tedious, and another one using the *incomplete beta* and *inverse beta* function, see next pages.

<u>Analytical approach:</u>

The first approach is using an analytical integration of P as a function of F, but one then needs to solve numerically $P(F)=P_{Targetted}$ for defining F as a function of $P_{Targetted}$.

$$\begin{aligned}
&in\ general:\\
&Starting\ with\ \ Coef_1 = 1\ \ \&\ \ P_1 = 1-(1-F)^N\\
&Coef_i = Coef_{i-1}.\frac{N-i+2}{i-1}\\
&P_i = P_{i-1} - Coef_i.F^{i-1}.(1-F)^{N-i+1} \qquad \left(also: P_i = P_{i-1} - \frac{N!}{(i-1)!.(N-i+1)!}.F^{i-1}.(1-F)^{N-i+1}\right)\\
&hence:\\
&P_1 = 1-(1-F)^N\\
&P_2 = 1-(1-F)^N - N.F.(1-F)^{N-1}\\
&P_3 = 1-(1-F)^N - N.F.(1-F)^{N-1} - \frac{N.(N-1)}{2}.F^2.(1-F)^{N-2}\\
&P_4 = 1-(1-F)^N - N.F.(1-F)^{N-1} - \frac{N.(N-1)}{2}.F^2.(1-F)^{N-2} - \frac{N.(N-1).(N-2)}{2*3}.F^3.(1-F)^{N-3}\\
&P_5 = 1-(1-F)^N - N.F.(1-F)^{N-1} - \frac{N.(N-1)}{2}.F^2.(1-F)^{N-2} - \frac{N.(N-1).(N-2)}{2*3}.F^3.(1-F)^{N-3}\\
&\qquad - \frac{N.(N-1).(N-2).(N-3)}{2*3*4}.F^4.(1-F)^{N-4}\\
&P_6 = 1-(1-F)^N - N.F.(1-F)^{N-1} - \frac{N.(N-1)}{2}.F^2.(1-F)^{N-2} - \frac{N.(N-1).(N-2)}{2*3}.F^3.(1-F)^{N-3}\\
&\qquad - \frac{N.(N-1).(N-2).(N-3)}{2*3*4}.F^4.(1-F)^{N-4} - \frac{N.(N-1).(N-2).(N-3).(N-4)}{2*3*4*5}.F^5.(1-F)^{N-5}\\
&P_7 = P_6 - \frac{N.(N-1).(N-2).(N-3).(N-4).(N-5)}{2*3*4*5*6}.F^6.(1-F)^{N-6}\\
&P_8 = P_7 - \frac{N.(N-1).(N-2).(N-3).(N-4).(N-5).(N-6)}{2*3*4*5*6*7}.F^7.(1-F)^{N-7}
\end{aligned} \tag{3}$$

$$\begin{aligned}
&P_N = F^N\\
&P_{N-1} = N.F^{N-1} - (N-1).F^N\\
&P_{N-2} = \frac{N.(N-1)}{2}.F^{N-2} - N.(N-2).F^{N-1} + \frac{(N-1).(N-2)}{2}.F^N\\
&P_{N-3} = \frac{N.(N-1).(N-2)}{2*3}.F^{N-3} - \frac{N.(N-1).(N-3)}{2}.F^{N-2} + \frac{N.(N-2).(N-3)}{2}.F^{N-1} - \frac{(N-1).(N-2).(N-3)}{2*3}.F^N
\end{aligned} \tag{4}$$

<u>Use of the *incomplete beta* function and *inverse beta* function:</u>

It has also been shown in [2] that P and F can be directly defined using the *incomplete beta* and *inverse beta* function.

$$\begin{aligned}
&P = \frac{N!}{(N-i)!.(i-1)!}.\int_0^F x^{i-1}.(1-x)^{N-i}.dx \quad = \quad Beta(F,i,N-i+1)\\
&F = InvBeta(P,i,N-i+1)
\end{aligned} \tag{5}$$

Median values of F can therefore be calculated exactly using one of the two previously described methods, but can also be approximated using Johnson's relationship or the one called *other* in our previous paper:

$$Johnson1: F_{P=0.5} \approx 1 - 2^{-\frac{1}{N}} + \frac{i-1}{N-1}.\left\{2^{\left(1-\frac{1}{N}\right)} - 1\right\} \quad (6)$$

$$other\ approx: F_{P=0.5} \approx \frac{i-0.305}{N+0.39}$$

The lower and upper bounds of F (corresponding to P = 0.05 and 0.95 respectively) can also be calculated (using Eqs. (3), (4) or (5)) for defining the 90% confidence interval of F.

First-in-*N* testing strategy using *NR* test rigs (*N*=4, *NR*=6 for example):

To reduce the testing time of an endurance test, a first-in-N (N= 4 for example) testing strategy is often used as shown by Houpert in [3].

This testing strategy consists of using NR test rigs (NR = 6 for example) having each N bearings (N=4 for example) under test, and to suspend the test on a given test rig when the first bearing (out of N) fails. NR lives representative of the $L_{15.91}$ bearing life are therefore available and are usually analyzed using a two-parameter Weibull model and a median value of F (called $F1$ in the next table showing also the 90% variation range of F) defined using NR values and i = 1 to NR.

$$F1 = InvBeta\left(P, i, NR - i + 1\right) \quad (7)$$

i	F1_0.05	median F1	F1_0.95	(using NR=6)
1	0.008512445	0.109101282	0.393037769	
2	0.062849892	0.264449983	0.581803409	
3	0.153161118	0.421407191	0.728661627	
4	0.271338373	0.578592809	0.846838882	
5	0.418196591	0.735550017	0.937150108	
6	0.606962231	0.890898718	0.991487555	

Table 1: Exact median, 5% lower and 95% upper bounds of $F1$ (using NR = 6)

The peculiar points, demonstrated next in this paper, is that the latter approach correctly defines the Weibull slope. Furthermore, the interpolated L_{50} value of the latter distribution corresponds to the true $L_{15.91}$ value.

Analysis of *P* and *F* using the first-in-*N* approach with *NR* test rigs

One will now call in this chapter $g(x)$ the density and $G(x)$ the cumulative probability of the failure rate x corresponding to the first failure out of N.

$$g(x) = N.(1-x)^{N-1} \qquad G(x) = \int_0^x g(x').dx' = -\left|(1-x')^N\right|_0^x = 1-(1-x)^N \quad (8)$$

It is possible to derive analytically the density f and the cumulative probability P of the NR first-in-four failure rates (sorted in ascending order)

$$f(x) = \frac{NR!}{(i-1)!.(NR-i)!}.(G(x))^{i-1}.(1-G(x))^{NR-i}.g(x) \quad (9)$$

The cumulative probability P is obtained by integrating analytically the latter relationship.

$$P = \frac{NR!}{(i-1)!.(NR-i)!}.N.\int_0^F (G(x))^{i-1}.(1-G(x))^{NR-i}.(1-x)^{N-1}\,dx \quad (10)$$

Hence:

$$P = \frac{NR!}{(i-1)!.(NR-i)!}.N.\int_0^F \left(1-(1-x)^N\right)^{i-1}.(1-x)^{N*(NR+1)-1-N*i}\,dx \qquad (NR\ first\ in\ N\ failures) \quad (11)$$

The latter relation can be compared to Eq. (5) corresponding to N failures.

As conducted in the previous chapter or in [2], P can be calculated analytically (for a given set of N and NR, N=4 and NR=6 for example) or numerically for any set of N and NR. Also, some approximated relationships are suggested in the appendix of [3] for defining the median value of F when considering suspended items.

Analytical calculations of P

Analytical calculations are conducted next by fixing $N = 4$ and $NR = 6$

$$
\begin{aligned}
&f_1(x) = 24 * (1-x)^3 \\
&P_1(x) = 1-(1-x)^{24} \\
&f_2(x) = 120 * \left[1-(1-x)^4\right] * (1-x)^{19} \\
&P_2(x) = 1+5*(1-x)^{24} - 6*(1-x)^{20} \\
&f_3(x) = 240 * \left[1-(1-x)^4\right]^2 * (1-x)^{15} \\
&P_3(x) = 1-10*(1-x)^{24} + 24*(1-x)^{20} - 15*(1-x)^{16} \\
&f_4(x) = 240 * \left[1-(1-x)^4\right]^3 * (1-x)^{11} \\
&P_4(x) = 1+10*(1-x)^{24} - 36*(1-x)^{20} + 45*(1-x)^{16} - 20*(1-x)^{12} \\
&f_5(x) = 120 * \left[1-(1-x)^4\right]^4 * (1-x)^{7} \\
&P_5(x) = 1-5*(1-x)^{24} + 24*(1-x)^{20} - 45*(1-x)^{16} + 40*(1-x)^{12} - 15*(1-x)^{8} \\
&f_6(x) = 24 * \left[1-(1-x)^4\right]^5 * (1-x)^{3} \\
&P_6(x) = 1+(1-x)^{24} - 6*(1-x)^{20} + 15*(1-x)^{16} - 20*(1-x)^{12} + 15*(1-x)^{8} - 6*(1-x)^{4}
\end{aligned}
\tag{12}
$$

Following are the Figures calculated using the previously analytical relationships:

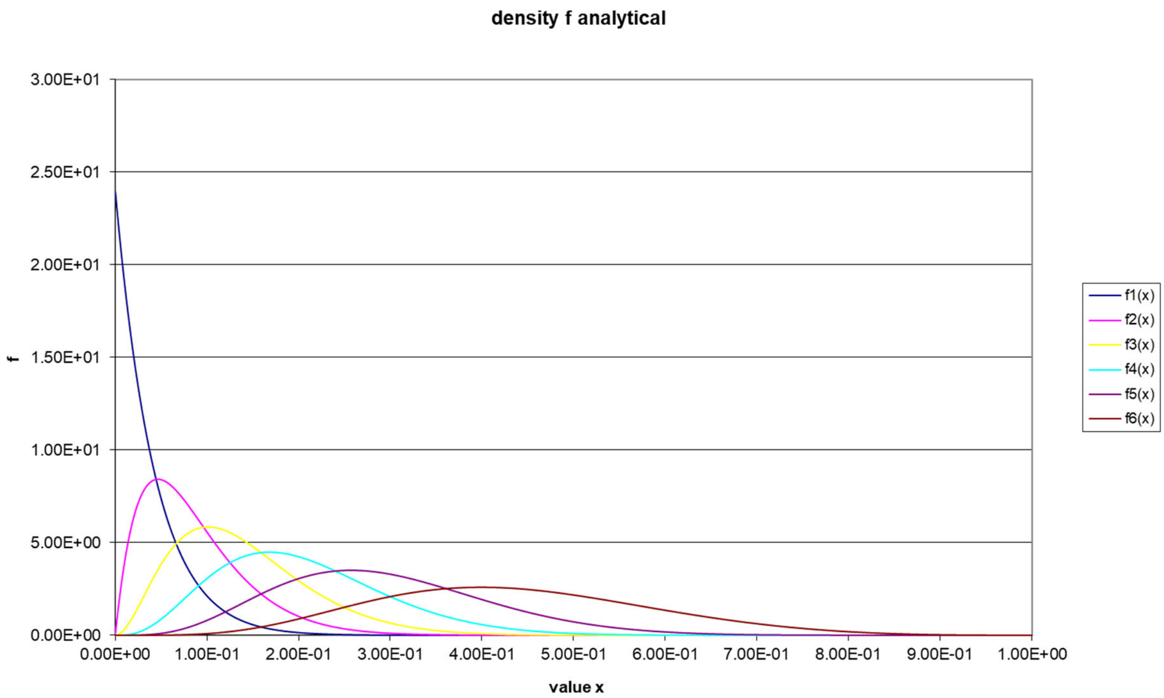


Fig. 1: Distribution of the density function f versus x ($x=F$)

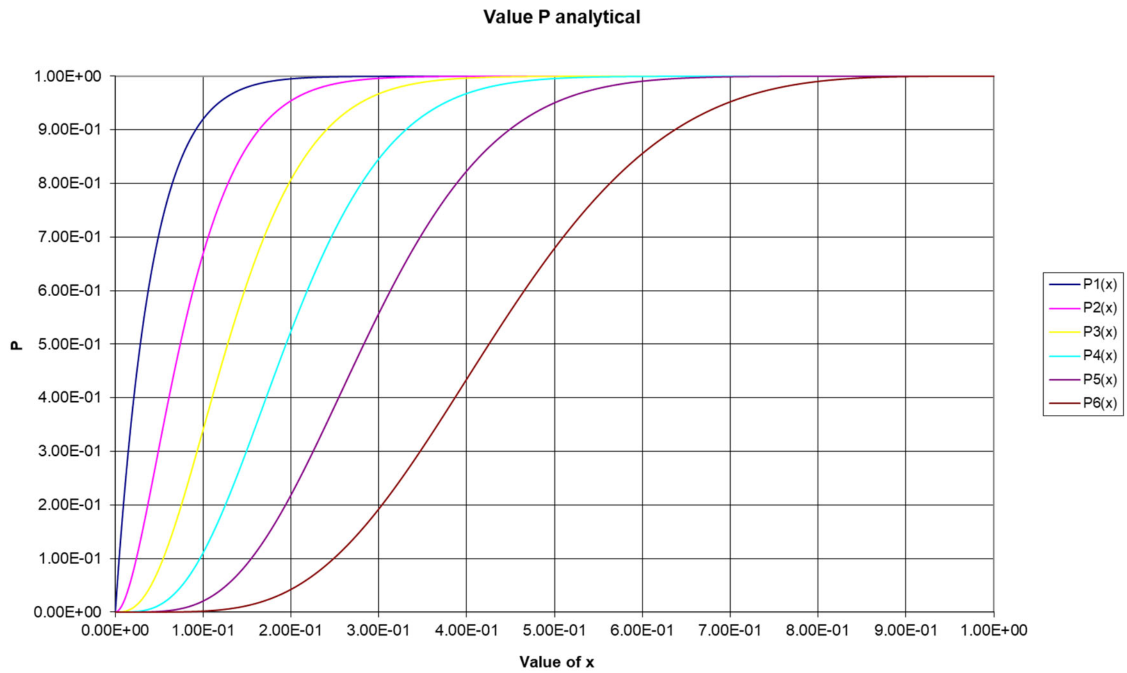


Fig. 2: Distribution of the cumulative probability *P* versus *x* (*x=F*)

Of interest to users is the reverse relationship *x=P(x)*

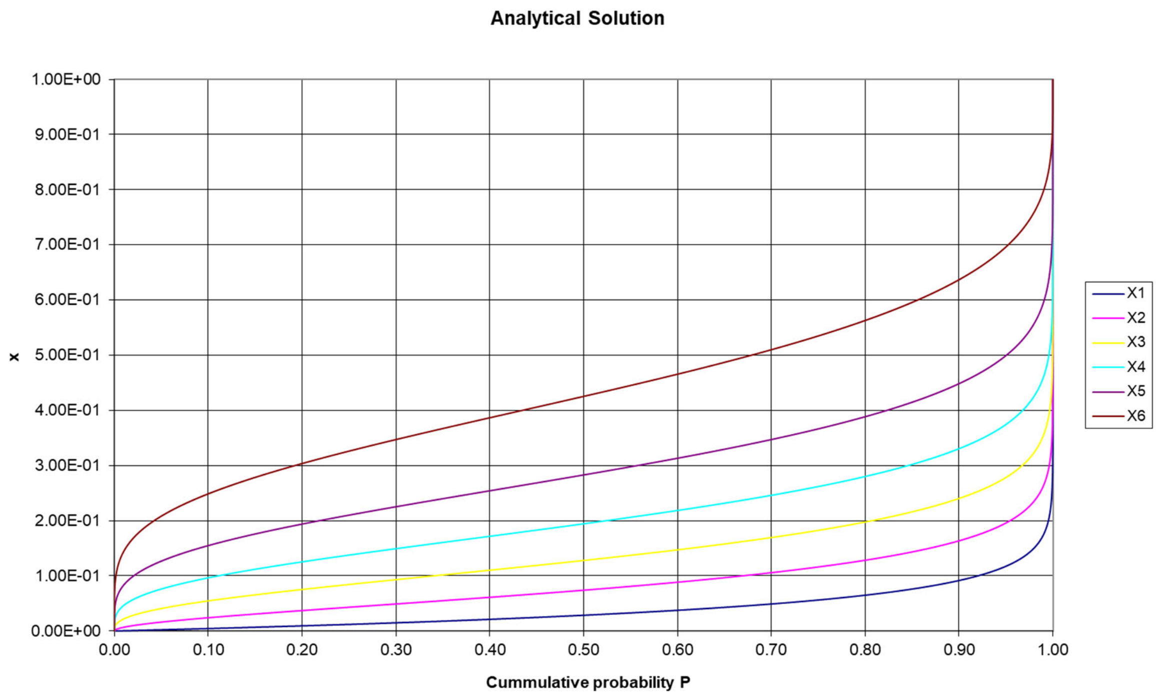


Fig. 3: Distribution of the cumulative failure probability *F* (*F=x*) versus *P*

Let's call *F2* the value of *F* defined using this second approach (which considers suspended items).
The value of *F2* corresponding to a given probability *P* can only be defined analytically for the first failure:

$$F2 = 1 - (1 - P)^{\frac{1}{24}} \qquad (13)$$

The other solution *F2* (when *i* is not equal to 1) requires using a solver for defining *F2* as a function of *P*.
Following is a table showing the results obtained using $P = 0.05$ (lower bound of the 90 % interval), 0.5 (median rank) and 0.95 (upper bound of the 90 % confidence interval).

				Approximations using Failed & Suspended increment orders and approximated F relationships				
i	F2_0.05	median F2	F2_0.95	i with increments	Median F2_Johnson	Median F2 others	rel error Johnson	rel error other
1	2.13493836970E-03	2.84680588464E-02	1.17346156155E-01	1.00000000	2.8468058846E-02	2.8688524590E-02	0.0000E+00	7.7443E-03
2	1.60969883221E-02	7.39103033498E-02	1.95835083696E-01	2.14285714	7.5328375980E-02	7.5526932084E-02	1.9186E-02	2.1873E-02
3	4.07093808822E-02	1.27845837259E-01	2.78264937927E-01	3.48739496	1.3045816084E-01	1.3063094090E-01	2.0433E-02	2.1785E-02
4	7.60861555648E-02	1.94296063572E-01	3.74413840309E-01	5.14221073	1.9831020375E-01	1.9845125944E-01	2.0660E-02	2.1386E-02
5	1.26638455257E-01	2.82889822892E-01	4.99301680895E-01	7.34863176	2.8877959429E-01	2.8887835084E-01	2.0820E-02	2.1169E-02
6	2.08212627164E-01	4.25278474630E-01	6.96251896687E-01	10.87890541	4.3353061916E-01	4.3356169706E-01	1.9404E-02	1.9477E-02

Table 2: Exact median, 5% lower and 95% upper bounds of *F2* (using N=4 and NR = 6)

The last five columns correspond to an approximated approach described in the appendix of Ref. [3] where Failed and Suspended items are considered (among $N*NR$ bearings, hence 24 bearings) for calculating a failure number increment and final failure order number i, not any longer equal to an integer. The latter value of i can then be used in Eq. (6) with 24 items for approximating the median value of *F2*. The relative error using the approximated solution is of the order of 0.02, but the latter approach can be used with any number of suspended items (and not systematically 1 failed and 3 suspended items).

Exact numerical calculations of *P* and *F*

Eq. (11) can be further studied by introducing a change of variable:

$$P = \frac{NR!}{(i-1)!.(NR-i)!}.N.\int_0^F \left(1-(1-x)^N\right)^{i-1}.(1-x)^{N*(NR+1)-1-N*i}\,dx$$

$$X = 1-(1-x)^N \qquad\qquad (1-x) = (1-X)^{\frac{1}{N}}$$

$$dX = N.(1-x)^{N-1}.dx \quad = \quad N.(1-X)^{\frac{N-1}{N}}.dx$$

$$dx = \frac{1}{N}.(1-X)^{-\frac{N-1}{N}}.dX$$

$$P = \frac{NR!}{(i-1)!.(NR-i)!}.\int_0^{1-(1-F)^N} X^{i-1}.(1-X)^{\frac{N*(NR+1)-1-N*i}{N}}.(1-X)^{-\frac{N-1}{N}}.dX$$

$$= \frac{NR!}{(i-1)!.(NR-i)!}.\int_0^{1-(1-F)^N} X^{i-1}.(1-X)^{(NR+1)-\frac{1}{N}-i}.(1-X)^{-1+\frac{1}{N}}.dX \qquad (14)$$

$$P = \frac{NR!}{(i-1)!.(NR-i)!}.\int_0^{1-(1-F)^N} X^{i-1}.(1-X)^{NR-i}.dX = Beta\left(1-(1-F)^N, i, NR-i+1\right)$$

The final 'exact' numerical solution reads therefore:

$$P = Beta\left(1-(1-F)^N, i, NR-i+1\right)$$

$$F = 1-\left[1-InvBeta\left(P_i, i, NR-i+1\right)\right]^{\frac{1}{N}} \quad or \qquad (15)$$

$$F2 = 1-\left[1-InvBeta\left(P_i, i, NR-i+1\right)\right]^{\frac{1}{N}}$$

where the symbols *F2* reminds the reader that the second approach (accounting for NR Failed and $NR*(N-1)$ suspended items) is used. The merits of the latter relationships are plural. No solver is required for defining *F2* as a function of *P*. Also, these relationships apply to any set (N, NR) while the previously defined analytical relationships giving *P* as a function of *F2* had to be developed analytically for a given set (N=4, NR=6).

One can now use a two-parameter Weibull model for retrieving the slope and any life (L_{10} or $L_{15.91}$ for example) using the latter median values of *F2*, with no need of interpolating L_{50} for having the best estimated of $L_{15.91}$ (as done when using *F1*).

Demonstration that the two approaches (using either *F1* or *F2*) are similar

When defining $Y = ln(t)$, $X1 = ln(-ln(1-F1))$ and $X2 = ln(-ln(1-F2))$, one can now demonstrate that the two approaches are similar. Using the previous relationships, one can write:

$$1-(1-F2)^{N} = InvBeta\left(P,i,NR-i+1\right) = F1$$

$$F2 = 1-\left[1-F1\right]^{\frac{1}{N}}$$

$$1-F2 = \left[1-F1\right]^{\frac{1}{N}} \quad (16)$$

$$\ln\left(1-F2\right) = \frac{1}{N}.\ln\left[1-F1\right]$$

$$\ln\left(-\ln\left(1-F2\right)\right) = \ln\left(\frac{1}{N}\right) + \ln\left(-\ln\left[1-F1\right]\right)$$

or:

$$X2 = \ln\left(\frac{1}{N}\right) + X1 = X1 + Translation$$

$$with\ \ X2 = \ln\left(-\ln\left(1-F2\right)\right) \quad X1 = \ln\left(-\ln\left(1-F1\right)\right) \ \ and\ \ Translation = \ln\left(\frac{1}{N}\right) \quad (17)$$

A constant translation between *X1* and *X2* is therefore observed (at any *P* value), explaining why the same slope is retrieved using the first or second approach, see next table and Figure.

i	median F1	median F2	X1	X2	Translation=X2-X1
1	0.109101282	2.84680588464E-02	-2.15827239	-3.544566751	-1.386294361
2	0.264449983	7.39103033498E-02	-1.180462231	-2.566756592	-1.386294361
3	0.421407191	1.27845837259E-01	-0.603020751	-1.989315112	-1.386294361
4	0.578592809	1.94296063572E-01	-0.146002302	-1.532296663	-1.386294361
5	0.735550017	2.82889822892E-01	0.285256492	-1.10103787	-1.386294361
6	0.890898718	4.25278474630E-01	0.795468469	-0.590825892	-1.386294361

Table 3: median *F1* and *F2* values confirming a constant translation

The following Figure also confirms the same slope but can be used for observing graphically that L_{50} using $F1$ = $L_{15.91}$ using $F2$.

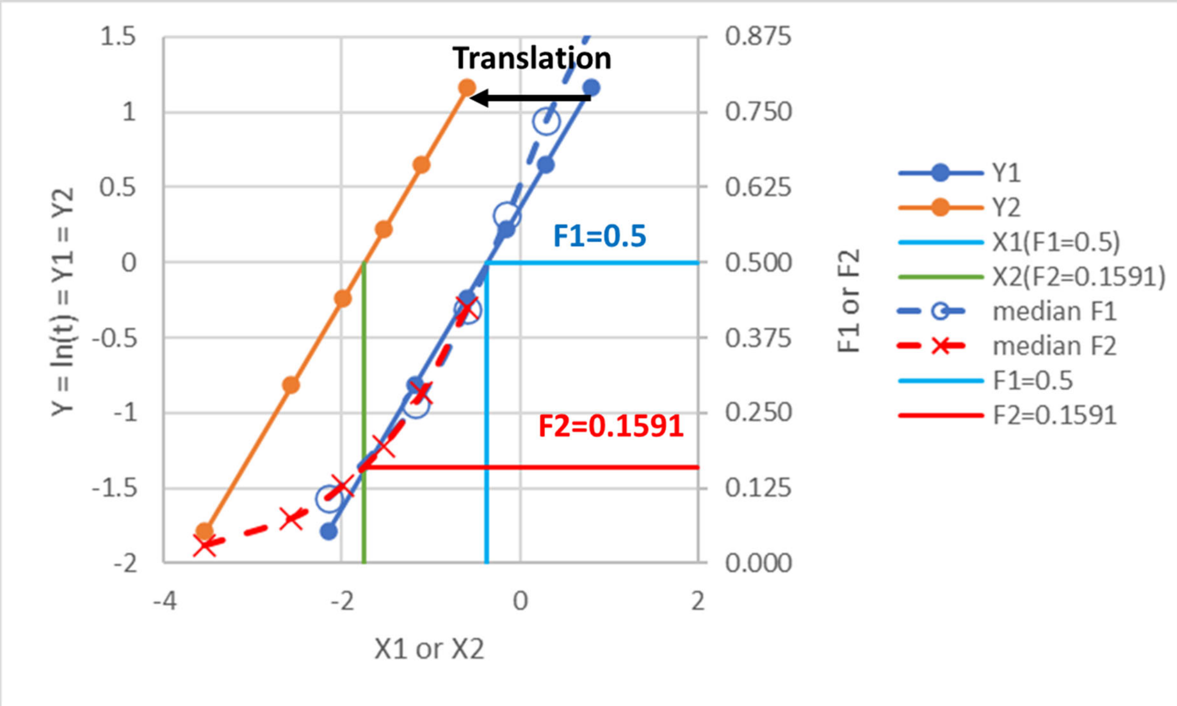


Fig. 4: Calculated results when plotting *Y* versus *X1* or *X2* using $L_{15.91}$=1 and β=1

The latter statement can also be demonstrated analytically.
Using *F2*, hence the exact relationship, one can write:

$$Y_2 = \ln(t) = \ln\left\{\frac{L_{15.91}}{\left[-\ln(1-0.1591)\right]^{\frac{1}{\beta}}}.\left[\left[-\ln(1-F2)\right]^{\frac{1}{\beta}}\right]\right\} = \ln\left\{\left[\frac{4}{\ln(2)}\right]^{\frac{1}{\beta}}.L_{15.91}\right\} + \frac{1}{\beta}.X2 \qquad (18)$$

When using *F1*, one already knows that the slope is the same, so that when using any of the 6 *Y* values, one can write:

$$Y_1 = b_1 + \frac{1}{\beta}.X1 = Y_2 = \ln\left\{\left[\frac{4}{\ln(2)}\right]^{\frac{1}{\beta}}.L_{15.91}\right\} + \frac{1}{\beta}.X2 \Rightarrow$$

$$b_1 = \ln\left\{\left[\frac{4}{\ln(2)}\right]^{\frac{1}{\beta}}.L_{15.91}\right\} + \frac{1}{\beta}.(X2 - X1) = \ln\left\{\left[\frac{4}{\ln(2)}\right]^{\frac{1}{\beta}}.L_{15.91}\right\} - \frac{1}{\beta}.Translation \qquad (19)$$

$$Y_1 = \ln\left\{\left[\frac{4}{\ln(2)}\right]^{\frac{1}{\beta}}.L_{15.91}\right\} - \frac{1}{\beta}.Translation + \frac{1}{\beta}.X1 = \ln\left\{\left[\frac{4}{\ln(2)}\right]^{\frac{1}{\beta}}.L_{15.91}\right\} - \frac{1}{\beta}.\ln(4) + \frac{1}{\beta}.\ln(-\ln(1-F1))$$

$$= \ln\left\{\left[\frac{4}{\ln(2)}\right]^{\frac{1}{\beta}}.L_{15.91}\right\} + \ln\left[\frac{\ln(1-F1)}{4}\right]^{\frac{1}{\beta}} = \ln\left\{\left[\frac{4}{\ln(2)}.\frac{\ln(1-F1)}{4}\right]^{\frac{1}{\beta}}\right\} + \ln(L_{15.91})$$

$$When\ F1 = 0.5,\quad Y1 = \ln(L_{15.91})\quad \&\quad t = L_{15.91} \qquad (20)$$

The next Figure shows an example of calculated *Y2* and *t* values plotted versus *X2* with the corresponding 90% variation range.

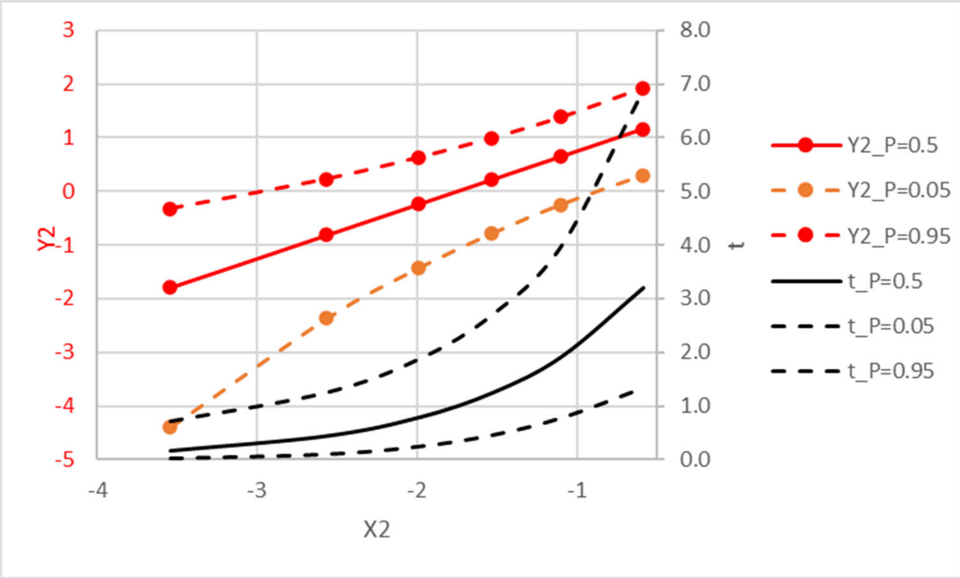


Fig. 5: Calculated median, lower & upper bounds of $Y2$ and t using $L_{15.91}$=1 and β=1

$X2$ can now be used with any reliability model using for example a three or four parameter reliability model for example.

In the following, one will use the true F values, hence the previously defined $F2$ value, with a first-in-N testing strategy, N being fixed to 4.

New four parameters reliability model and curve-fitting of the four parameters:

The new model described in [2] contains four unknowns to define: a_1, b_1, a_2 and b_2 representative of a two parameter Weibull model in the large and low range of F respectively:

$$a_1 = \frac{1}{\beta_1} \qquad b_1 = \ln(\eta_1) \qquad a_2 = \frac{1}{\beta_2} \qquad b_2 = \ln(\eta_2) \tag{21}$$

$$Y = \ln(t) = Y_1 + \frac{Y_2 - Y_1}{1 + \left(\frac{F}{F_{\text{intersection}}}\right)^2} \tag{22}$$

with:

$$Y_1 = b_1 + a_1 * X \qquad with \quad X = \ln\left(-\ln\left(1-F\right)\right)$$
$$Y_2 = b_2 + a_2 * X$$
$$F_{\text{intersection}} = 1 - \exp\left[-\exp(X_{\text{intersection}})\right] \qquad X_{\text{intersection}} = \frac{b_2 - b_1}{a_1 - a_2} \tag{23}$$

Note that the symbols Y_1 and Y_2 used in this chapter differ from the ones used in the previous chapter. In the previous chapter, $Y1$ and $Y2$ were associated to the use of $X1$ and $X2$ ($X2$=$X1$+$Translation$), while they now represent the two asymptotic lines calculated with the true value of X (corresponding to $X2$ defined with $F2$, hence including the translation or suspended items).

Values of a_2 and b_2 have been correlated in [2] to Rosemann's exponent c (varying from 2 to 175 as show in the next table) while maintaining the other three Rosemann's parameters constant: L_0 = 0.2, L_{10}=1 and β=1.

The value $L_{0.1}$ used next corresponds to the life when F = 0.001 while L_{10} is the life corresponding to F = 0.1

c	b2	a2	L0.1	b1	a1	L10
2	0.7834	0.5306	0.0560	2.0873	0.9447	1
4	-0.0229	0.297	0.1256	2.0655	0.9377	1
10	-0.5716	0.1636	0.1824	2.0650	0.9375	1
175	-0.8241	0.1109	0.2039	2.0650	0.9375	1

Table 4: From [2]: correlation between c, a_2 and b_2 when L_{10}=1, a_1=1 and L_0 = 0.2

In the following, one will test our new model using a_1, b_1, a_2 and b_2, but may also use as input: a_1=1, L_{10}=1, a_2 and $L_{0.1}$ as defined in the previous table and refer sometimes (for simplicity) to c=2, 4, 10 and 175 (easier to refer to rather than referring to a_2 and b_2).

L_{10} and $L_{0.1}$ can indeed be used as input for defining b_1 and b_2 via the following linear approximation:

$$b_1 \approx ln\, L_{10} - a_1 . ln(-ln(0.9))$$

$$b_2 \approx ln\, L_{0.1} - a_2 . ln(-ln(0.999)) \qquad (24)$$

hence b_1 = 2.2504 and b_2 = 0.7826 when a_1=1, L_{10}=1, a_2=0.5306 and $L_{0.1}$= 0.056 for example.

L_{10} and $L_{0.1}$ can be also used as input for defining, via some iterations, the exact values of b_1 and b_2 using our full four parameter non-linear model: b_1 = 2.3096 and b_2 = 0.7838 when a_1=1, L_{10}=1, a_2=0.5306 and $L_{0.1}$= 0.056 for example.

Note that any other life can be used as reference for defining b_1, for example $L_{15.91}$ = 1 with a_1=1, leading to b_1 = 1.7528 (using F=0.1591 and the linear model).

Following is one example of results obtained with our new model and 1000 points generated using either F_{median} or random values of F sorted in ascending order. The corresponding Y values ($Y=ln(t)$) are then plotted versus the median values of $X_{median}=ln(-ln(F_{median}))$. The 90% variation range of Y is also shown.

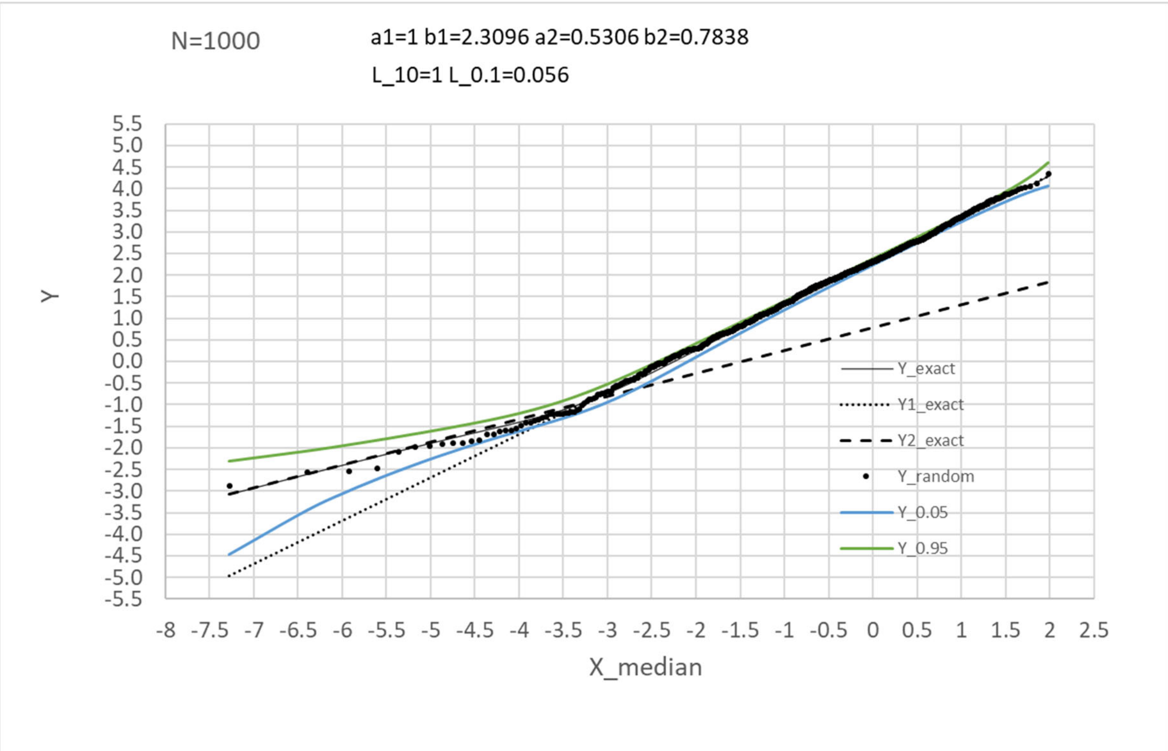


Fig. 6: Example of results obtained using the new 4 parameter non-linear model, N=1000

The next step consists of curve-fitting this database for retrieving the four parameters of our model: a_1, b_1, a_2 and b_2 or a_1, L_{10}, a_2 and $L_{0.1}$.

Non-linear curve-fitting

The full model is non-linear and requires a challenging non-linear curve-fitting for defining the four unknowns a_1, b_1, a_2 and b_2:

$$Y = b_1 + a_1.X + \frac{(b_2 - b_1) + (a_2 - a_1).X}{1 + \left(\frac{1 - \exp\left[-\exp(X)\right]}{1 - \exp\left[-\exp(\frac{b_2 - b_1}{a_1 - a_2})\right]} \right)^2} \quad (25)$$

An iterative non-linear curve-fitting approach is suggested in Appendix 1 with the previous example curve-fitted in the following Figure 7.

The approach called Method 1 in appendix 1 minimizes the sum of the square of the vertical differences, $\left(Y_{\exp\_i} - Y_{cf\_i}\right)^2$

Other approaches are available, like the one called Method 2 in Ref. [2], consisting of minimizing the sum of the square of the horizontal differences. This approach has been fully tested in Ref. [2].
Mike Kotzalas used in Ref. [4] the Hazard method to determine the median rank and then used method 2.
Another well-known approach is the Maximum Likelihood Estimate (MLE) approach, consisting of maximizing the product of all density distributions f_i (associated to each of the failed L_i) times the product of all S_i survival probabilities (associated to suspended lives L_i). This approach is very attractive since there is therefore no need to define median ranks while suspended items can be easily considered.

While using a standard 2-parameter Weibull model, five approaches (including the MLE one) have been tested in ref. (5], unfortunately only available upon request. It has been demonstrated that the results obtained using the MLE can be quite biased when using low N values, but also that all five approaches lead to similar confidence intervals once correcting for the biased median ratio.

When using a standard 2-parameter Weibull model, Methods 1 or 2 require using simple linear curve-fitting while a non-linear curve-fitting is required when using MLE, explaining perhaps why users often favour the use of Method 1 or 2.

When using the here-in described new 4-parameter Weibull model, non-linear curve-fitting cannot be avoided, even when using method 1.

Since all approaches are probably equivalent once correcting for the biased ratio, only method 1 is fully described in appendix 1 and tested in this paper.

Last, bearing users interested in advanced information on reliability may read Ref. [6] to [9].

Simplified linear curve-fittings

The simplest curve-fitting requires us to conduct two linear curve-fittings for defining a_1, b_1, a_2 and b_2 by fixing the lower bound F_{1T} of range 1 ($F > F_{1T}$) and upper bound F_{2T} of range 2 ($F < F_{2T}$), for example F_{1T}=0.05 and F_{2T} = 0.01, and rely on the smooth transition between Y_1 and Y_2 using $F_{intersection}$ and n=2.

$$\begin{array}{ll} \textit{Linear curve-fitting:} & \\ & \\ Y_1 = b_1 + a_1 * X & \textit{when } F \geq F_{1T} = 0.05 \;\; \textit{or } X \geq X_{1T} = \ln(-\ln(1 - F_{1T})) \\ Y_2 = b_2 + a_2 * X & \textit{when } F \leq F_{2T} = 0.01 \;\; \textit{or } X \leq X_{2T} = \ln(-\ln(1 - F_{2T})) \end{array} \quad (26)$$

The following Figure 7 shows the results obtained with the previous example curve-fitted using the non-linear and linear curve-fitting. Minor differences are observed in this example, the sum Si2 being only slightly reduced when using the non-linear curve-fitting.

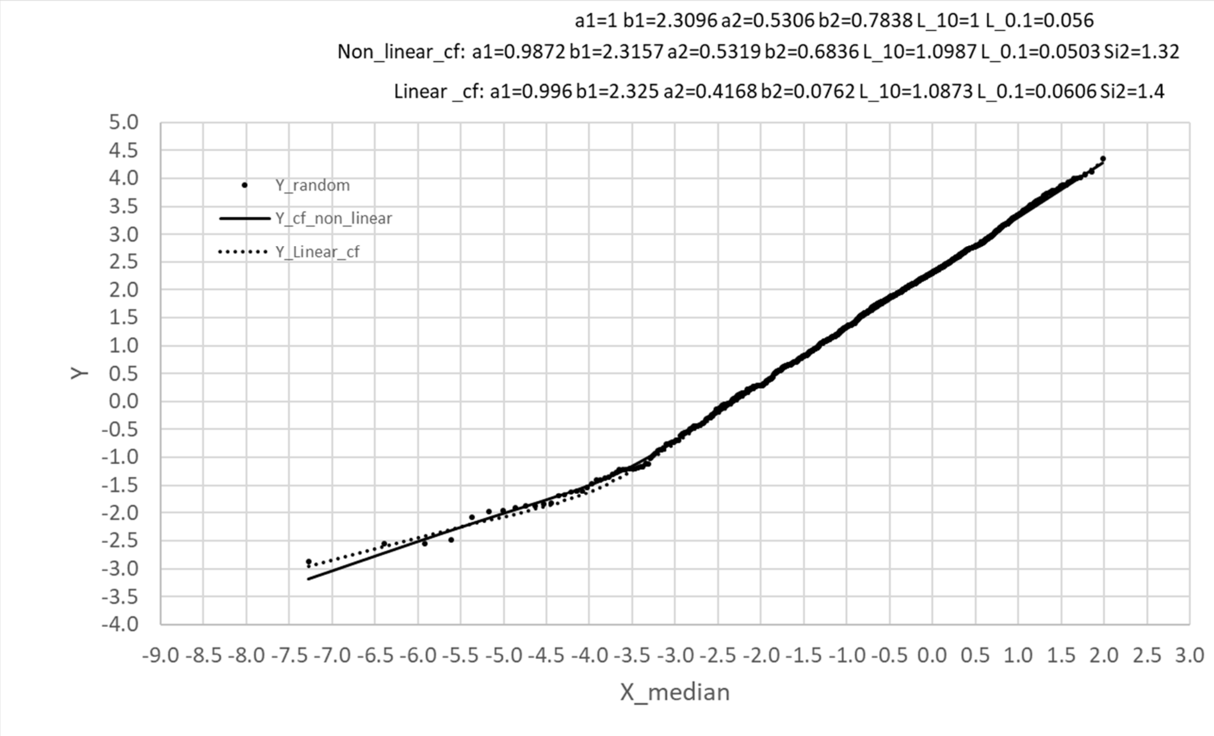


Fig. 7: Example of curve-fitted results obtained with N=1000.

The objectives being to describe life reliability models at low F values, one could therefore simply suggest (and test later for confirming our suggestion) a linear curve-fitting between Y and X at low F values, hence use $Y_2=a_2.X+b_2$ when $F < F_{2T} = 0.01$.

Note that even when using N=1000 (hence many bearing failures), only 10 points corresponding to $F < 0.01$ are available, the lowest value of F_{median} being then equal to $6.9291*10^{-4}$.

The main problem when using a three or four parameter reliability model remains therefore to obtain a large database for having results to curve-fit at low F values.

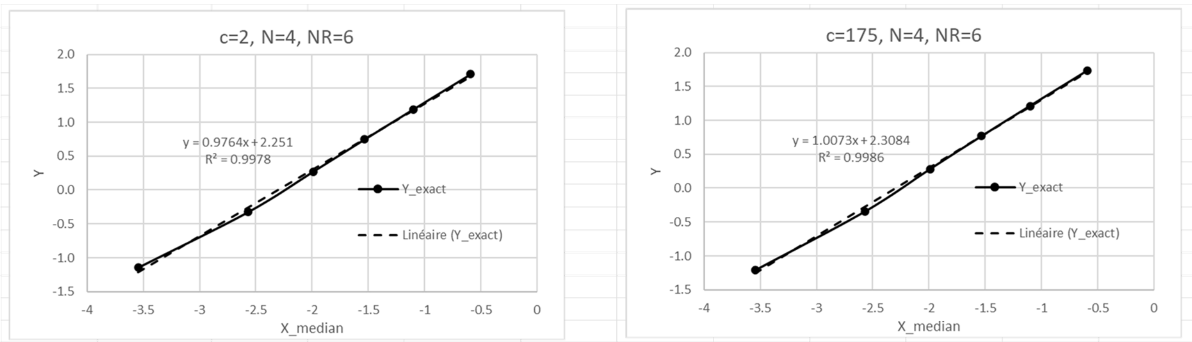
One will demonstrate next that all parameters (a_2, b_2 especially, but also a_1, b_1) can be retrieved using relative lives defined for example using 100 times 6 failures corresponding to first-in-4 failures.

Generation of a large database

For creating a large endurance test database, one idea used by M. Kotzalas [4] (using a three-parameter Weibull model) consists of analyzing relative lives L_{rel_i}, each bearing life L_i being divided by the estimated $L_{15.91\_G}$ of the tested group when using the first-in-4 (N=4) bearing failure and 6 (NR = 6) failures for example.

$$L_{rel\_i} = \frac{L_i}{L_{15.91\_G}} \quad (27)$$

It can be shown that when using N=4 and NR=6, the value of $L_{15.91\_G}$ can be estimated using a linear regression, the error on $L_{15.91\_G}$ being of the order of 2% when c=2, see next Figure, and surprisingly even less when c= 175.



c=2			c=4			c=10			c=175		
Linear regression		Exact values	Linear regression		Exact values	Linear regression		Exact values	Linear regression		Exact values
a1	0.976362267	1	a1	0.980263423	1	a1	0.99596857	1	a1	1.00729784	1
b1	2.25095972	2.30963377	b1	2.269333772	2.33262093	b1	2.29393329	2.33803497	b1	2.30793618	2.33577372
a2		0.5306	a2		0.297	a2		0.1636	a2		0.1109
b2		0.78377331	b2		-0.02110275	b2		-0.56300707	b2		-0.8203938
L_0.001		0.056	L_0.001		0.1256	L_0.001		0.1834	L_0.001		0.2039
L10	1.055258727	1	L_10	1.065432683	1	L_10	1.05404817	1	L10	1.04200425	1
L15.91	1.715294575	1.68012775	L_15.91	1.73519692	1.69780755	L_15.91	1.73012268	1.70535092	L15.91	1.72002218	1.70556851
L_10 Linear/L_10_exact		1.055	Ratio L_10 Linear/exact		1.065	Ratio L_10 Linear/exact		1.054	L_10 Linear/L_10_exact		1.042
L_15.91 Linear/L_15.91_exact		1.021	Ratio L_15.91 Linear/exact		1.022	Ratio L_15.91 Linear/exact		1.015	L_15.91 Linear/L_15.91_exa		1.008

Fig. 8: Examples of linear regression conducted on 6 first-in-4 'exact' failures.

Using now several endurance databases (for example 100 of them, each leading to 6 relative lives), 600 relative lives are available for being sorted in ascending order.

The lowest median F value (of the first point) is then quite small: $2.8877*10^{-4}$ (with 23 points corresponding to $F < 0.01$) versus $6.9291*10^{-4}$ when using 1000 first-in-1 failures (and only 10 points corresponding to $F < 0.01$).

Appendix 2 shows that the relative life corresponding to the two asymptotic linear curves can be defined using the same slopes a_1 and a_2 and two relative value b_{1_rel} and b_{2_rel}:

$$Y1_{rel} = a_1.(X - X_{0.1591}) = a_1.X + b_{1\_rel}$$
$$with \quad X = \ln(-\ln(1-F)) \quad \& \quad b_{1\_rel} = -a_1.X_{0.1591} \qquad (28)$$

$$Y2_{rel} = a_2.X + b_{2\_rel}$$
$$b_{2\_rel} = b_2 - a_1.(X_{0.1591} - X_{ref}) - \ln(L_{ref}) = b_2 - a_1.(X_{0.1591} - X_{0.1}) - \ln(L_{10}) \qquad (29)$$

The full non-linear model reads of course:

$$Y_{rel} = Y1_{rel} + \frac{Y2_{rel} - Y1_{rel}}{1+\left(\frac{F}{F_{intersection}}\right)^2} = b_{1\_rel} + a_1.X + \frac{(b_{2\_rel} - b_{1\_rel}) + (a_2 - a_1).X}{1+\left(\frac{1-\exp[-\exp(X)]}{1-\exp\left[-\exp(\frac{b_{2\_rel}-b_{1\_rel}}{a_1 - a_2})\right]}\right)^2} \qquad (30)$$

Curve-fitting the relative lives

Using one example, the full non-linear (as explained in Appendix 1) and the two simple linear curve-fittings are conducted next for defining the curve-fitted values of our four unknowns, now using the relative lives.

$$Y_{rel\_cf} = f(X, a_{1\_cf}, b_{1\_rel\_cf}, a_{2\_cf}, b_{2\_rel\_cf}, n=2) \quad to\ compare\ to\ Y_{rel} \qquad (31)$$

$$Y1_{rel\_cf} = a_{1\_cf}.X + b_{1\_rel\_cf} \quad to\ compare\ to\ Y1_{rel} = a_1.X + b_{1\_rel} \quad with \quad b_{1\_rel} = -a_1.X_{0.1591} \qquad (32)$$

$$Y2_{rel\_cf} = a_{2\_cf}.X + b_{2\_rel\_cf} \quad to\ compare\ to\ Y2_{rel} = a_2.X + b_{2\_rel}$$
$$with \quad b_{2\_rel} = b_2 - a_1.(X_{0.1591} - X_{0.1}) - \ln(L_{10}) \qquad (33)$$

The four non-linear curve-fitted values (a_{1_cf}, b_{1_cf}, a_{2_cf} and b_{2_cf}) used in Eq. (31) differ of course slightly from the linear curve-fitted ones used in Eq. (32) and (33) as shown for the example in Fig. 9, but the same symbols (a_{1_cf}, b_{1_cf}, a_{2_cf} and b_{2_cf}) will be kept for simplicity.

In the following, one will follow the ratios a_1/a_{1_cf} and a_2/a_{2_cf}, as well as some other ratios of interests to users.

Calculation of the derived exact and curve-fitted ratio: $L_{0.1}/L_{10}$

Eq.(31) can be used three times (with $X_{0.001}$, $X_{0.1}$ and $X_{0.1591}$) for defining the ratios:

$$\text{Non-linear curve-fitted ratio:}\quad \frac{L_{0.1\_cf}}{L_{10\_cf}} = \exp\left(Y_{0.1\_rel\_cf} - Y_{10\_rel\_cf}\right) \;\&\; \frac{L_{0.1\_cf}}{L_{15.91\_cf}} = \exp\left(Y_{0.1\_rel\_cf} - Y_{15.91\_rel\_cf}\right) \tag{34}$$

while the linear curve-fittings lead to:

$$\text{Linear curve-fitting ratio:}\quad \frac{L_{0.1\_cf}}{L_{10\_cf}} = \exp\left(Y2_{0.1\_rel\_cf} - Y1_{10\_rel\_cf}\right) \;\&\; \frac{L_{0.1\_cf}}{L_{15.91\_cf}} = \exp\left(Y2_{0.1\_rel\_cf} - Y1_{15.91\_rel\_cf}\right) \tag{35}$$

When using the linear models, simple analytical relationships can be further developed for defining the linear curve-fitted ratio to compare to the exact ratio:

$$\begin{array}{ll} \text{Exact ratio} & \text{Linear curve-fitted ratio:} \\ \dfrac{L_{0.1}}{L_{10}} = \exp\left(Y_{0.1\_rel} - Y_{10\_rel}\right) & \dfrac{L_{0.1\_cf}}{L_{10\_cf}} = \exp\left[a_{2\_cf}.X_{0.001} + b_{2\_rel\_cf} - a_{1\_cf}.X_{0.1} - b_{1\_rel\_cf}\right] \\ \dfrac{L_{0.1}}{L_{15.91}} = \exp\left(Y_{0.1\_rel} - Y_{15.91\_rel}\right) & \dfrac{L_{0.1\_cf}}{L_{15.91\_cf}} = \exp\left[a_{2\_cf}.X_{0.001} + b_{2\_rel\_cf} - a_{1\_cf}.X_{0.1591} - b_{1\_rel\_cf}\right] \end{array} \tag{36}$$

These ratios are of major interest to users. One will compare later (using Monte Carlo simulations) the exact ratios to the ones obtained using non-linear and linear curve-fittings. It is hoped that the median ratios a_2/a_{2_cf} and $\dfrac{L_{0.1}}{L_{10}} \Big/ \dfrac{L_{0.1\_cf}}{L_{10\_cf}}$ will be close to 1 and that the 90 % confidence interval of the latter ratio will not be too large for demonstrating that relative lives can indeed be used for access to a_2 and $L_{0.1}/L_{10}$.

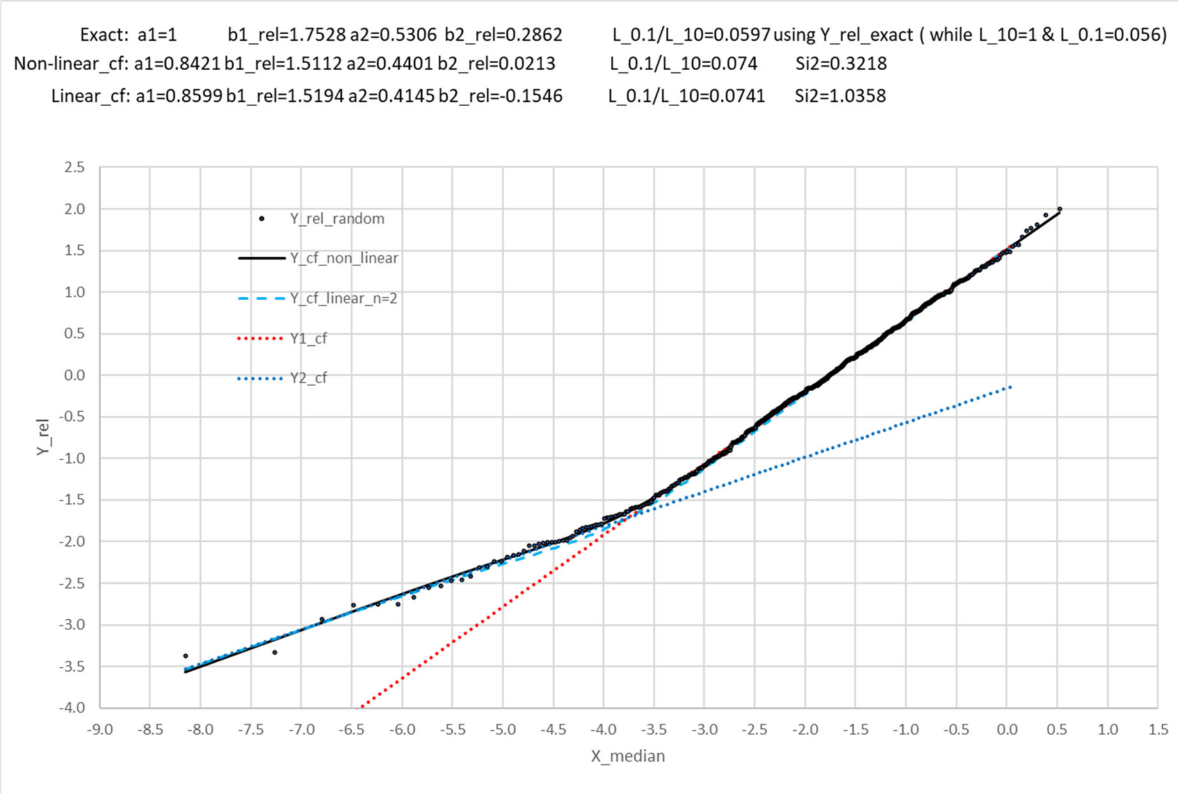


Fig. 9: Example of results obtained using one large database of 600 relative lives.

In this example, the ratio $L_{0.1\_cf}/L_{10\_cf}$ is of the order of 0.074 instead of 0.056 and a_{2_cf} are equal to about 0.44 or 0.41 instead of 0.5306

Note that when using the exact 4 parameters (a_{1_cf}, $b_{1\_rel\_cf}$, a_{2_cf} *and* $b_{2\_rel\_cf}$), $L_{0.1}/L_{10}$ is equal to 0.0597 instead of 0.056 while the relative lives have been created using $L_{0.1}$=0.056 & L_{10} =1. The differences are attributed to the use of relative lives.

Estimation of L_{10} and $L_{0.1}$:

Although relative lives are used in our large database, one can still try to retrieve or estimate the single values of L_{10_cf} and $L_{0.1\_cf}$.

For estimating L_{10_cf} and $L_{0.1\_cf}$, one will use the curve-fitted results representative of $ln(L_{10\_cf}/L_{15.41\_G})$ and $ln(L_{0.1\_cf}/L_{15.41\_G})$, but needs however to estimate the value of $L_{15.91\_G}$ used as the denominator of the latter ratio. This value has been estimated 100 times when using 600 relative lives in our example, but a single value must be estimated and kept now for the next steps.

For estimating $L_{15.91\_G}$, one must assume knowing the origin of our experimental database (describing $L/L_{15.91\_G}$), hence the median value for example of $L_{15.91\_G}$.

In our numerical simulation, the large database has been created using L_{10} as initial reference and a slope a_1 (that can be assumed correctly estimated when using a two-parameter Weibull model), but also a_2 and b_2. One can assume that the median value of $L_{15.91\_G}$ is equal to an extrapolated 'exact' or curve-fitted value $L_{15.91\_G\_cf}$ defined with the exact L_{10} value (taken as reference) and the slope a_1 or a_{1_cf} respectively. Furthermore, these 'exact' values of $L_{15.91\_G}$ and curve-fitted value $L_{15.91\_G\_cf}$ will be estimated next using our simplified linear relationship giving *Y1* (with either the exact a_1 or a_{1_cf} slope) since the 100 values of $L_{15.91\_G}$ have been defined using a linear curve-fitting, hence *Y1*.

The quotes around the word 'exact' are used since a simplified linear relationship (*Y1*) has been used for defining $L_{15.91\_G}$ while the real behavior is non-linear, but unknown to the user. In other words, the true exact values of $L_{15.91}$, but also L_{10} and $L_{0.1}$ are never known since the four parameters used in our model (a_2 and b_2 especially) are not known.

Note also that the distinction between a_1 and a_{1_cf} would not be needed *if* $L_{15.91}$ (instead of L_{10}) would have been taken as initial reference, but it is common practice to use L_{10} as reference. The 'exact' and curve-fitted values of $L_{15.91\_G}$ can finally be estimated using:

$$L_{15.91\_G} \approx L_{10} \cdot \left(\frac{\ln(1-0.1591)}{\ln(1-0.1)} \right)^{a_1} \approx L_{15.91} \quad \& \quad L_{15.91\_G\_cf} \approx L_{10} \cdot \left(\frac{\ln(1-0.1591)}{\ln(1-0.1)} \right)^{a_{1\_cf}} \qquad (37)$$

Note also that a_{1_cf} can be defined as corresponding to the non-linear or linear curve-fitting.
When using the 'exact' value of $L_{15.91\_G}$ with the non-linear Y_{cf} or linear relationships Y_{1cf} and Y_{2cf}, one obtains:

$$L_{15.91\_G} = L_{10} \cdot \left(\frac{\ln(1-0.1591)}{\ln(1-0.1)} \right)^{a_1}$$

$$\begin{aligned} &Non-linear: \\ &L_{0.1\_cf} = \exp\left[Y_{0.1\_cf}\right].L_{15.91\_G} \\ &L_{10\_cf} = \exp\left[Y_{10\_cf}\right].L_{15.91\_G} \\ &L_{15.91\_cf} = \exp\left[Y_{15.91\_cf}\right].L_{15.91\_G} \end{aligned} \qquad (38)$$

$$\begin{aligned} &Linear: \\ &L_{0.1\_cf} = \exp\left[Y2_{0.1\_cf}\right].L_{15.91\_G} = \exp\left[a_{2\_cf}.\ln(-\ln(1-0.001)) + b_{2\_rel\_cf}\right].L_{15.91\_G} \\ &L_{10\_cf} = \exp\left[Y1_{10\_cf}\right].L_{15.91\_G} = \exp\left[a_{1\_cf}.\ln(-\ln(1-0.1)) + b_{1\_rel\_cf}\right].L_{15.91\_G} \\ &L_{15.91\_cf} = \exp\left[Y1_{15.91\_cf}\right].L_{15.91\_G} = \exp\left[a_{1\_cf}.\ln(-\ln(1-0.1591)) + b_{1\_rel\_cf}\right].L_{15.91\_G} \end{aligned}$$

Similar relationships are of course suggested when using $L_{15.91\_G\_cf}$ using a_{1_cf} defined with either the non-linear or linear curve-fitting.
The next table shows the results obtained in our last example corresponding to Fig. 9:

L_15.91 exact	1.6801			
L_15.91_G	1.6447	(linearly extrapolated using L10 & a1)		
L_15.91_G_cf 1	1.5204	(linearly extrapolated using L10 & non-linear a1_cf)		
L_15.91_G_cf2	1.5340	(linearly extrapolated using L10 & linear a1_cf)		
	Non-linear _L0.1_cf/L_10_cf =0.074		Linear L_0.1_cf/L_10_cf =0.0741	
	L_0.1_cf	L_10_cf	L_0.1_cf	L_10_cf
Using L_15.91_G	0.0802	1.0846	0.0805	1.0854
Using L_15.91_G_cf 1	0.0742	1.0026	0.0744	1.0033
Using L_15.91_G_cf 2	0.0748	1.0115	0.0750	1.0123

Table 5: Example of estimates of $L_{0.1\_cf}$ and L_{10_cf} using the relative lives, L_{10}=1, $L_{0.1}$=0.056

The true and exact value of $L_{15.91}$ is here equal to 1.68012 but is usually not known by the user conducting linear regression on 6 first-in-four failures. Having defined the four unknows, one can follow some ratios of interest to users, see appendix 2.

Besides following the ratio a_1/a_{1_cf} (equal to β_{1_cf}/β_1) and a_2/a_{2_cf} (equal to β_{2_cf}/β_2), one can follow the following ratios when using $L_{15.91\_G}$, hence using a_1 in Eq. (38):

$$\frac{L_{10}}{L_{10\_cf}} = \exp\left[Y_{10} - Y_{10\_cf}\right] \text{ (non-linear model)}$$

$$= \exp\left[\left(a_1 - a_{1\_cf}\right).X_{0.1} + b_{1\_rel} - b_{1\_rel\_cf}\right] \text{ (linear model)}$$

$$\left(\frac{L_{10}}{L_{10\_cf}}\right)^{\beta_{1\_cf}} = \left(\frac{L_{10}}{L_{10\_cf}}\right)^{\frac{1}{a_{1\_cf}}} = \exp\left[\frac{Y_{10} - Y_{10\_cf}}{a_{1\_cf}}\right] \text{ (non-linear model)}$$

$$= \exp\left[\frac{\left(a_1 - a_{1\_cf}\right)}{a_{1\_cf}}.X_{0.1} + \frac{\left(b_{1\_rel} - b_{1\_rel\_cf}\right)}{a_{1\_cf}}\right] \text{ (linear model)}$$

$$\frac{L_{15.91}}{L_{15.91\_cf}} = \exp\left[Y_{15.91} - Y_{15.91\_cf}\right] \text{ (non-linear model)}$$

$$= \exp\left[\left(a_1 - a_{1\_cf}\right).X_{0.1591} + b_{1\_rel} - b_{1\_rel\_cf}\right] \text{ (linear model)}$$

$$\left(\frac{L_{15.91}}{L_{15.91\_cf}}\right)^{\beta_{1\_cf}} = \left(\frac{L_{15.91}}{L_{15.91\_cf}}\right)^{\frac{1}{a_{1\_cf}}} == \exp\left[\frac{Y_{15.91} - Y_{15.91\_cf}}{a_{1\_cf}}\right] \text{ (non-linear model)}$$

$$= \exp\left[\frac{\left(a_1 - a_{1\_cf}\right)}{a_{1\_cf}}.X_{0.1591} + \frac{\left(b_{1\_rel} - b_{1\_rel\_cf}\right)}{a_{1\_cf}}\right] \text{(linear model)} \quad (39)$$

$$\frac{L_{0.1}}{L_{0.1\_cf}} = \exp\left[Y_{0.1} - Y_{0.1\_cf}\right] \text{ (non-linear model)}$$

$$= \exp\left[\left(a_2 - a_{2\_cf}\right).X_{0.001} + \left(b_{2\_rel} - b_{2\_rel\_cf}\right)\right] \text{ (linear model)}$$

$$\left(\frac{L_{0.1}}{L_{0.1\_cf}}\right)^{\beta_{2\_cf}} = \left(\frac{L_{0.1}}{L_{0.1\_cf}}\right)^{\frac{1}{a_{2\_cf}}} = \exp\left[\frac{Y_{0.1} - Y_{0.1\_cf}}{a_{2\_cf}}\right] \text{(non-linear model)} \quad (40)$$

$$= \exp\left[\frac{\left(a_2 - a_{2\_cf}\right)}{a_{2\_cf}}.X_{0.001} + \frac{\left(b_{2\_rel} - b_{2\_rel\_cf}\right)}{a_{2\_cf}}\right] \text{ (linear model)}$$

When using $L_{15.91\_G\_cf}$ defined with a_{1_cf}, the same relationships can be used with a correction factor f:

$$f=\left(\frac{\ln(1-0.1591)}{\ln(1-0.1)}\right)^{a_{1\_cf}-a_1}$$

$$\frac{L_{10}}{L_{10\_cf}}=\exp\left[Y_{10}-Y_{10\_cf}\right].f \text{ (non-linear model)}$$

$$=\exp\left[\left(a_1-a_{1\_cf}\right).X_{0.1}+b_{1\_rel}-b_{1\_rel\_cf}\right].f \text{ (linear model)}$$

$$\left(\frac{L_{10}}{L_{10\_cf}}\right)^{\beta_{1\_cf}}=\left(\frac{L_{10}}{L_{10\_cf}}\right)^{\frac{1}{a_{1\_cf}}}=\exp\left[\frac{Y_{10}-Y_{10\_cf}}{a_{1\_cf}}\right].f^{\frac{1}{a_{1\_cf}}} \text{ (non-linear model)}$$

$$=\exp\left[\frac{\left(a_1-a_{1\_cf}\right)}{a_{1\_cf}}.X_{0.1}+\frac{\left(b_{1\_rel}-b_{1\_rel\_cf}\right)}{a_{1\_cf}}\right].f^{\frac{1}{a_{1\_cf}}} \text{ (linear model)}$$

$$\frac{L_{15.91}}{L_{15.91\_cf}}=\exp\left[Y_{15.91}-Y_{15.91\_cf}\right].f \text{ (non-linear model)}$$

$$=\exp\left[\left(a_1-a_{1\_cf}\right).X_{0.1591}+b_{1\_rel}-b_{1\_rel\_cf}\right].f \text{ (linear model)}$$

$$\left(\frac{L_{15.91}}{L_{15.91\_cf}}\right)^{\beta_{1\_cf}}=\left(\frac{L_{15.91}}{L_{15.91\_cf}}\right)^{\frac{1}{a_{1\_cf}}}==\exp\left[\frac{Y_{15.91}-Y_{15.91\_cf}}{a_{1\_cf}}\right].f^{\frac{1}{a_{1\_cf}}} \text{ (non-linear model)} \quad (41)$$

$$=\exp\left[\frac{\left(a_1-a_{1\_cf}\right)}{a_{1\_cf}}.X_{0.1591}+\frac{\left(b_{1\_rel}-b_{1\_rel\_cf}\right)}{a_{1\_cf}}\right].f^{\frac{1}{a_{1\_cf}}} \text{ (linear model)}$$

$$f=\left(\frac{\ln(1-0.1591)}{\ln(1-0.1)}\right)^{a_{1\_cf}-a_1}$$

$$\frac{L_{0.1}}{L_{0.1\_cf}}=\exp\left[Y_{0.1}-Y_{0.1\_cf}\right].f \text{ (non-linear model)}$$

$$=\exp\left[\left(a_2-a_{2\_cf}\right).X_{0.001}+\left(b_{2\_rel}-b_{2\_rel\_cf}\right)\right].f \text{ (linear model)} \quad (42)$$

$$\left(\frac{L_{0.1}}{L_{0.1\_cf}}\right)^{\beta_{2\_cf}}=\left(\frac{L_{0.1}}{L_{0.1\_cf}}\right)^{\frac{1}{a_{2\_cf}}}=\exp\left[\frac{Y_{0.1}-Y_{0.1\_cf}}{a_{2\_cf}}\right].f^{\frac{1}{a_{2\_cf}}} \text{ (non-linear model)}$$

$$=\exp\left[\frac{\left(a_2-a_{2\_cf}\right)}{a_{2\_cf}}.X_{0.001}+\frac{\left(b_{2\_rel}-b_{2\_rel\_cf}\right)}{a_{2\_cf}}\right].f^{\frac{1}{a_{2\_cf}}} \text{ (linear model)}$$

One expects these ratios to be close to 1 and will demonstrate next that their median ratios are indeed close to 1 when conducting Monte Carlo simulations of this exercise.

More precisely, 1000 curve-fittings (linear and non-linear) of large databases (of relative lives) have been conducted, each database having been obtained by simulating 100 times 6 first-in-four failures (with random F values) for defining 100 times 6 relative lives (via the use of $L_{15.91\_G}$).

This exercise also will also lead to the derivation of the confidence intervals assigned to all ratios, as done in [2] and [3].

Median values and confidence intervals obtained via 1000 Monte Carlo simulations

Several outputs can be provided, especially when distinguishing how $L_{15.91\_G}$ and $L_{15.91\_G\_cf}$ are defined (using a_1, a_{1_cf} with the non-linear or linear curve-fitting). Following is one example corresponding to c=2:

	EXACT VALUES											
	a1	b1_rel	a2	b2_rel	F_Transition							
	1	1.75280728	0.5306	0.28621326	0.04301087							
	CURVE FITTING NON-LINEAR						CURVE FITTING LINEAR					
	a1/a1_cf	a2/a2_cf	a2_cf/a1_cf	r_exact/r_cf	L0.1cf/L10cf	r_exact/r_cf	a1/a1_cf	a2/a2_cf	a2_cf/a1_cf	r_exact/r_cf	L0.1cf/L10cf	r_exact/r_cf
Lower_0.05	1.0007	0.8553	0.2662	0.8124	0.0509	0.4693	1.0035	0.7515	0.2970	0.7068	0.0457	0.5208
Median_0.5	**1.0870**	**1.1110**	0.5128	**1.0348**	0.0760	**0.7368**	**1.0638**	**1.1621**	0.4886	**1.0859**	0.0742	**0.7545**
Upper_0.95	1.1625	2.2073	0.6531	1.9931	0.1193	1.1005	1.1310	1.8796	0.7507	1.7862	0.1075	1.2264
	CURVE-FITTED NON-LINEAR using L15.91_G						CURVE-FITTED LINEAR using L15.91_G					
	L0.1/L0.1_cf	L10/L10_cf	L15.91/L15.91_cf				L0.1/L0.1_cf	L10/L10_cf	L15.91/L15.91_cf			
Lower_0.05	0.4718	0.8866	0.9503				0.5165	0.9454	0.9977			
Median_0.5	**0.7345**	**0.9383**	**0.9821**				**0.7452**	**0.9871**	**1.0164**			
Upper_0.95	1.1026	0.9949	1.0309				1.2107	1.0278	1.0350			
	using a1_cf_non-linear for defining L15.91_G_cf:											
	CURVE-FITTED NON-LINEAR using L15.91_G_cf						CURVE-FITTED LINEAR using L15.91_G_cf					
	L0.1/L0.1_cf	L10/L10_cf	L15.91/L15.91_cf				L0.1/L0.1_cf	L10/L10_cf	L15.91L15.91_cf			
Lower_0.05	0.4476	0.8310	0.8918				0.4897	0.9018	0.9453			
Median_0.5	**0.7081**	**0.9019**	**0.9436**				**0.7213**	**0.9496**	**0.9778**			
Upper_0.95	1.0893	0.9912	1.0240				1.1874	1.0013	1.0186			
	using a1_cf_linear for defining L15.91_G_cf:											
	CURVE-FITTED NON-LINEAR using L15.91_G_cf						CURVE-FITTED LINEAR using L15.91_G_cf					
	L0.1/L0.1_cf	L10/L10_cf	L15.91/L15.91_cf				L0.1/L0.1_cf	L10/L10_cf	L15.91/L15.91_cf			
Lower_0.05	0.4595	0.8635	0.9255				0.5030	0.9208	0.9717			
Median_0.5	**0.7154**	**0.9138**	**0.9564**				**0.7257**	**0.9614**	**0.9899**			
Upper_0.95	1.0738	0.9689	1.004				1.1791	1.001	1.008			

Table 6: Results obtained using 1000 Monte Carlo simulations corresponding to c = 2

In the latter table, the symbol r (in the green cells) represents the left column number ratio, r_{exact}/r_{cf} beside a_{2_cf}/a_{1_cf} being for example equal to $\dfrac{a_2/a_1}{a_{2\_cf}/a_{1\_cf}}$.

One first can notice that most of the relevant median ratios are close to 1, (except the ratio involving $L_{0.1}$ which are slightly biased, for example with the non-linear curve-fitting: $L_{0.1\_cf}/L_{10\_cf}$= 0.076 instead of 0.056, leading to a ratio r_{exact}/r_{cf} =0.7368), confirming that the relative live can be used for retrieving useful information about the third and fourth reliability parameters. Also, most of the confidence 90% intervals are quite narrow.

Also, the difference observed using miscellaneous options are minor, so that one decided next to only show the results corresponding to the linear curve-fitting (circled in red), which was also the initial attractive point of our newly suggested model.

The following table summarizes the results obtained using two linear curve-fittings (in the respective range F <0.01 for $Y2$ and F > 0.05 for $Y1$) and four sets (a_2, $L_{0.1}$) corresponding to four values of c:

		USING 2 LINEAR CURVE-FITTINGS (F<0.01 & F > 0.05)								
		a1/a1_cf	a2/a2_cf	a2_cf/a1_cf	r_exact/r_cf	L0.1cf/L10cf	r_exact/r_cf	L0.1/L0.1_cf	L10/L10_cf	L15.91/L15.91_cf
c=2	Lower_0.05	1.0035	0.7515	0.2970	0.7068	0.0457	0.5208	0.5030	0.9208	0.9717
	Median_0.5	**1.0638**	**1.1621**	0.4886	**1.0859**	0.0742	**0.7545**	**0.7257**	**0.9614**	**0.9899**
	Upper_0.95	1.1310	1.8796	0.7507	1.7862	0.1075	1.2264	1.1791	1.0010	1.0080
c=4	Lower_0.05	1.0080	0.7249	0.1891	0.6761	0.1009	0.7518	0.7441	0.9409	0.9927
	Median_0.5	**1.0631**	**1.0810**	0.2912	**1.0199**	0.1333	**0.9405**	**0.9226**	**0.9797**	**1.0103**
	Upper_0.95	1.1276	1.6671	0.4393	1.5708	0.1667	1.2419	1.2131	1.0209	1.0273
c=10	Lower_0.05	1.0005	0.5634	0.1512	0.5296	0.1325	0.9733	0.9789	0.9497	0.9989
	Median_0.5	**1.0569**	**0.7809**	0.2213	**0.7394**	0.1590	**1.1435**	**1.1293**	**0.9912**	**1.0174**
	Upper_0.95	1.1200	1.1409	0.3089	1.0821	0.1868	1.3724	1.3545	1.0307	1.0350
c=175	Lower_0.05	0.9938	0.4008	0.1433	0.3767	0.1362	1.0870	1.0837	0.9503	0.9976
	Median_0.5	**1.0535**	**0.5633**	0.2063	**0.5375**	0.1609	**1.2622**	**1.2575**	**0.9888**	**1.0146**
	Upper_0.95	1.1154	0.8149	0.2944	0.7741	0.1869	1.4917	1.4659	1.0336	1.0330

Table 7: Summary using 1000 Monte Carlo simulations and two linear curve-fittings

One sees that the median ratios are indeed often close to 1, although slightly biased in some cases.

Of particular interest for example is the median ratio a_2/a_{2_cf} and $L_{0.1}/L_{0.1\_cf}$ plotted next versus the ratio a_{2_cf}/a_{1_cf} (defined using the median values of a_{2_cf} and a_{1_cf}):

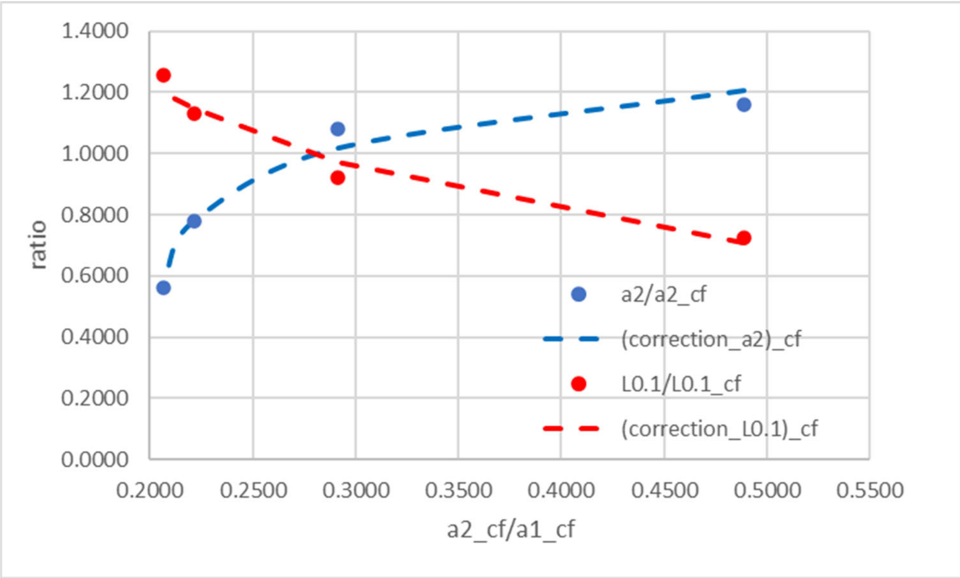


Fig. 10: Median ratios (full line) & correction factor (dotted lines) versus median a_{2_cf}/a_{1_cf}

A correction factor can be introduced (equal the latter median ratio) and curve-fitted for defining an unbiased or best estimate of a_2 and $L_{0.1}$ using:

$$a_2 = \left(correction\ \ for\ \ a_2\right)_{cf} * a_{2\_cf} \quad with \quad \left(correction\ \ for\ \ a_2\right)_{cf} = 1.409 + 0.1633 * \ln\left(\frac{a_{2\_cf}}{a_{1\_cf}} - 0.2\right)$$

$$L_{0.1} = \left(correction\ \ for\ L_{0.1}\right)_{cf} * L_{0.1\_cf} \ \ with \quad \left(correction\ \ for\ L_{0.1}\right)_{cf} = 0.459 * \left(\frac{a_{2\_cf}}{a_{1\_cf}}\right)^{-0.6084} \qquad (43)$$

The corrected curve-fitted ratios are then almost unbiased as shown next. The lower and upper bounds have also been corrected by the same correction factors and can be easily curve-fitted too.

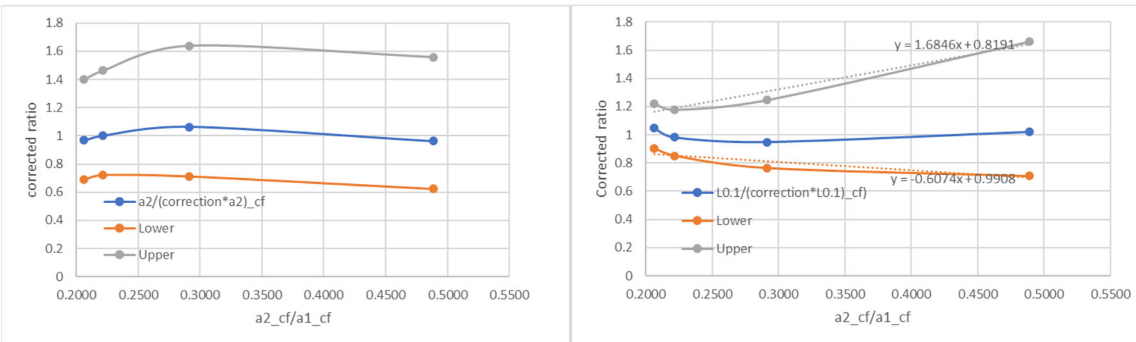


Fig. 11: Unbiased ratio obtained when using the correction factors

Using 600 points (hence 100 endurance tests with 6 first-in-4 failures), a_2 and $L_{0.1}$ can be correctly estimated with an accuracy often smaller than about ± 40%.

In the case of Fig. 9 for example, the best estimate of $L_{0.1}$ is obtained using the curve-fitted value of $L_{0.1\_cf}$= 0.0741 when using Eq. (35) to multiply by a correction factor equal to 0.72 when *a2_cf/a1_cf* is equal to 0.48, leading to a final estimate of $L_{0.1}$ equal to 0.053 instead of 0.056.

Conclusions

Following some work initiated in [2], a new reliability model is suggested in this paper where the failure probability F is calculated as a function of the life and 4 parameters: a_1, b_1, a_2 and b_2 or: a_1, L_{10}, a_2 and $L_{0.1}$ where L_{10} and $L_{0.1}$ are the lives corresponding to $F = 0.1$ and 0.001 respectively and a_1 and a_2 are the inverse of the Weibull slopes noticed in the large and low F range respectively (with $a_2 \leq a_1$). Two asymptotic linear Weibull models are therefore used with a non-linear smooth transition between these asymptotic lines when using Weibull plots.

For reducing the duration of endurance tests, a first-in-N testing strategy is often used with NR test rigs as described in [3] using N=4 and N=6 for example.

It has been demonstrated that the 6 failures can be analyzed using a standard two-parameter Weibull model for estimating the Weibull slope and the interpolated L_{50} life is then representative of the true $L_{15.91\_G}$ life of the group of 6 bearings.

But the 6 failures can also be analyzed using the 18 suspended items and this paper offers an exact calculation of F using the *inverse beta* function, the cumulative probability P, as well as N and NR in a general case. Using P =0.5 defines the median value of F_{median} to use for conducting the 2 or 4 parameter Weibull study. The Weibull slope and $L_{15.91\_G}$ life can then be defined directly, without passing via L_{50}.

P can also be fixed to 0.05 and 0.95, for defining $F_{0.05}$ and $F_{0.95}$ and understanding the 90 % variation range of the life using any models.

Defining a_2 and $L_{0.1}$ at low F values requires having access to a large database containing for example 1000 lives (sorted in ascending order). Such a database can be created numerically by simulating randomly 1000 values of F, sorted then in ascending order for defining the live corresponding to our four parameter models using fixed set of (a_1, L_{10}, a_2 and $L_{0.1}$).

An appropriate non-linear curve-fitting technique is suggested for defining the curve-fitted values of (a_1, L_{10}, a_2 and $L_{0.1}$) that can be compared to the curve-fitted values of (a_1 and L_{10}) and (a_2 and $L_{0.1}$) obtained using a simple linear curve-fitting in the respective range $F > 0.05$ and $F < 0.01$

This exercise confirmed the possibility of relying on two linear regressions for defining the set of 4 unknowns.

But having access to a real endurance database containing 1000 points is not realistic, so that one tested the idea suggested in [4] of using relative lives. The relative life represents the ratio $L/ L_{15.91\_G}$ where $L_{15.91\_G}$ is the life of the group of 6 bearing for example.

Using 100 endurance tests leads for example to 600 points to analyze using the non-linear and linear approach. The lowest median value of F is then equal to $2.888 \cdot 10^{-4}$ corresponding to the first failure out of 2400 tested bearings. The first 23 values of F are then smaller than 0.01, and we will demonstrate that 23 points are sufficient for analyzing the life corresponding to low F values.

Such a database can be obtained experimentally as used in [4] but can also be simulated and studied numerically by generating random values of F as done herein.

Of particular interest are the curve-fitted values of a_{2_cf} and $L_{0.1\_cf}/L_{10\_cf}$ that can be compared to the exact values a_2 and $L_{0.1}/L_{10}$.

The values of $L_{15.91\_G}$ can also be estimated using the 600 relative lives, L_{10} as reference and the exact slope a_1 or curve-fitted slope a_{1_cf}, leading to an estimate of $L_{0.1\_cf}$ and $L_{0.1}/L_{0.1\_cf}$ ratio.

This ratio, as well as many additional ones are further studied by conducting Monte Carlo simulations, duplicating for example 1000 times this exercise (using also miscellaneous sets of (a_2, $L_{0.1}$) for defining the median values of these ratios as well as their 90% confidence intervals.

Median ratios are often close to 1, confirming the possibility of using relative lives for retrieving the 4 parameters of our model.

Results obtained using the non-linear curve-fitting are only slightly more accurate that the ones obtained using two simple linear curve-fittings, hence the linear curve-fitting can finally be suggested.

Some of these ratios may be slightly biased, varying for example as a function of a_{2_cf}/a_{1_cf}, but a correction factor has been introduced and curve fitted as a function of a_{2_cf}/a_{1_cf} for a better estimation of a_2 and $L_{0.1}$ defined as a function of a_{2_cf} and $L_{0.1\_cf}$ and the latter correction factors.

As a summary, it can be said that the main benefits and novelties of this paper are the following:

- Exact calculations of F (median values and variation range) are provided using the *inverse beta* function applied to first in N testing strategy, NR test rigs or 100*NR test rigs.
- A new 4-parameter reliability model, duplicating quite well Rosemann's model at low and large failure rate F, is suggested.

- A simple linear curve-fitting can be used for retrieving the third and fourth parameter required for defining the life at low *F* values.
- Relative lives can be used for retrieving these four parameters and obtaining results at very low *F* values. Satisfactory confidence intervals about these four parameters have been obtained.

Finally, it can be recommended to see the main bearing manufacturers testing and sharing, in the frame of some ISO/DIN working committees for example, their relative live results for deriving estimates of $\left(\frac{L_{0.1}}{L_{10}} \ \& \ a_2\right)$ and abandoning the current conservative ISO suggestion $\left(\frac{L_0}{L_{10}} = 0.05\right)$.

Acknowledgements

The authors would like to acknowledge SMT for having sponsored this study and Prof. Poll for his interest in this work.

Appendix 1: Non-linear curve-fitting using the new model

The new model can be written:

$$Y_{cf\_i} = b_1 + a_1.X_i + \frac{(b_2-b_1)+(a_2-a_1).X_i}{1+\left(\frac{1-\exp[-\exp(X_i)]}{1-\exp\left[-\exp(\frac{b_2-b_1}{a_1-a_2})\right]}\right)^2} = b_1 + a_1.X_i + \frac{N_i}{D_i}$$

$$\begin{aligned}
&with: \\
&N_i = (b_2-b_1)+(a_2-a_1).X_i \\
&D_i = 1+\left(1-\exp[-\exp(X_i)]\right)^2.\left(1-\exp\left[-\exp(\frac{b_2-b_1}{a_1-a_2})\right]\right)^{-2} \\
&\quad = 1+\left(1-\exp[-\exp(X_i)]\right)^2.\left(1-\exp[-E]\right)^{-2} \qquad with \;\; E=\exp(\frac{b_2-b_1}{a_1-a_2})=\exp\left(X_{\text{intersection}}\right) \\
&\quad = 1+\left(1-\exp[-\exp(X_i)]\right)^2.\left(1-H\right)^{-2} \qquad with \;\; H=\exp[-E] \\
&\quad = 1+\left(1-\exp[-\exp(X_i)]\right)^2.G \qquad with \;\; G=\left(1-H\right)^{-2}=F_{\text{intersection}}^{-2}
\end{aligned} \tag{44}$$

One will now use Method 1 described in [2], minimizing the sum S^2 (also called $Si2$), for defining the four unknowns, the challenge being to calculate analytically the 4 partial derivatives.

$$S^2 = Si2 = \sum_{i=1,N} S_i^2 = \sum_{i=1,N}\left(Y_{cf\_i}-Y_{\exp\_i}\right)^2 = \min \quad (Method\,1) \tag{45}$$

$$\begin{cases}
f_1(a_1,b_1,a_2,b_2) = \dfrac{dS^2}{da_1} = 2.\displaystyle\sum_{i=1,N} S_i.\dfrac{dS_i}{da_1} = 2.\displaystyle\sum_{i=1,N} S_i.\dfrac{dY_{cf_i}}{da_1} = 0 \;\; with \;\; S_i=\left(Y_{cf_i}-Y_{\exp_i}\right) \\
f_2(a_1,b_1,a_2,b_2) = \dfrac{dS^2}{db_1} = 2.\displaystyle\sum_{i=1,N} S_i.\dfrac{dS_i}{db_1} = 2.\displaystyle\sum_{i=1,N} S_i.\dfrac{dY_{cf_i}}{db_1} = 0 \\
f_3(a_1,b_1,a_2,b_2) = \dfrac{dS^2}{da_2} = 2.\displaystyle\sum_{i=1,N} S_i.\dfrac{dS_i}{da_2} = 2.\displaystyle\sum_{i=1,N} S_i.\dfrac{dY_{cf_i}}{da_2} = 0 \\
f_4(a_1,b_1,a_2,b_2) = \dfrac{dS^2}{db_2} = 2.\displaystyle\sum_{i=1,N} S_i.\dfrac{dS_i}{db_2} = 2.\displaystyle\sum_{i=1,N} S_i.\dfrac{dY_{cf_i}}{db_2} = 0
\end{cases} \tag{46}$$

Let's first recall that when calling one of the four unknowns v:

$$\frac{d\left[\exp(f(v)\right]}{dv} = \frac{df}{dv}.\left[\exp(f(v)\right] \tag{47}$$

leading to the following successive calculations:

$$\frac{dE}{da_1} = -E.(b_2 - b_1).(a_1 - a_2)^{-2} \quad with \quad E = \exp\left(\frac{b_2 - b_1}{a_1 - a_2}\right)$$
$$\frac{dE}{db_1} = -E.(a_1 - a_2)^{-1}$$
$$\frac{dE}{da_2} = E.(b_2 - b_1).(a_1 - a_2)^{-2} = -\frac{dE}{da_1}$$
$$\frac{dE}{db_2} = E.(a_1 - a_2)^{-1} = -\frac{dE}{db_1} \tag{48}$$

$$\frac{dH}{dv} = -\exp(-E).\frac{dE}{dv} \qquad with \quad v = a_1 \; or \; b_1 \; or \; a_2 \; or \; b_2$$
$$\frac{dG}{dv} = 2.(1-\exp(-E))^{-3}.\frac{dH}{dv} = -2.(1-\exp(-E))^{-3}.\exp(-E).\frac{dE}{dv}$$
$$\frac{dD_i}{dv} = \left[1-\exp(-\exp(X_i))\right]^2.\frac{dG}{dv} = -\left[1-\exp(-\exp(X_i))\right]^2.2.(1-\exp(-E))^{-3}.\exp(-E).\frac{dE}{dv}$$

Or:

$$\frac{dD_i}{da_1} = \left[1-\exp(-\exp(X_i))\right]^2.2.F_{\text{intersection}}^{-3}.\exp(-E).E.(b_2 - b_1).(a_1 - a_2)^{-2}$$
$$\frac{dD_i}{db_1} = \left[1-\exp(-\exp(X_i))\right]^2.2.F_{\text{intersection}}^{-3}.\exp(-E).E.(a_1 - a_2)^{-1}$$
$$\frac{dD_i}{da_2} = -\frac{dD_i}{da_1}$$
$$\frac{dD_i}{db_2} = -\frac{dD_i}{db_1} \tag{49}$$

The next calculated steps are:

$$\frac{dN_i}{da_1} = -X_i \qquad \frac{dN_i}{db_1} = -1 \qquad \frac{dN_i}{da_2} = X_i \qquad \frac{dN_i}{db_2} = 1$$
$$\frac{d\left(\frac{N_i}{D_i}\right)}{dv} = \frac{\frac{d(N_i)}{dv}.D_i - \frac{d(D_i)}{dv}.N_i}{D_i^2}$$
$$\frac{dY_{cf\_i}}{dv} = \frac{d\left(\frac{N_i}{D_i}\right)}{dv} \tag{50}$$
$$add \; +X_i \;\; when \;\; calculating \;\; \frac{dY_{cf\_i}}{da_1}$$
$$add \; +1 \;\; when \;\; calculating \;\; \frac{dY_{cf\_i}}{db_1}$$

The set of 4 equations (Eq. 46) can now be solved using an iterative Newton-Raphson approach:

$$f_j(a_1+\Delta a_1, b_1+\Delta b_1, a_2+\Delta a_2, b_2+\Delta b_2) = f_j(a_1,b_1,a_2,b_2) + \frac{df_j}{da_1}.\Delta a_1 + \frac{df_j}{db_1}.\Delta b_1 + \frac{df_j}{da_2}.\Delta a_2 + \frac{df_j}{db_2}.\Delta b_2 = 0$$
$$or:$$
$$\frac{df_j}{da_1}.\Delta a_1 + \frac{df_j}{db_1}.\Delta b_1 + \frac{df_j}{da_2}.\Delta a_2 + \frac{df_j}{dc}.\Delta b_2 = -f_j(a_1,b_1,a_2,b_2) \quad for \quad j=1 \ to \ 4 \tag{51}$$

The partial derivatives df_j/dv are calculated using:

$$\frac{df_j}{da_1} = \frac{f_j(a_1+da_1,b_1,a_2,b_2) - f_j(a_1-da_1,b_1,a_2,b_2)}{2.da_1}$$
$$\frac{df_j}{db_1} = \frac{f_j(a_1,b_1+db_1,a_2,b_2) - f_j(a_1,b_1-db_1,a_2,b_2)}{2.db_1}$$
$$\frac{df_j}{da_2} = \frac{f_j(a_1,b_1,a_2+da_2,b_2) - f_j(a_1,b_1,a_2-da_2,b_2)}{2.da_2}$$
$$\frac{df_j}{db_2} = \frac{f_j(a_1,b_1,a_2,b_2+db_2) - f_j(a_1,b_1,a_2,b_2-db_2)}{2.db_2} \tag{52}$$

Appendix 2: Study of the relative life ratio

At large F

Using exact values and the linear asymptotic trend:

$$Y = \ln(L)$$

$$At \text{ large } F : Y1 = a_1.X + b_1$$

$$At\ F_{ref}\ \ (F_{ref} = 0.10\ \ for\ example) : X = X_{ref} = \ln\left(-\ln\left(1 - F_{ref}\right)\right)\ \&\ L = L_{ref} :$$

$$\ln(L_{ref}) = a_1.X_{ref} + b_1$$

$$b_1 = \ln(L_{ref}) - a_1.X_{ref}$$

$$For\ example : b_1 = \ln(L_{10}) - a_1.X_{0.10}\ \ \ with\ \ \ X_{0.10} = \ln\left(-\ln\left(1 - 0.10\right)\right)$$

$$So\ in\ general : \qquad Y1 = a_1.\left(X - X_{ref}\right) + \ln(L_{ref})$$

$$or\ in\ our\ case : \qquad Y1 = a_1.\left(X - X_{0.1}\right) + \ln(L_{10})$$

$$Y_{15.91} = \ln\left(L_{15.91}\right) = a_1.\left(X_{0.1591} - X_{ref}\right) + \ln(L_{ref})\ \ \ or\ \ \ a_1.\left(X_{0.1591} - X_{0.1}\right) + \ln(L_{10}) \qquad (53)$$

$$L_{15.91} = L_{ref}.\exp\left[a_1.\left(X_{0.1591} - X_{ref}\right)\right] = L_{10}.\exp\left[a_1.\left(X_{0.1591} - X_{0.1}\right)\right]$$

When using the relative life:

$$Y1_{rel} = \ln\left(\frac{L}{L_{15.91}}\right) = Y1 - Y_{15.91}$$

$$= a_1.\left(X - X_{ref}\right) + \ln(L_{ref}) - a_1.\left(X_{0.1591} - X_{ref}\right) - \ln(L_{ref}) \qquad (54)$$

$$= a_1.\left(X - X_{0.1591}\right)$$

$$Y1_{rel} = a_1.\left(X - X_{0.1591}\right) = a_1.X + b_{1\_rel}$$
$$with\ \ b_{1\_rel} = -a_1.X_{0.1591} \qquad irrespective\ of\ \ L_{ref} \qquad (55)$$

Using curve-fitted values and the linear asymptotic trend:

$Y1_{rel\_cf} = a_{1\_cf}.X + b_{1\_rel\_cf}$ to compare to $Y1_{rel} = a_1.X + b_{1\_rel}\quad with\quad b_{1\_rel} = -a_1.X_{0.1591}$ (56)

$$\frac{L_{10}}{L_{10\_cf}} = \frac{\dfrac{L_{10}}{L_{15.91}}}{\dfrac{L_{10\_cf}}{L_{15.91\_G}}} \approx \frac{\exp\left(Y1_{10\_rel}\right)}{\exp\left(Y1_{10\_rel\_cf}\right)} = \exp\left(Y1_{10\_rel} - Y1_{10\_rel\_cf}\right) \qquad (57)$$

$$= \exp\left[\left(a_1 - a_{1\_cf}\right).X_{0.1} + b_{1\_rel} - b_{1\_rel\_cf}\right]$$

In the latter relationship, it has been implicitly assumed that the linear relationships can be used and that $L_{15.91\_G}$ can be extrapolated using L_{10} as reference and the exact slope a_1 (as explained in the core of this paper, Eq. (38)).

$$\left(\frac{L_{10}}{L_{10\_cf}}\right)^{\beta_{1\_cf}} \approx \left(\frac{L_{10}}{L_{10\_cf}}\right)^{\frac{1}{a_{1\_cf}}} = \exp\left[\frac{\left(a_1 - a_{1\_cf}\right)}{a_{1\_cf}}.X_{0.1} + \frac{\left(b_{1\_rel} - b_{1\_rel\_cf}\right)}{a_{1\_cf}}\right] \quad (58)$$

Using the non-linear model:

The latter two relationships can be extrapolated to the use of the non-linear model, leading to:

$$\frac{L_{10}}{L_{10\_cf}} = \frac{\dfrac{L_{10}}{L_{15.91}}}{\dfrac{L_{10\_cf}}{L_{15.91\_G}}} \approx \frac{\exp\left(Y_{10\_rel}\right)}{\exp\left(Y_{10\_rel\_cf}\right)} = \exp\left(Y_{10\_rel} - Y_{10\_rel\_cf}\right) \quad (59)$$

$$\left(\frac{L_{10}}{L_{10\_cf}}\right)^{\beta_{1\_cf}} \approx \left(\frac{L_{10}}{L_{10\_cf}}\right)^{\frac{1}{a_{1\_cf}}} = \exp\left(\frac{Y_{10\_rel} - Y_{10\_rel\_cf}}{a_{1\_cf}}\right) \quad (60)$$

At low F:

Using exact values and the asymptotic linear trend:

$$At\ low\ F : Y2 = a_2.X + b_2$$

$$b_2\ defined\ \text{using}\ L_{0.1}\ corresponding\ to\ F = 0.001\ \ for\ example \quad (61)$$

$$b_2 = \ln\left(L_{0.1}\right) - a_2.X_{0.001} \quad with \quad X_{0.001} = \ln\left(-\ln\left(1-0.001\right)\right)$$

When using the relative life:

$$Y2_{rel} = \ln\left(\frac{L}{L_{15.91}}\right) = a_2.X + b_2 - a_1.\left(X_{0.1591} - X_{ref}\right) - \ln\left(L_{ref}\right)$$

$$= a_2.X + b_2 - a_1.\left(X_{0.1591} - X_{0.1}\right) - \ln\left(L_{10}\right) \quad (62)$$

$$if\ \ X_{ref} = X_{0.1}\ \ and\ \ L_{ref} = L_{10}$$

$$Y2_{rel} = a_2.X + b_{2\_rel}$$

$$b_{2\_rel} = b_2 - a_1.\left(X_{0.1591} - X_{ref}\right) - \ln\left(L_{ref}\right) = b_2 - a_1.\left(X_{0.1591} - X_{0.1}\right) - \ln\left(L_{10}\right) \quad (63)$$

Using curve-fitted values and the asymptotic linear trend:

$Y2_{rel\_cf} = a_{2\_cf}.X + b_{2\_rel\_cf}$ to compare to $Y2_{rel} = a_2.X + b_{2\_rel}$ (64)

$$\frac{L_{0.1}}{L_{0.1\_cf}} = \frac{\dfrac{L_{0.1}}{L_{15.91}}}{\dfrac{L_{0.1\_cf}}{L_{15.91\_G}}} \approx \exp\left[\left(a_2 - a_{2\_cf}\right).X_{0.001} + \left(b_{2\_rel} - b_{2\_rel\_cf}\right)\right] \quad (65)$$

$$\left(\frac{L_{0.1}}{L_{0.1\_cf}}\right)^{\beta_{2\_cf}} = \left(\frac{L_{0.1}}{L_{0.1\_cf}}\right)^{\frac{1}{a_{2\_cf}}} \approx \exp\left[\frac{\left(a_2 - a_{2\_cf}\right)}{a_{2\_cf}}.X_{0.001} + \frac{\left(b_{2\_rel} - b_{2\_rel\_cf}\right)}{a_{2\_cf}}\right] \quad (66)$$

Using the non-linear model:

$$\frac{L_{0.1}}{L_{0.1\_cf}} = \frac{\dfrac{L_{0.1}}{L_{15.91}}}{\dfrac{L_{0.1\_cf}}{L_{15.91\_G}}} \approx \exp\left[Y_{0.1} - Y_{0.1\_cf}\right] \qquad (67)$$

$$\left(\frac{L_{0.1}}{L_{0.1\_cf}}\right)^{\beta_{2\_cf}} = \left(\frac{L_{0.1}}{L_{0.1\_cf}}\right)^{\frac{1}{a_{2\_cf}}} \approx \exp\left[\frac{Y_{0.1} - Y_{0.1\_cf}}{a_{2\_cf}}\right] \qquad (68)$$

Ratio $L_{0.1}/L_{10}$ when using the linear model

Of main interest to users is also the ratio $L_{0.1}/L_{10}$:

$$Exact\ Calculations:$$

$$Y2_{0.1\_rel} = a_2 X_{0.001} + b_{2\_rel}$$
$$with\ b_{2\_rel} = b_2 - a_1.\left(X_{0.1591} - X_{0.1}\right) - \ln\left(L_{10}\right)$$

$$Y1_{10\_rel} = a_1.X_{0.1} + b_{1\_rel}$$
$$with\ \ b_{1\_rel} = -a_1.X_{0.1591} \qquad irrespective\ of\ \ L_{ref}$$

$$\ln\left(\frac{\dfrac{L_{0.1}}{L_{15.91\_G}}}{\dfrac{L_{10}}{L_{15.91\_G}}}\right) = \ln\left(\frac{L_{0.1}}{L_{10}}\right) = Y2_{0.1\_rel} - Y1_{10\_rel}$$
$$= a_2 X_{0.001} + b_{2\_rel} - a_1.X_{0.1} - b_{1\_rel}$$
$$\frac{L_{0.1}}{L_{10}} = \exp\left[a_2 X_{0.001} + b_{2\_rel} - a_1.X_{0.1} - b_{1\_rel}\right]$$
$$\left(\frac{L_{0.1}}{L_{10}}\right)^{\frac{1}{a_2}} = \exp\left[X_{0.001} + \frac{b_{2\_rel}}{a_2} - \frac{a_1}{a_2}.X_{0.1} - \frac{b_{1\_rel}}{a_2}\right]$$

$$Curve - fitted\ results:$$

$$Y2_{0.1\_rel\_cf} = a_{2\_cf} X_{0.001} + b_{2\_rel\_cf}$$
$$Y1_{10\_rel\_cf} = a_{1\_cf}.X_{0.1} + b_{1\_rel\_cf}$$

$$\ln\left(\frac{\dfrac{L_{0.1\_cf}}{L_{15.91\_G\_cf}}}{\dfrac{L_{10\_cf}}{L_{15.91\_G\_cf}}}\right) = \ln\left(\frac{L_{0.1\_cf}}{L_{10\_cf}}\right) = Y2_{0.1\_rel\_cf} - Y1_{10\_rel\_cf}$$
$$= a_{2\_cf} X_{0.001} + b_{2\_rel\_cf} - a_{1\_cf}.X_{0.1} - b_{1\_rel\_cf}$$
$$\frac{L_{0.1\_cf}}{L_{10\_cf}} = \exp\left[a_{2\_cf} X_{0.001} + b_{2\_rel\_cf} - a_{1\_cf}.X_{0.1} - b_{1\_rel\_cf}\right] \tag{69}$$
$$\left(\frac{L_{0.1\_cf}}{L_{10\_cf}}\right)^{\frac{1}{a_{2\_cf}}} = \exp\left[X_{0.001} + \frac{b_{2\_rel\_cf}}{a_{2\_cf}} - \frac{a_{1\_cf}}{a_{2\_cf}}.X_{0.1} - \frac{b_{1\_rel\_cf}}{a_{2\_cf}}\right]$$

Note that the curve-fitted ratio $L_{0.1\_cf}/L_{10\_cf}$ is independent about how $L_{15.91\_G}$ has been defined (using a_1 or a_{1_cf}, see previous discussion in this paper)

References

1. H. Rosemann, 'The Weibull Distribution and the Problem of Guaranteed Minimum Lifetimes', Bearing World Journal (2021), https://doi.org/10.15488/11469

2. L. Houpert, J. Clarke 'A study of the four parameter Rosemann's reliability model; Suggestion of a New four-parameter reliability model', Bearing World Journal (2022)

3. L. Houpert, 'An engineering approach to confidence intervals & endurance test strategies', STLE Tribology Trans., Vol 46 (2003), 2, 248-259.

4. M. N. Kotzalas, 'Statistical Distribution of Tapered Roller Bearing Fatigue Lives at High Levels of Reliability', ASME Jour. Tribology, Oct. 2005, Vol. 127, pp 1-6

5. L. Houpert, 2003, "Five methods for defining statistical confidence intervals", available upon request and permission of the Timken Company

6. J. McCool, "Evaluating Weibull Endurance Data by the Method of Maximum Likelihood", ASLE Transactions, 1970.

7. D. R. Thoman, L. J. Bain and C. E. Antle, "Inferences on the parameters of the Weibull Distribution", Technometrics, 1969.

8. S. Blachere, "A new bias correction technique for Weibull parametric estimation", Quality Engineering Applications and Research, 2015.

9. S. Blachere, A. Gabelli, "Monte Carlo comparison of Weibull two and three parameters in the context of the statistical analysis of rolling bearings fatigue testing", ASTM, 2012