



HAL
open science

On the Use of Cox Regression for Statistical Analysis of Fatigue Life Results

Kubat Narynbek Ulu, Bertrand Huneau, Erwan Verron, Anne-Sophie Béranger, Patrick Heuillet

► **To cite this version:**

Kubat Narynbek Ulu, Bertrand Huneau, Erwan Verron, Anne-Sophie Béranger, Patrick Heuillet. On the Use of Cox Regression for Statistical Analysis of Fatigue Life Results. *Journal of Testing and Evaluation*, 2020, 48 (2), pp.20180541. 10.1520/JTE20180541 . hal-04371162

HAL Id: hal-04371162

<https://hal.science/hal-04371162v1>

Submitted on 13 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On the Use of Cox Regression for Statistical Analysis of Fatigue Life Results

February 2, 2019

K. Narynbek Ulu^{a, 1}, B. Huneau^a, E. Verron^a, A.S. Béranger^b, P. Heuillet^b

^aInstitut de Recherche en Génie Civil et Mécanique (GeM), UMR CNRS 6183, Ecole Centrale de Nantes, Nantes, France

^bLRCCP, Vitry-sur-Seine, France

Abstract

Presence of scatter is inherent to fatigue life studies of materials and structures, for analysis of which a wide variety of statistical tools exist. Among them, Cox (proportional hazards) regression is widely encountered in survival studies of medical sciences, but has not found frequent use in fatigue of materials. Of interest is the ability to treat scatter in both applied load and fatigue life, as well as censored measurements of the latter. In the present work, this statistical approach is adapted for fatigue life studies. For ease of its application, it is numerically implemented as a Python module and provided for permissive free use. The module allows to carry out Cox regression on raw fatigue life results. Survival estimates, which provide a relationship between probability of failure and fatigue duration, can be additionally plotted with the module. Finally, the module allows to plot probabilistic Wöhler ($S-N$) curves, which lead to a more accurate representation of scatter in fatigue life results.

Keywords: fatigue life; statistical analysis; probabilistic $S-N$ curve.

Declarations of interest: none.

¹Corresponding Author: kubat.narynbek-ulu@ec-nantes.fr, 1 rue de la Noe, 44321 Nantes Cedex, France

Nomenclature

$\vec{\beta}$	=	vector of Cox regression coefficients
CSV	=	comma-separated values
$f(N)$	=	probability density function
$h(N)$	=	hazard function
$h_0(N)$	=	Cox baseline hazard function
HR	=	hazard ratio
HNBR	=	Hydrogenated Nitrile Butadiene Rubber
R	=	software environment for statistical computing
$S(N)$	=	survival function
$\hat{S}(N)$	=	survival function estimate
\vec{x}	=	vector of Cox covariates

1 Introduction

It is widely acknowledged that data scatter is inherent to fatigue life testing of materials and structures. Thus, statistical analysis of the results - namely, its adaptability and flexibility - plays an important role in the interpretation of experimental observations [1]. In the present work, because of expertise of the authors in the field of elastomer fatigue, the initial discussion is to be mostly focused on this material family; however, the general discussion on statistical analysis can be appropriately extended to a wide range of materials (including metals, composites, *etc.*) and structures/parts. In general, scatter appears irrespective of the type of control prescribed by testing machines - displacement, force, strain energy, or true stress. In elastomers and other materials, for example, scatter is mainly due to material defects introduced during processing and their influence on fatigue crack initiation [2, 3, 4, 5]. Noticeably, the magnitude of scatter depends also on the tested material. There exist a wide variety of statistical tools and strategies for analysis of fatigue life results. For the sake of simplicity, the two most frequently encountered strategies are discussed below; for a wider range of statistical approaches, the authors encourage familiarization with the reviews of Lee [6] and Crowder [7].

First, simple regression analysis is carried out due to its simplicity: a specific function (*e.g.* applied load as a function of fatigue life) is fitted to the available data. Moreover, there exists a phenomenological significance in availability of “laws”, which specify the function itself. Traditionally, the Basquin’s law of fatigue is often considered in analyzing fatigue life results, which constitutes a simple power law fit of the Wöhler curve [8, 9]; however, many more different and complex relationships exist [10]. The wide use of this type of regression analysis leads to differences in their application on fatigue life data. In general, these can be summarized into two approaches: regression is carried out for all available data points (at all loading levels); or it is carried out on average (arithmetic or geometric) at each separate loading level. However, the simplicity has its drawbacks: the results are poorly representative in presence of large scatter; and, treatment of statistically censored² results becomes complicated.

Second, parametric models consider statistical distributions - such as normal or log-normal - on sets of fatigue life measurements, in most cases grouped around a specific loading level.

²In fatigue life, right-censoring is commonly observed, where tests have to be stopped without achieving failure criteria, and it is known that fatigue life is above a certain duration but the exact fatigue life remains unknown.

Of special interest is the Weibull distribution, which is widely used for parametric statistical analysis in engineering and other fields [11]. For example, it is the *de facto* method employed for the statistical discussion of fatigue of metals and metallic parts [12]. This is a parametric approach, where the Weibull distribution is controlled by k , ρ , and γ , the so-called shape, scale, and location parameters, respectively. One of the main advantages of the Weibull distribution is that the shape parameter k can be changed to equal or approximate different types of distributions, *e.g.* $k = 1$ to the exponential, $k = 2$ to the Rayleigh, $k = 2.5$ approximates the log-normal, and $k = 3.6$ approximates the normal distributions. As evident, the Weibull and other statistical distribution analyses are mostly performed on only one variable (referred to as univariate analyses). Hence, presence of scatter in the loading (*i.e.* stress of the S - N curve) cannot be taken into account in the basic Weibull form. Moreover, methods exist to take into account right-censored data, *e.g.* for Weibull distribution [13, 14, 15]; however, they are limited by their parametric nature, which require large sample populations for a proper fit.

At this point, it should be mentioned that there are also non-parametric approaches. They are useful in cases where data does not fit a particular distribution or if the number of samples is small; for example, the fit of a Basquin's law or of a parametric distribution (such as Weibull) can be difficult and eventually provide erroneous conclusions. Here, we focus on the Kaplan-Meier estimator [16], which is a univariate non-parametric approach that is widely used in medical sciences to carry out survival analysis. Its main advantage lies in being a useful comparative tool when analyzing time-to-event studies, especially those with incomplete observations [17]. Such approach has been considered in fatigue studies of various metal alloys. However, the influence of its widespread use in medical sciences is evident as these metals are used for applications in prostheses or implants [18, 19], medical equipment [20, 21], *etc.*

Building upon these observations, the present work aims to develop a statistical tool, which tackles the following objectives:

- Obj1. Consider scatter of several variables (multivariate analysis);
- Obj2. Take into account right-censored data;
- Obj3. Ease of application in both academic and industrial environments; release of the numerical implementation under open source/permissive use;
- Obj4. Naturally, such tool should be extended to the plotting of S - N or Wöhler curves.

At the same time, it has to be considered that fatigue life results present a non-linear behavior for some materials at relatively low loading levels in form of endurance limits; hence, the statistical analysis should also reflect this phenomenon.

Thus, the Cox regression (or Cox proportional hazards model), widely used in medical sciences, can be introduced to solve these objectives [22]. Cox regression is a semi-parametric approach and has been applied in general reliability studies (as outlined in the review of Kumar and Klefjo [23]). However, interestingly, it appears that the Cox regression is rarely encountered in fatigue studies; there are few studies that focus on fatigue of prostheses [24] and medical equipment [21], as well as of oil pipelines [25], and engine turbine blades [26]. However, these are limited in their scope and relatively not highly cited; thus with the present work, the authors hope for a more general application of this powerful tool in fatigue studies. The application of the model, both mathematically and numerically, is presented in the following sections. For ease of future application, the corresponding code is presented within the text.

2 Cox Regression

2.1 FROM MEDICAL SCIENCE TO FATIGUE STUDIES

The extensive base of knowledge developed for application of Cox regression in medical sciences [27] can be applied to fatigue life studies, because a direct parallel can be drawn between the two [23, 28]. Putting it in simple terms, survival of a patient is parallel to fatigue life; similar connections can be made for treatment of categorical covariates (*e.g.* sex vs. material type), continuous covariates (*e.g.* age vs. applied load), censored data (*e.g.* patients leaving a study vs. end of experiments before failure), *etc.*

2.2 THEORY

The Cox proportional hazards model is based on the hazard function

$$h(N) = \frac{f(N)}{S(N)} \quad (1)$$

where $f(N)$ is the probability density function (PDF) and $S(N)$ is the survival function. Practically, the hazard function represents the probability that a specimen at cycle N will undergo failure in the next instant; often, it is referred to as the conditional (instantaneous) failure rate [11]. The hazard function of the Cox model is generalized to

$$h(N, \vec{x}) = h_0(N) \times \exp(\vec{\beta} \cdot \vec{x}) \quad (2)$$

where N is the fatigue life, $\vec{x} = (x_1, x_2, \dots, x_m)$ is a vector of m continuous or categorical covariates (such as load or material type, respectively), $\vec{\beta} = (\beta_1, \beta_2, \dots, \beta_m)$ is a vector of unknown regression coefficients that measure the impact of corresponding covariates, and $h_0(N)$ is referred to as the baseline hazard and corresponds to the hazard when $\vec{x} = 0$. The Cox proportional hazards model can naturally take into account both multivariate analysis and right-censored data, tackling objectives 1 and 2.

The exponent of a given coefficient β is referred to as the hazard ratio (HR) as

$$\text{HR}_i = \exp \beta_i \quad (3)$$

and its value indicates the following:

- $\text{HR}_i = 1$ - no effect of covariate i ;

- $HR_i > 1$ - as value of covariate i increases, *e.g.* load, the event hazard increases and subsequently duration of survival decreases;
- $HR_i < 1$ - the opposite, as value of covariate i increases, *e.g.* load, the event hazard decreases and duration of survival increases.

It is important to note that the *proportional* term of the “Cox proportional hazards model” name comes from the key assumption of the model: hazard curves are proportional for groups of observations. In other words, if applied on fatigue life data, this assumption implies that if one specimen (tested under one given condition) has a risk of failure (hazard) that is for example 10 times higher than under some other condition at some time t , then the risk of failure remains 10 times higher at any other time. Thus, it is important to examine the validity of this assumption (hence, it is implemented in the given module as will be shown in the following section). Further details and nuances of the Cox proportional hazards model have been extensively outlined by Cox [22], Grambsch & Therneau [29], Bradburn *et al.* [30], and many others.

2.3 NUMERICAL IMPLEMENTATION

To tackle objectives 3 and 4, *i.e.* the practical application of the method and subsequent plotting of Wöhler curves, a small *Python* module has been developed that allows for statistical analysis of fatigue data. The module (and its documentation) is released under permissive use of the MIT license [31], and the code and its documentation are available from the project’s *GitHub* page³. *Python* is chosen for its wide-ranging use in engineering applications (allowing for further integration), being open source, and available on leading operating systems. On the other hand, extensive and robust implementation of Cox analysis has been developed for the *R-programming language and environment* in the widely used, tested, and optimized *survival* library [32, 33]. Hence, the present module has been built on top of the *rpy2 Python* module that acts as an interface for the *R-environment*.

Within the *R-environment*, the *survival* library’s `coxph()` function is called: it fits a Cox proportional hazards model corresponding to equation (2). The function is called with the default parameters, particularly ties are resolved by the Efron approximation [33]. Moreover, functions `survfit()` and `cox.zph()` are called, respectively, for creation of survival function estimate curves and to test the proportionality hazards assumption for a Cox regression model

³<https://github.com/kubat-n/cox-fatigue-life>

fit; for the former, the Kaplan-Meier method [16] is used by default. For more details, refer to the extensive documentation of the library [33, 34].

From a user perspective, one is required to input just the following data on each tested specimen:

1. Test time durations (in consistent units);
2. Survival statuses (according to fatigue life criteria) at these respective durations: 1 for failed and 0 for survived;
3. Load magnitudes;
4. Any other covariates (variables) that are tested, *e.g.* material type, *etc.*

The input is arranged in a m -by- n matrix as shown in Table 1 and can be additionally imported in form of a CSV (comma-separated value) spreadsheet. Afterwards, with calling appropriate functions of the module, one can extract statistical information of the fit: the hazard ratio, statistical significance (p -values and other tests), *etc.* Furthermore, one is able to plot a Kaplan-Meier survival estimate for any combination of covariates: the corresponding figure shows the evolution of probability with respect to the units of fatigue life measurement. Finally, Cox regression is then adapted for creation of probabilistic Wöhler (S - N) curves. A practical example is presented in the following section.

3 Practical Example

3.1 *EXPERIMENTAL PROGRAM*

Fatigue life tests are performed on two formulations (materials A and B) of an elastomer, HNBR (hydrogenated nitrile butadiene rubber); the differences between the two materials are on a chemical level, where the weight percentage of acrylonitrile monomers and the percentage of hydrogenation are changed for both materials. The relevant details of the experimental program are the following:

- Flat dumbbell specimens are utilized in uniaxial tension with a gauge length of 10 mm;
- The applied load is controlled in true stress with four target constant amplitudes (4, 6, 8, and 10 MPa) at $R=0$; here, *target* implies that the applied true stress is not exact at each loading level and for each tested specimen, as would be ideally required in order to carry out any classic univariate approach; as a consequence, analysis of data takes a multivariate character by introduction of “scatter” along the loading axis;
- End of failure criterion is considered as complete failure of a specimen; at the lowest amplitude of 4 MPa, tests have been stopped after 2 million cycles (about 6 days in duration), thereby introducing right-censoring in some of the results.

For more information, the details of the procedure are available in [35, 36].

3.2 *APPLICATION AND RESULTS*

Table 2 contains the raw fatigue data for the two formulations and Fig. 1 shows the corresponding values. It is evident that the present data set exhibits the two aforementioned difficulties – scatter in both loading (stress) and fatigue life, and censored points around 2 million cycles – which justify the use of the Cox model.

The Cox analysis starts by creating of a `CoxFatigueAnalysis` object (where `analysis` is the name assigned to the object).

```
In [1]: analysis = CoxFatigueAnalysis(['stress', 'material'],[float, str])
```

Here, the input variables are two lists containing the names of the covariates to be tested and their subsequent *Python* data types. Second, the data are imported from a file that is arranged in a matrix as depicted in Table 1.

```
In [1]: analysis.import_from_csv('input_data.csv')
```

Subsequently, Cox regression is performed on the data by calling the function `cox_regression`.

The proportionality assumption of the Cox model can be tested by the function `cox_zph`.

```
In [1]: analysis.cox_regression()
```

```
In [2]: analysis.cox_zph()
```

The results of the regression and the proportionality test can be obtained by `print(analysis)` or by calling `print_cox_r_output` (which displays the output generated by the `summary` function of the *R-environment*). The results of the cox regression as applied on the data of Tab. 2 are:

```
In [1]: print(analysis)
```

```
Out [2]:
```

```
Cox proportional hazards model output:
```

```
Cox regression on variables stress , materialMaterialB , :
```

	beta	se(beta)	HR-.95	HR	HR+.95	p
stress	0.97237	0.11074	2.1283	2.6442	3.2852	0.0
materialMaterialB	-1.3573	0.29184	0.14525	0.25735	0.45598	3.3063e-06

```
Proportionality assumption from cox_zhp() (p>0.05):
```

	rho	chisq	p
stress:	0.0599	0.1637	0.68573
materialMaterialB:	0.0486	0.1281	0.72045
GLOBAL:	nan	0.5049	0.77688

With respect to the regression results, the coefficient (**beta**), the hazard ratio (**HR**) and its lower (**HR-.95**) and upper (**HR+.95**) 95% confidence intervals, as well as the log-rank test *p*-value are of most interest. For the continuous covariate **stress**, the hazard ratio $HR = 2.64 > 1$ indicates a strong relationship between the increase in true stress amplitude and the increased risk of failure. This can be interpreted as: in most general terms, an additional increase in amplitude by 1 MPa induces the hazard of failure (at each respective cycle) by a factor of 2.64 for both **MaterialA** and **MaterialB**. Of course, such trend is clear even from simple observations of Fig. 1 and, moreover, they are confirmed to be statistically significant from the *p*-value of the log-rank test $p = 0 < 0.05$. A similar interpretation can be made in comparison of the categorical covariate **material**; note that only result for **MaterialB** is shown since **MaterialA** is taken as reference by default. Thus, the hazard of failure of **MaterialB** with respect to **MaterialA** decreases by around four times ($HR = 0.257$) with statistical significance

$p = 3.31 \times 10^{-6} < 0.05$. The relatively low p -values of each respective covariate can be said to be indicative that the data are generally well described by the model.

With respect to the proportionality test results, of great interest is the p -value for global test and each covariate. High statistical significance, $p < 0.05$, indicates a strong evidence of non-proportionality of hazards of each covariate; the proportionality assumption is supported by a non-significant relationship, *i.e.* when $p > 0.05$. In the present case, the p -values for **stress**, **MaterialB**, and globally are $p = 0.69 > 0.05$, $p = 0.72 > 0.05$, and $p = 0.78 > 0.05$ respectively and, thus, indicating that one can assume proportional hazards.

Furthermore, the present module allows to visualize the survival function estimate obtained from the Cox regression; as mentioned earlier, the Kaplan-Meier method is utilized to plot the survival estimate. The Kaplan-Meier method is a non-parametric approach and thus represents the experimental estimate of the survival function. For example, one can visualize with the function `plot_cox_survival_1var` the survival estimate for Material B at loading levels of 4, 6, 8, and 10 MPa (since the Cox model can take into account continuous variables, the choice of these four specific numbers is rather arbitrary and any other value(s) of stress can be considered).

```
In [1]:
analysis.plot_cox_survival_1var(axes, # matplotlib axes
                               'stress', # covariate that changes
                               [4,6,8,10], # values of the stress
                               ['material'], # constant covariates
                               ['MaterialB']) # their respective values
```

The corresponding plot is shown in Fig. 2a. For each curve corresponding to a stress value, the probability of survival decreases with increasing number of cycles as expected. Similarly, if one compares the survival curves at four stress values, it is evident that the probability of survival decreases faster with increasing stress.

For comparison, classic Weibull analysis is carried out to estimate the survival function of the corresponding data (even though, as mentioned earlier, the Weibull distribution is not an optimal choice when the statistical sample size is small and right-censoring is present). The parameters are calculated in *R-environment* using the *survival* library. For visual analysis, the results are plotted in Fig. 2b on top of the Kaplan-Meier curves to aid in the comparison of results. It is clear that the parametric approach of the Weibull fitting does not exactly

correspond to the non-parametric approach of the Kaplan-Meier method. It is especially evident for short fatigue lives at 10 MPa and 8 MPa target loading levels (where the survival function is underestimated by the Weibull fit), for those longer than 5 thousand cycles at 8 MPa and 6 MPa (overestimated), and for those longer than 80 thousand cycles at 4 MPa (overestimated).

Furthermore, the function `plot_wohler_curve` is used to plot the probabilistic Wöhler curves, *e.g.* the following code plots such curve for Material A from 4 to 10 MPa (as well as extrapolations to 2 MPa and 12 MPa).

```
In [1]:
mat_a = analysis.plot_wohler_curve(axw, # axes object of matplotlib
                                   'stress', # covariate load on S-N curve
                                   [2,12], # list of load values [min, max]
                                   0.1, # set the resolution of the load-axis
                                   ['material'], # name(s) of constant covariate(s)
                                   ['MaterialA'] # corresponding values
                                   linecolors=('y','y','k','k','k') # line colors
                                   )
```

Figs. 3 and 4 show the corresponding probabilistic Wöhler curves for materials A and B, respectively. For ease of visualization, the probabilities are plotted as contour regions of different colors, each corresponding to a decrease of 5% in survival probability. Actual experimental data points are also shown for comparison. Furthermore, solid lines correspond to almost zero (indicated as $\widehat{S}(N) = 0.001$) and 100% (indicated as $\widehat{S}(N) = 0.999$) probabilities of survival, as well as the mean survival (indicated as $\widehat{S}(N) = 0.5$); the upper and lower 95% probability bounds are shown as dashed lines (indicated as $\widehat{S}(N) = 0.05$ and $\widehat{S}(N) = 0.95$ respectively). Unlike what would be expected for a realistic material response, it should be noted that these lines are not smooth and appear somewhat irregular; this is a consequence of the discrete nature of the estimation of the survival function as shown in Fig. 2a. It is expected that as more specimens are tested, the smoother these curves will become. Moreover, there is presence of horizontal asymptotic lines around 3 MPa (extrapolated stress) for Material A (Fig. 3) and around 4.5 MPa for Material B (Fig. 4) corresponding to the mean lines; such horizontal asymptotes are also visible for the lines for the almost zero probability (indicated as 0.001) and the lower 95% bound (indicated as 0.050). It is of importance to note that although presence of horizontal asymptotes could indicate presence of an endurance limit, such fact does not explicitly guarantee it; these phenomena are caused by missing data (observed

broken specimens or non-censored data) within the range of the fatigue lives from about 350 thousand cycles to final end of experiments at 2 million cycles as indicated by the hatched region. Thus, there is a loss of resolution in the probability values within this range and it is impossible to determine them without carrying out additional experiments. This presents a limitation of the method in terms of available experimental points; however, at the same time, it allows to identify additional required testing. Finally, it should be noted that the time domain is limited by the last experimental observation made, i.e. at 2 million cycles. However, as stated earlier, special attention has to be paid in interpreting the results within the domain region where horizontal lines are observed (hatched regions in Figs. 3 and 4).

Otherwise, the probabilistic Wöhler curves of the two materials correspond to the general hazard ratios obtained from Cox regression. For both materials, as expected, survival probability decreases as number of cycles increases and as applied load increases. For ease of comparison, we implemented a specific function to graphically compare two Wöhler curves for two different values of some covariate. Here, a curve of each of the two materials, corresponding either to almost zero or almost complete survival, the mean, or the upper and lower 95% probability bounds, can be plotted for comparison. For example, the mean of Material A and Material B and their corresponding curves shown in Figs. 3 and 4 can be plotted together with the following code:

```
In [1]:
analysis.compare_curves(ax, # axes object of matplotlib
                        mat_a, # wohler curve of material A
                        mat_b, # wohler curve of material B
                        'mean', # selection of which curve to plot
                        # other options:
                        # 'max95', 'min95', 'survival', 'failure'
                        ['Material_A', 'Material_B'] # labels for the legend
                        )
```

The corresponding plot is shown in Fig. 5, which clearly highlights, in a conventional manner, that Material B has better fatigue resistance than Material A; for comparison, the original data points are also plotted.

4 Conclusions

Cox regression (or Cox proportional hazards model) is widely used in medical sciences; however, its application in fatigue studies has been limited even considering the direct relationship between survivability and fatigue studies. Its use should be considered however, since the model allows to tackle many commonly encountered problems. First, being a multivariate analysis, it allows taking into consideration multiple covariates (variables); moreover, these covariates can be both categorical (such as type of material) or continuous (such as applied load). From the perspective of fatigue tests, this allows to deal with scatter not only in fatigue life durations, but in other variables (such as loading). Second, it allows to treat censored results, especially right-censoring when fatigue life tests have to be stopped before reaching failure.

In the present paper, the theory behind the Cox regression analysis has been summarized and a subsequent *Python* module is developed (and provided) for its straightforward application on fatigue life results. The module computes important Cox analysis values, such as the hazard ratios and the statistical significance; additionally, the proportionality assumption of the model can be tested. Furthermore, survival estimate curves are available for a combination of various covariates. The module also plots probabilistic Wöhler ($S-N$) curves, which propose a more complete representation of scatter in fatigue life results. Finally, two different Wöhler curves can be compared. A practical example on actual experimental results is given to demonstrate the relevance of the developed module.

By releasing the present module under permissive free use license, the authors hope for greater use of Cox regression in fatigue studies as well as for further improvement of the module itself.

References

- [1] W. Schütz, “A history of fatigue,” *Eng. Fract. Mech.* 54, no. 2 (May 1996): 263–300, [http://dx.doi.org/10.1016/0013-7944\(95\)00178-6](http://dx.doi.org/10.1016/0013-7944(95)00178-6)
- [2] S. Svensson, “Testing methods for fatigue properties of rubber materials and vibration isolators,” *Polym. Test.* 2, no. 3 (July/September 1981): 161–174, [http://dx.doi.org/10.1016/0142-9418\(81\)90002-7](http://dx.doi.org/10.1016/0142-9418(81)90002-7)
- [3] W. V. Mars and A. Fatemi, “A literature survey on fatigue analysis approaches for rubber,” *Int. J. Fatigue* 24, no. 9 (September 2002): 949–961, [http://dx.doi.org/10.1016/S0142-1123\(02\)00008-7](http://dx.doi.org/10.1016/S0142-1123(02)00008-7)
- [4] W. V. Mars and A. Fatemi, “Factors that affect the fatigue life of rubber: A literature survey,” *Rubber Chem. Technol.* 77, no. 3 (July 2004): 391–412, <http://dx.doi.org/10.5254/1.3547831>
- [5] Y. L. Tee, M.S. Loo, and A. Andriyana, “Recent advances on fatigue of rubber after the literature survey by Mars and Fatemi in 2002 and 2004,” *Int. J. Fatigue* 110 (May 2018): 115–129, <http://dx.doi.org/10.1016/J.IJFATIGUE.2018.01.007>
- [6] Y.-L. Lee, *Fatigue Testing and Analysis: Theory and Practice* (Oxford: Butterworth-Heinemann, 2005).
- [7] M. J. Crowder, *Statistical Analysis of Reliability Data* (New York: Routledge, 2017).
- [8] O. H. Basquin, “The exponential law of endurance tests,” *Proc. Am. Soc. Test. Mater.* 10, no. 2 (1910): 625–630.
- [9] F. Kun, H. A. Carmona, J. S. Andrade, and H. J. Herrmann, “Universality behind Basquin’s Law of Fatigue,” *Phys. Rev. Lett.* 100, no. 9 (March 2008): 094301, <http://dx.doi.org/10.1103/PhysRevLett.100.094301>
- [10] E. Castillo and A. Fernández-Canteli, *A Unified Statistical Methodology for Modeling Fatigue Damage* (Springer Netherlands, 2009).
- [11] J. P. Klein and M. L. Moeschberger, *Survival Analysis: Techniques for Censored and Truncated Data* (New York: Springer, 2003).

- [12] R. I. Stephens, A. Fatemi, S. R. Robert, and H. O. Fuchs, *Metal Fatigue in Engineering* (New York: Wiley, 2001).
- [13] R. A. Lockhart and M. A. Stephens, "Estimation and tests of fit for the three-parameter Weibull distribution," *J. R. Stat. Soc. Ser. B* 56, no. 3 (1994): 491-500, <http://dx.doi.org/10.2307/2346122>
- [14] R. B. Abernethy, *The New Weibull Handbook* (Oxford: Elsevier Science and Technology, 1996).
- [15] D. Cousineau, "Fitting the three-parameter Weibull distribution: review and evaluation of existing and new methods," *IEEE Trans. Dielectr. Electr. Insul.* 16, no. 1 (February, 2009): 281–288, <http://dx.doi.org/10.1109/TDEI.2009.4784578>
- [16] E. L. Kaplan and P. Meier, "Nonparametric estimation from incomplete observations," *J. Am. Stat. Assoc.* 53, no. 282 (1958): 457-481, <http://dx.doi.org/10.2307/2281868>
- [17] J. T. Rich, J. G. Neely, R. C. Paniello, C. C. J. Voelker, B. Nussenbaum, and E. Wang, "A practical guide to understanding Kaplan-Meier curves," *Otolaryngol. Head. Neck Surg.* 143, no. 3 (September 2010): 331–336, <http://dx.doi.org/10.1016/j.otohns.2010.05.007>
- [18] S. S. Scherrer, M. Cattani-Lorente, E. Vittecoq, F. De Mestral, J.A. Griggs, and H. W. A. Wiskott, "Fatigue behavior in water of Y-TZP zirconia ceramics after abrasion with 30 um silica-coated alumina particles," *Dent. Mater.* 27, no. 2 (February 2011): e28-42, <http://dx.doi.org/10.1016/j.dental.2010.10.003>
- [19] K. S. Chan, M. Koike, R. L. Mason, and T. Okabe, "Fatigue life of titanium alloys fabricated by additive layer manufacturing techniques for dental implants," *Metall. Mater. Trans. A Phys. Metall. Mater. Sci.* 44, No. 2 (February 2013): 1010–1022, <http://dx.doi.org/10.1007/s11661-012-1470-4>
- [20] S. G. Fitzgerald, R. A. Cooper, M. L. Boninger, and A. J. Rentschler, "Comparison of fatigue life for 3 types of manual wheelchairs," *Arch. Phys. Med. Rehabil.* 82, no. 10 (October 2001): 1484–1488, <http://dx.doi.org/10.1053/apmr.2001.26139>

- [21] E. Fišerová, M. Chvosteková, B. Silvie, and M. Bumbálek, “Survival analysis of factors influencing cyclic fatigue of nickel-titanium endodontic instruments,” *Adv. Mater. Sci. Eng.* 2015, (2015): 1-6, <http://dx.doi.org/10.1155/2015/189703>
- [22] D. R. Cox, “Regression models and life-tables,” *J. R. Stat. Soc. Ser. B* 34, no. 2 (1972): 187–220, http://dx.doi.org/10.1007/978-1-4612-4380-9_37
- [23] D. Kumar and B. Klefsjö, “Proportional hazards model: a review,” *Reliab. Eng. Syst. Saf.* 44, no. 2 (1994): 177–188, [http://dx.doi.org/10.1016/0951-8320\(94\)90010-8](http://dx.doi.org/10.1016/0951-8320(94)90010-8)
- [24] J. A. Tang, J. M. Leasure, J. S. Smith, J. M. Buckley, D. Kondrashov, and C. P. Ames, “Effect of severity of rod contour on posterior rod failure in the setting of lumbar pedicle subtraction osteotomy (PSO),” *Neurosurgery* 72, no. 2 (February 2013): 276–283, <http://dx.doi.org/10.1227/NEU.0b013e31827ba066>
- [25] K. Farhangdoost and M. Siahpoosh, “On the fatigue life prediction of die-marked drillpipes,” *Design and Analysis* 3, (July 2006): 461–468, <http://dx.doi.org/10.1115/PVP2006-ICPVT-11-93181>
- [26] J. Zhang, J. Lin, G. Zhang, and H. Liu, “High cycle fatigue life prediction and reliability analysis of aeroengine blades,” *Trans. Tianjin Univ.* 18, no. 6 (December 2012): 456–464, <http://dx.doi.org/10.1007/s12209-012-1785-7>
- [27] R. Singh and K. Mukhopadhyay, “Survival analysis in clinical trials: Basics and must know areas,” *Perspect. Clin. Res.* 2, no. 4 (October 2011): 145–148, <http://dx.doi.org/10.4103/2229-3485.86872>
- [28] J. F. Lawless, *Statistical models and methods for lifetime data* (Hoboken, NJ: Wiley-Interscience, 2003).
- [29] P. M. Grambsch and T. M. Therneau, “Proportional hazards tests and diagnostics based on weighted residuals,” *Biometrika* 81, no. 3 (September 1994): 515–526, <http://dx.doi.org/10.1093/biomet/81.3.515>
- [30] M. J. Bradburn, T. G. Clark, S. B. Love, and D. G. Altman, “Survival analysis part II: multivariate data analysis—an introduction to concepts and methods,” *Br. J. Cancer* 89, no. 3 (August 2003): 431–437, <http://dx.doi.org/10.1038/sj.bjc.6601119>

- [31] A. Morin, J. Urban, and P. Sliz, “A Quick Guide to Software Licensing for the Scientist-Programmer,” *PLoS Comput. Biol.* 8, no. 7 (July 2012): e1002598, <http://dx.doi.org/10.1371/journal.pcbi.1002598>
- [32] J. Fox and S. Weisberg, “Cox proportional-hazards regression for survival data in R,” in *An R Companion to Applied Regression, Second Edition* (London: Sage Publications, 2011).
- [33] T. M. Therneau and T. Lumley, ”Survival Analysis Guide,” CRAN, 2017, <http://web.archive.org/web/20190127202342/https://cran.r-project.org/web/packages/survival/survival.pdf>
CRAN. <http://web.archive.org/web/20190127202342/https://cran.r-project.org/web/packages/survival/survival.pdf>
- [34] T. M. Therneau and P. M. Grambsch, *Modeling Survival Data: Extending the Cox Model* (New-York: Springer, 2000).
- [35] K. Narynbek Ulu, B. Huneau, E. Verron, A.-S. Béranger, and P. Heuillet, “True stress controlled fatigue life experiments for elastomers,” *Int. J. Fatigue* 104, (November 2017): 171–182, <http://dx.doi.org/10.1016/j.ijfatigue.2017.07.007>
- [36] K. Narynbek Ulu, “Fatigue of HNBR: Effects of Formulation and Thermal Aging” (PhD diss., Ecole Centrale de Nantes, 2018).

5 Tables with captions

Table 1: Arrangement of input data.

Rows	Columns 1 to n					
1 to m	fatigue_life	fatigue_survival	load (covariate.1)	covariate_2	...	covariate_(n-2)
1	<i>duration 1</i>	<i>status 1</i>	<i>load 1</i>	<i>value_2 1</i>	<i>...</i>	<i>value_(n-2) 1</i>
...	<i>...</i>	<i>...</i>	<i>...</i>	<i>...</i>	<i>...</i>	<i>...</i>
2	<i>duration m</i>	<i>status m</i>	<i>load m</i>	<i>value_2 m</i>	<i>...</i>	<i>value_(n-2) m</i>

Table 2: Fatigue lives of materials A and B; survival status of 1 indicates a broken specimen, 0 - censored point.

Material A			Material B		
Fatigue Life (Cycles)	Status	True Stress (MPa)	Fatigue Life (Cycles)	Status	True Stress (MPa)
566	1	10.37	3085	1	10.34
931	1	10.04	790	1	9.95
610	1	10.01	564	1	9.95
70	1	9.99	1845	1	9.91
285	1	9.83	933	1	9.91
550	1	9.80	622	1	9.87
98	1	9.76	2359	1	9.49
262	1	9.70	797	1	9.44
877	1	8.43	5752	1	8.14
4434	1	8.30	2665	1	7.99
1373	1	8.27	3222	1	7.99
1218	1	8.05	16082	1	7.99
3332	1	8.02	781	1	7.93
884	1	7.97	915	1	7.79
2419	1	7.78	2858	1	7.69
1353	1	7.49	37484	1	7.68
3565	1	6.15	2223	1	7.67
2632	1	6.07	8113	1	7.55
1136	1	6.00	12488	1	6.06
9599	1	6.00	54051	1	6.00
2700	1	5.94	37485	1	5.99
3595	1	5.88	169456	1	5.99
13599	1	5.85	21553	1	5.96
1372	1	5.70	114465	1	5.92
199734	1	4.22	44332	1	5.87
1999500	0	4.13	57267	1	5.87
353376	1	3.98	1999900	0	4.02
49040	1	3.97	85904	1	4.02
288358	1	3.95	1999800	0	4.00
57521	1	3.94	1999900	0	3.96
207959	1	3.93	1999900	0	3.94
6474	1	3.93	1999900	0	3.93
1999500	0	3.90	76374	1	3.93
30048	1	3.88	1999900	0	3.92
72278	1	3.80	1999900	0	3.91
313566	1	3.78	221901	1	3.82

6 List of figure captions

Figure 1: Raw fatigue lifetime data for materials A and B.

(a)

(b)

Figure 2: Survival function estimates for different stress values of Material B; (a) Kaplan-Meier as returned by `plot_cox_survival_1var()`; (b) Comparison with a classic Weibull fit; shape, k , and scale, λ , parameters are indicated in the legend.

Figure 3: Probabilistic Wöhler curve for Material A; actual experimental points are shown; hatched region indicates loss of resolution in probability (see the text for details).

Figure 4: Probabilistic Wöhler curve for Material B; actual experimental points are shown; hatched region indicates loss of resolution in probability (see the text for details).

Figure 5: Comparison of mean curves for materials A and B; hatched region indicates loss of resolution in probabilities.

