Fitting of lifetime distributions with Bathtub-Shaped Failure Rate to Data Using Optimization Algorithms

Aalto University School of Chemical Engineering

Teemu J. Ikonen[a], Francesco Corona[a], Iiro Harjunkoski[a,b]

[a]Aalto University, Finland   [b]Hitachi Energy Research, Germany

Abstract

Equipment lifetime distributions with bathtub-shaped failure rate can be fitted to data by the maximum likelihood estimate. In the literature, a commonly used method is to find a point in the parameter space where the partial derivatives of the log-likelihood function are zero. As the log-likelihood function is typically non-convex, this approach may yield a sub-optimal fit. In this work, we maximize the log-likelihood function by three nonlinear optimization algorithms (i.e., Nelder-Mead with adaptive parameters, SLSQP, and L-BFGS-B) by performing a multi-start of 100 optimization procedures. We perform a systematic study of refitting 10 key lifetime distributions with bathtub-shaped failure rate from the literature to two widely studied datasets [1, 10].

Motivation

- An industrial plant generates (as a byproduct) a large amount of lifetime data of its components (e.g., pumps, motors, and drives)
- The data (time of failure) may look, for example, like this

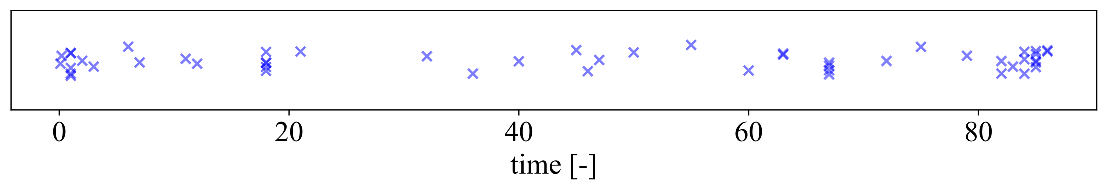
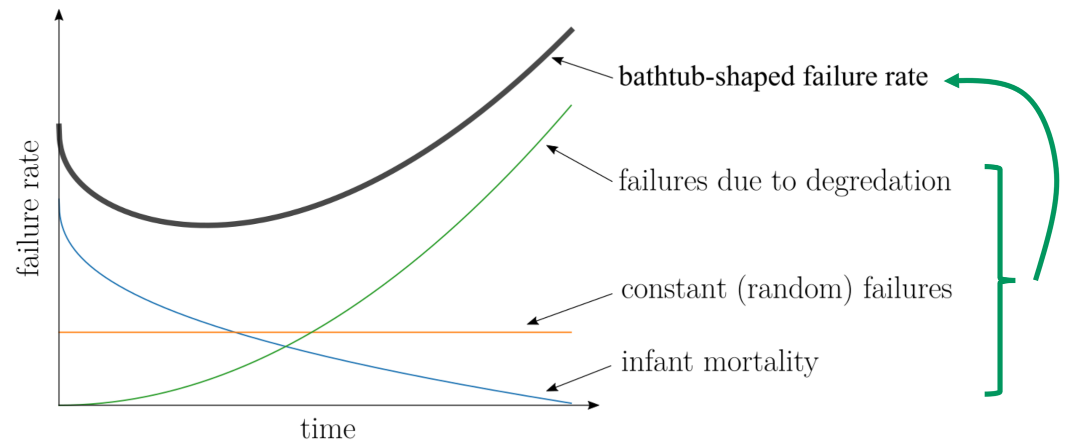


Fig. 1: Failure times of 50 electrical components [1].

- The following questions are raised:
  - How can the data be used to make predictions of future failures at the plant?
  - Which components should we replace/repair during the next maintenance shutdown?

Bathtub-shaped failure rate

- An engineering component typically has a bathtub-shaped failure rate
- Many distributions with such a failure rate have been proposed in the literature



Maximum likelihood estimate

For a complete dataset, consisting of n failure times, the likelihood is

$$\mathcal{L}(\Theta) = \prod_{i=1}^{n} f(t_i;\Theta)$$

and the log-likelihood

$$\ell(\Theta) = \sum_{i=1}^{n} \ln f(t_i;\Theta).$$

Commonly used in the literature: solve $\hat{\Theta}$ numerically from the system of equations

$$\frac{\partial \ell(\hat{\Theta})}{\partial \Theta} = 0$$

Limitation: When $\ell(\Theta)$ is non-convex, $\hat{\Theta}$ can be sub-optimal!

Log-likelihood maximization by optimization algorithms with multi-start

- Instead of the partial derivative approach, we explicitly maximize the log-likelihood
- Optimization problem

$$\max_{\Theta} \ \ell(\Theta)$$
$$\text{subject to } \ \boldsymbol{b}_{\text{lo}} \leq \Theta \leq \boldsymbol{b}_{\text{up}}$$
$$\boldsymbol{c}_{\text{lo}} < \Theta < \boldsymbol{c}_{\text{up}}$$

- Optimization algorithms used from the Scipy.optimize library
  - Nelder-Mead with augmenting parameters [6]
  - SLSQP [8]
  - L-BFGS-B [3]
- For SLSQP and L-BGFS-B, we pass the Jacobian $\frac{\partial \ell(\Theta)}{\partial \Theta}$ as an argument
- We perform a multi-start from 100 randomized starting points

Results on the EMWE distribution [12]

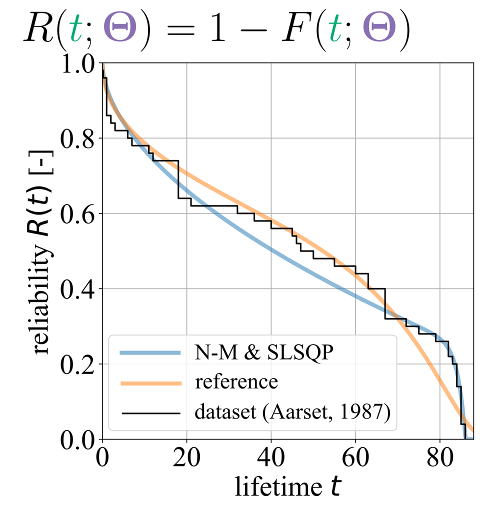
- Cumulative distribution function

$$F(\alpha,\beta,\gamma,\lambda) = [1 - \exp\{\lambda\alpha(1 - e^{(t/\alpha)^{\beta}})\}]^{\gamma} \qquad \alpha,\beta,\gamma,\lambda > 0$$

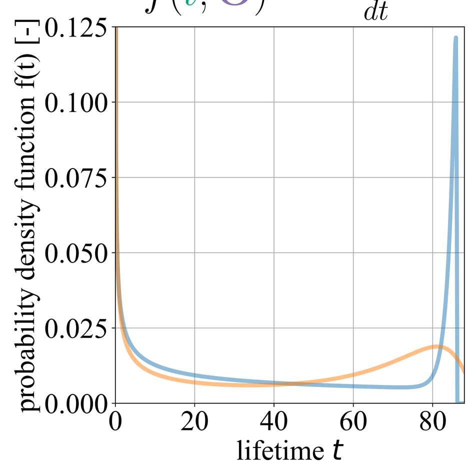
- Likelihood function

$$\ell(\Theta) = n[\alpha\lambda + (1-\beta)\ln\alpha + \ln\beta + \ln\lambda + \ln\gamma] - \alpha\lambda\sum_{i=1}^{n} e^{(t_i/\alpha)^{\beta}} + \frac{1}{\alpha^{\beta}}\sum_{i=1}^{n} t_i^{\beta}$$
$$+ (\beta-1)\sum_{i=1}^{n}\ln t_i + (\gamma-1)\sum_{i=1}^{n}\ln(1 - \exp\{\lambda\alpha(1 - e^{(t_i/\alpha)^{\beta}})\})$$

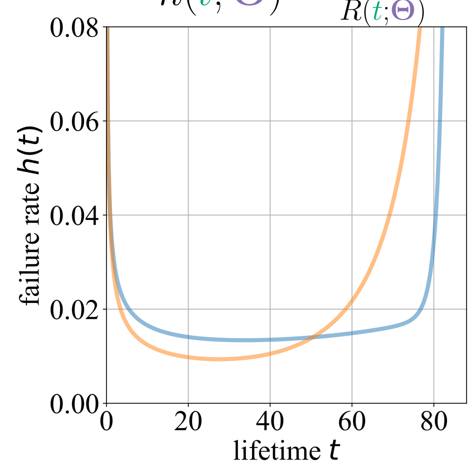
Reliability function $R(t;\Theta) = 1 - F(t;\Theta)$



Probability density function $f(t;\Theta) = \frac{dF(t;\Theta)}{dt}$



Failure rate $h(t;\Theta) = \frac{f(t;\Theta)}{R(t;\Theta)}$



- Reference parameters [12]:
  - log-likelihood $\ell(\Theta) = -213.858$
- Multi-start Nelder-Mead & SLSQP:
  - log-likelihood $\ell(\Theta) = -203.697$

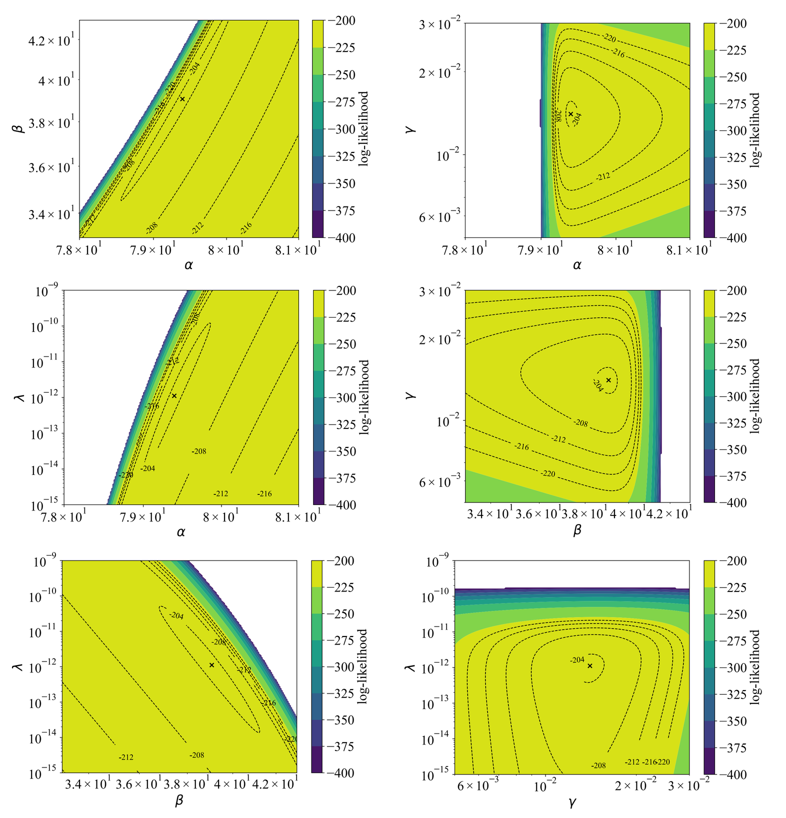


Fig. 4: The neighborhood of the optimized parameters of the EMWE distribution on the Aarset dataset [1].

Results on 10 key distributions from the literature

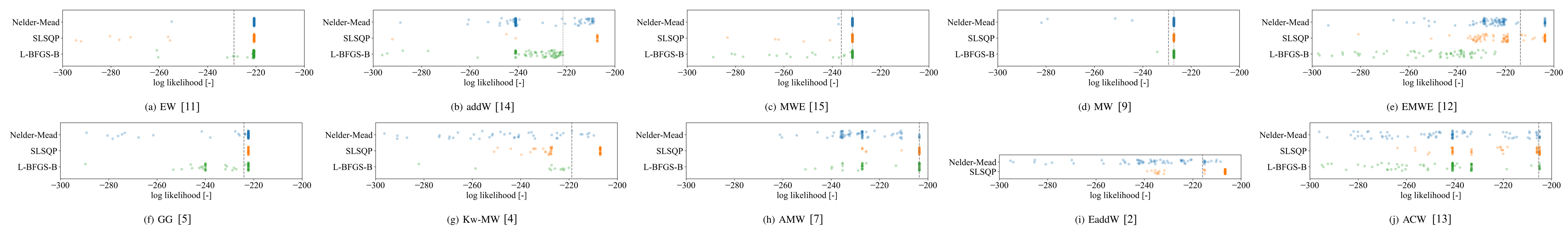


Fig. 5: Results of the multi-start optimization procedures on the Aarset dataset. Vertical lines show the log-likelihood of reference parameters reported in the literature.

References

[1] M. V. Aarset. How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, 36(1):106–108, 1987.
[2] A. E. A. Ahmad and M. G. M. Ghazal. Exponentiated additive Weibull distribution. *Reliability Engineering & System Safety*, 193:106663, 2020.
[3] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on scientific computing*, 16(5):1190–1208, 1995.
[4] G. M. Cordeiro, E. M. M. Ortega, and G. O. Silva. The Kumaraswamy modified Weibull distribution: theory and applications. *Journal of Statistical Computation and Simulation*, 84(7):1387–1411, 2014.
[5] A. El-Gohary, A. Alshamrani, and A. N. Al-Otaibi. The generalized Gompertz distribution. *Applied Mathematical Modelling*, 37(1-2):13–24, 2013.
[6] F. Gao and L. Han. Implementing the Nelder-Mead simplex algorithm with adaptive parameters. *Computational Optimization and Applications*, 51(1):259–277, 2012.
[7] B. He, W. Cui, and X. Du. An additive modified Weibull distribution. *Reliability Engineering & System Safety*, 145:28–37, 2016.
[8] D. Kraft. A software package for sequential quadratic programming. *Forschungsbericht- Deutsche Forschungs- und Versuchsanstalt fur Luft- und Raumfahrt*, 1988.
[9] C. D. Lai, M. Xie, and D. N. P. Murthy. A modified Weibull distribution. *IEEE Transactions on reliability*, 52(1):33–37, 2003.
[10] W. Q. Meeker and L. A. Escobar. *Statistical methods for reliability data*. John Wiley & Sons, 1998.
[11] G. S. Mudholkar and D. K. Srivastava. Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE transactions on reliability*, 42(2):299–302, 1993.
[12] A. M. Sarhan and J. Apaloo. Exponentiated modified Weibull extension distribution. *Reliability Engineering & System Safety*, 112:137–144, 2013.
[13] T. T. Thach and R. Briš. An additive Chen-Weibull distribution and its applications in reliability modeling. *Quality and Reliability Engineering International*, 2020.
[14] M. Xie and C. D. Lai. Reliability analysis using an additive Weibull model with bathtub-shaped failure rate function. *Reliability Engineering & System Safety*, 52(1):87–93, 1996.
[15] M. Xie, Y. Tang, and T. N. Goh. A modified Weibull extension with bathtub-shaped failure rate function. *Reliability Engineering & System Safety*, 76(3):279–285, 2002.

Acknowledgement

The financial support from Academy of Finland, through project RELOOP, is gratefully acknowledged.

ACADEMY OF FINLAND

Conclusions

- The approach yields improved fits in 14 out of 19 model-dataset pairs, for which reference parameters are available
- SLSQP was the best at finding the best-known parameters at least once in 100 optimization procedures, finding them in 17 out of 20 distribution-dataset pairs
- Proper fitting of lifetime distributions is important for
  - statisticians developing new lifetime distributions
  - engineers using the existing distributions as decision support tools in maintenance decision-making