Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

INFERRING RELIABILITY MODEL PARAMETERS FROM EXPERT OPINION

James Nutaro

Oak Ridge National Laboratory, Oak Ridge, TN, USA

ABSTRACT

We propose a method for constructing bathtub models of reliability from opinion. The method is intended for reliability studies early in the design and prototyping of a new system, before data concerning reliability has become available. A stylized bathtub curve is presented for soliciting best engineering judgement from technical experts. By pooling these stylized curves, we produce data that can be used to infer parameters for a piece-wise Weibull model of reliability. A numerical example demonstrates the practicality of the method while also highlighting potential pitfalls when working with subjective data.

1 INTRODUCTION

Reliability estimates are sometimes needed early in the design of a new system, before data is available concerning the reliability of the system or its components. There may be several motivations for these early reliability studies. For instance, to create designs that will satisfy *a priori* requirements for expected lifetime or to estimate economic losses due to product replacements.

Recently proposed methods for estimating reliability during early design cite expert opinion and engineering experience as potential sources of information; see, e.g., (Cheng 2017; Rajchel 2020). Both of these methods use information from expert opinion that is cast into a form suitable for Bayesian inference. That is, opinion is supplied as a probability or as parameters for a probability distribution.

We present a method for obtaining parameters for a probability distribution from opinion originating in *best engineering judgement*, derived from a group of engineers' prior experiences and informed intuitions. (I attribute the term *best engineering judgement* to Mike Kamrowski at Raytheon Missile Systems, who introduced the phrase to me when I was a new engineer in the mid 1990s.) To gather best engineering judgement in a form useful for analysis, the reliability engineer must articulate reliability questions to design engineers, technicians, and other technical professionals who may have little or no background in probabilistic models of system reliability. The stylized bathtub curve that we propose is intended to facilitate these conversations.

The stylized bathtub curve shown in Figure 1. This curve is motivated by polynomial models of the bathtub curve (Kogan 1988). Our curve builds on this concept by using the simplest possible polynomials consistent with a bathtub shape and combining them in a piece-wise fashion following the ideas of Peng, Liu, and Wang (2016).

The crucial difference here is that our stylized curve is not intended as a model of failure rates through time; it is not the reliability bathtub curve. Rather, this simple representation is intended to be easily understood and parameterized by technical professionals on the basis of their experience and judgement. Experience suggests that probability and duration are more easily estimated than rates. Indeed, introductory courses on statistics expose all engineers to probability, but hazard functions and other derived quantities are unlikely to be familiar without further study.

To obtain a useful reliability model, we present a pooling approach that uses data generated by simulations of the stylized curve to fit common models of probability. Specifically, we show that a piece-wise Weibull random variable can be closely fit to the data obtained by our method. Once the Weibull parameters have been found, the model thus obtained can be used for reliability analysis.

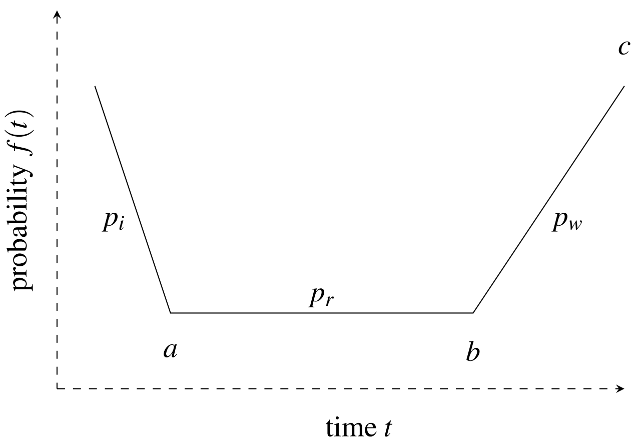

Figure 1: Elements of the stylized bathtub curve.

2 BACKGROUND AND RELATED WORK

The intuition behind our approach is the often observed and much studied wisdom of crowds (Larrick, Mannes, and Soll 2012; Wang, Hyndman, Li, and Kang 2022). Decades of research in a variety of fields has shown the benefits of combining information from knowledgeable people. The method of combination that we use here equally weights the opinions of our experts. This approach is simple and has been shown to be effective in many circumstances (see, e.g., the survey in (Larrick, Mannes, and Soll 2012)).

However, recent work by Cai, Lin, Han, Liu, and Zhang (2017) and Ji, Lu, and Zhang (2021) on the problem of combining probability distributions provided by experts offers alternative, but more complex, methods for combination that could be explored in future work. Alternatively, opinion pooling could be viewed as a simplistic form of consensus building. A method of automating consensus building, proposed by Pérez, Cabrerizo, Alonso, and Herrera-Viedma (2014), could offer another promising direction for future research.

Reliability is often concerned with low probability events. Experimental work underlying prospect theory has found that events with very low probabilities are generally over weighted (Kahneman and Tversky 1979); the perceived probability of such an event is greater than the actual probability. If this bias is consistent across the population of opinions, then it creates systemic errors in information derived from estimates offered by that population. This bias has been observed in betting markets, wherein the odds of the bookmaker reflect the wisdom of the betting public (Nutaro 2023). Our experimental data (see Sect. 6) suggests that this bias also exists in reliability estimates.

Finally, the cost of gathering useful opinions is particularly relevant in engineering applications. This problem appears explicitly in a survey of elicitation studies described in (Soares, Sharples, Morton, Claxton, and Bojke 2018). The summary of solicitation methods reviewed in (O'Hagan, Buck, Daneshkhah, Eiser, Garthwaite, Jenkinson, Oakley, and Rakow 2006) (Chapter 6) offers several approaches for obtaining an expert opinion, and the relative merits of each in our context would be another interesting topic for future work.

Experience with early prototypes may provide sparse reliability data. These data offer an opportunity to refine opinions, particularly when an initial choice of parameters is clearly inconsistent with new data. When reliability data becomes plentiful, we anticipate discarding the model built from opinion in favor of a model suitable for the analysis of, more or less plentiful, data. Bathtub models designed for this purpose have been a topic of research for half a century, and we will not discuss them further here; for example, see models by Leithead (1970), Hjorth (1980) from the prior century, very recent models proposed by Arshad, Iqbal, and Mutairi (2021), He, Cui, and Du (2016), and the brief survey by Yong (2004).

3 PROPERTIES OF THE STYLIZED BATHTUB

Three key points define failure probabilities in time for our piece-wise linear model. The point a is when early failures no longer occur and we pass into the random failure regime at the bottom of the bathtub curve. The point b is the end of random failures and the earliest time we expect to see end of life due to wear out. The point c is the time beyond which we cannot reasonably expect the system to operate.

Two values describe the frequency of failures in each regime. The fraction p_i is the likelihood of an early failure or, equivalently, the fraction of the population expected to suffer early failure. The fraction p_r is the likelihood of a random failure. The likelihood of a wear out failure is $p_w = 1 - p_i - p_r$.

These five numbers - three in time and two fractions - define our approximation to a bathtub curve. Estimates of these parameters might be solicited with five questions:

1. What fraction of systems do you expect to fail early, from causes related to manufacturing, material quality, and so forth? This establishes p_i.
2. How long does the system need to operate before you are confident that the causes of early failure are not applicable? This establishes a.
3. What is the intended life of the system? This establishes b.
4. What fraction of systems will fail before reaching the intended end of life? This establishes p_r.
5. What is the longest time for the system to operate that you feel is credible? This establishes c.

The stylized bathtub defines a random variable with density function f. The cumulative probability for the early failure region of the curve must equal p_i to ensure that fraction of failures occur in this region. Over the interval from 0 to a, the probability density declines from $f(0)$ to $f(a)$. Hence, $f(t)$ in this section is the line

$$f(t) = f(0) - \frac{f(0) - f(a)}{a} t \ , \ 0 \leq t \leq a \ . \tag{1}$$

The fractional number of failures, and so the probability of failure, p_i is the area under this region, which is obtained by integrating (1) from $t = 0$ to a yielding

$$p_i = \frac{a}{2}(f(0) + f(a)) \ . \tag{2}$$

The bottom of the bathtub curve has a constant probability density

$$f(t) = f(a) = f(b) \ , \ a \leq t \leq b \tag{3}$$

and integrating from a to b gives

$$p_r = f(a)(b - a) = f(b)(b - a) \ . \tag{4}$$

The probability density in the wear out region is

$$f(t) = \frac{f(c) - f(b)}{c - b}(t - b) + f(b) \ , \ b \leq t \leq c \tag{5}$$

and integrating from b to c we obtain

$$p_w = \frac{c - b}{2}(f(b) + f(c)) \ . \tag{6}$$

These results for p_i, p_r, and p_w are obvious from the geometric shape of each region.

Deriving expressions for the cumulative distribution function and for sampling the random variable in simulations is simplified by working with conditional probabilities. Proceeding in this way, let us first assume that the failure occurs in the early region. With this information, $p_i = 1$ and $p_w = p_r = 0$. From (4) we find $f(a) = 0$ and from (2) that $f(0) = 2/a$. Substituting these values into (1) and integrating with respect to time we find the cumulative distribution to be

$$F_i(t) = \int_0^t \frac{2}{a} - \frac{2}{a^2}\tau \, d\tau = \frac{t}{a}\left(2 - \frac{t}{a}\right) . \tag{7}$$

To sample this function in a simulation, we draw a uniform random number u between zero and one and calculate

$$t = F_i^{-1}(u) = a(1 - \sqrt{1-u}) \tag{8}$$

which is obtained from (7) by taking the root of the quadratic that ensures a between zero and one when $F_i(t) = u$ is given.

Reasoning in the same way, let us assume that the failure occurs in the wear out region. In this case, $p_w = 1$, $p_i = p_r = 0$, $f(b) = 0$, and from (6) we have $f(c) = 2/(c-b)$. Substituting these values into (5) and integrating gives the cumulative distribution

$$F_w(t) = \int_b^t \frac{2}{(c-b)^2}(\tau - b) \, d\tau = \frac{t^2 - b^2 - 2b(t-b)}{(c-b)^2} . \tag{9}$$

Given u in zero to one and solving the quadratic for t gives us

$$t = F_w^{-1}(u) = b + (c-b)\sqrt{u} . \tag{10}$$

Lastly, we assume $p_r = 1$. In this case, (4) and (3) give $f(t) = 1/(b-a)$. Integrating we find

$$F_r(t) = \int_a^t \frac{1}{b-a} \, d\tau = \frac{t-a}{b-a} \tag{11}$$

and

$$t = F_r^{-1}(u) = a + u(b-a) . \tag{12}$$

A simulation samples the complete curve by drawing a pair of numbers u and s uniformly distributed in zero to one and calculating

$$t = \begin{cases} a(1-\sqrt{1-u}) & s < p_i \\ a + u(b-a) & p_i \leq s \leq p_w \\ b + (c-b)\sqrt{u} & s > p_w \end{cases} \tag{13}$$

4 RATIONAL PARAMETER SELECTIONS

It is possible for an opinion to provide parameters that are incompatible with a bathtub shape. When these opinions are found, they should be discarded as irrational in relation to the assumption of a bathtub curve or be revised to be consistent with the bathtub shape.

The bathtub curve has its particular shape because we expect failure density in the wear out and early regions to exceed the failure density in the random failure region. That is, we expect

$$f(t) > f(a) \text{ , } t < a \text{ and} \tag{14}$$

$$f(t) > f(b) \text{ , } t > b . \tag{15}$$

For the early failure region, (1) shows it is sufficient for

$$f(0) > f(a) \tag{16}$$

and, rearranging (2) and (4) to get $f(0)$ and $f(a)$, we find

$$\frac{p_i}{a} > \frac{p_r}{b-a} \ . \tag{17}$$

In some analyses, the early failure region is ignored under the assumption that machines subject to early failure are found and removed prior to use. If this assumption is made, then we do not need to be concerned with equality (17). Otherwise, it must be satisfied for the model to behave rationally.

In the wear out region, (5) requires that

$$f(c) > f(b) \ . \tag{18}$$

Rearranging (4) and (6) to obtain $f(b)$ and $f(c)$, we find

$$\frac{p_w}{c-b} > \frac{p_r}{b-a} \ . \tag{19}$$

This constraint should always be satisfied when choosing parameters for the model.

5 A SIMULATION APPROACH TO OPINION POOLING

We now turn to the problem of synthesizing several opinions. It is a familiar experience that upon asking "when will you arrive" we receive the point estimate "at 6" or "around 6", or the bounded estimate "after 5 but no later than 7", or possibly some variant of the latter, say, "between 5 and 7". Accordingly, we expect a natural response to our questions in the form of an interval or point (i.e., interval of zero length) for each model parameter. We assume that each value in that interval is equally likely.

In this way, each opinion j defines a five dimensional volume V_j in the parameter space. A point $v_j \in V_j$ is constructed by selecting a single value along each dimension: p_i, p_r, a, b, and c. Given opinions V_1, V_2, ..., V_n a simulation based pooling of these opinions is accomplished by random sampling as follows.

1. Select, at random, the opinion V_j we will use and a point $v_j \in V_j$.
2. Is v_j a rational choice as described in Sect. 4? If so, continue to the next step. If not, return to the previous step.
3. Using (13), sample the time to fail given the model parameters v_j. Record this value.
4. Repeat from step 1 until sufficiently many samples have been recorded.

Using the data produced by this simulation, we may calculate the fraction of the population surviving up to a given time t as the fraction of data points greater than t. Mean and variance of the collected data may also be calculated.

6 NUMERICAL EXAMPLE

To illustrate the proposed model, data from Backblaze concerning hard drive reliability (Backblaze 2022) was compared with a pooled model of opinions solicited from scientists and engineers at Oak Ridge National Laboratory. The data from Backblaze contains daily snapshots of disk drives operating in the company's data center. The snapshots include, among other information, the S.M.A.R.T. statistics reported by each drive.

The topic of hard drives was chosen assuming that nearly everyone has experience with this technology and will have formed an intuitive understanding of a disk drive's useful lifetime. The solicitations were

Table 1: Parameters obtained by interviews. The horizontal lines group opinions that were gathered together. The bold entries offer irrational choices for the early failure region. All parameter selections are rational choices for the wear out region.

a	b	c	p_i	p_r
30-90 days	5-7 years	10 years	0.05	0.1
30-90 days	5-7 years	10 years	0.01	0.1
2 years	**6-7 years**	**9 years**	**0.01**	**0.19**
2 years	**8 years**	**9 years**	**0.03**	**0.17**
14 days	4 years	10 years	0.02	0.05
2-3 months	10 years	20 years	0.01	0.001
1 month	**5 years**	**8 years**	**0.0001**	**0.05**

done in person, beginning with a brief review of the bathtub curve in Figure 1 and then a request for parameters describing each region.

The parameters obtained in this way are shown in Table 1. The first two entries in the table come from persons interviewed together, the second two were interviewed together, and the other entries were obtained in separate interviews. Figure 2 shows the complete, cumulative distribution function $F(t)$ produced by this model.

In generating this data, we attempted 500,000 trials and discarded 43% of our selections as irrational (see Sect. 4). If early failures are removed then no trials are discarded. The three regions of early failure, random failure, and wear out are clearly apparent. The wear out region has a long, shallowly sloped tail, which is due to the single 20 year estimate for c in one of the opinions. Otherwise, the pooled observations produce a curve with the expected shape.

It will quickly become apparent that the best reliability estimates are for older drives when early failures are discounted. We offer possible explanations for these results in Section 8. Briefly, poor estimation of early failure by our interview subjects is anticipated by prospect theory. The good for fit for older drives and poorer fit for newer drives matches the experience of our interview subjects. Both explanations highlight the challenge of working with subjective rather than objective data.

Figure 3 compares the model to failure curves produced from data collected by Backblaze in the indicated years. We have selected the years for which a complete year of data was available at the time of analysis. The data of interest are the hours on power (SMART field 9 raw value) reported by the drive and, if the drive was seen to fail, the hours on power reported the day prior to the failure.

We discard data from any drive that does not report hours on power or that continues to appear in the data after reporting a failure. In the simulation, every drive is observed to fail. In the collected data, drives may be retired prior to failure and we use the Kaplan-Meier method to account for these non-events when constructing the failure curve (see (Kaplan and Meier 1958) or any readily available tutorial).

The reliability of hard drives improved year on year. Operating times past the intervals shown are very sparse in the available data and have been omitted from the plots. The unusual bump in failures seen at years two to three in the 2014 data is surprising and merits further consideration. In other respects, the data produce failure curves with the expected shape.

The opinions of our interview subjects anticipate a larger number of early failures than is seen in the data. However, the width and rates of failures in the random failure and wear our regions approximate what was observed historically in 2013-2017. Agreement is worse for more recent data, which reflects newer, and seemingly more reliable, drive technology.

To more clearly see how the rate of failure in the random and wear our regimes compare with the data, we set $p_i = 0$ for every opinion in Table 1 and reran the simulations. Figure 4 compares the failure curve produced by this model without early failures to the historical failure data.

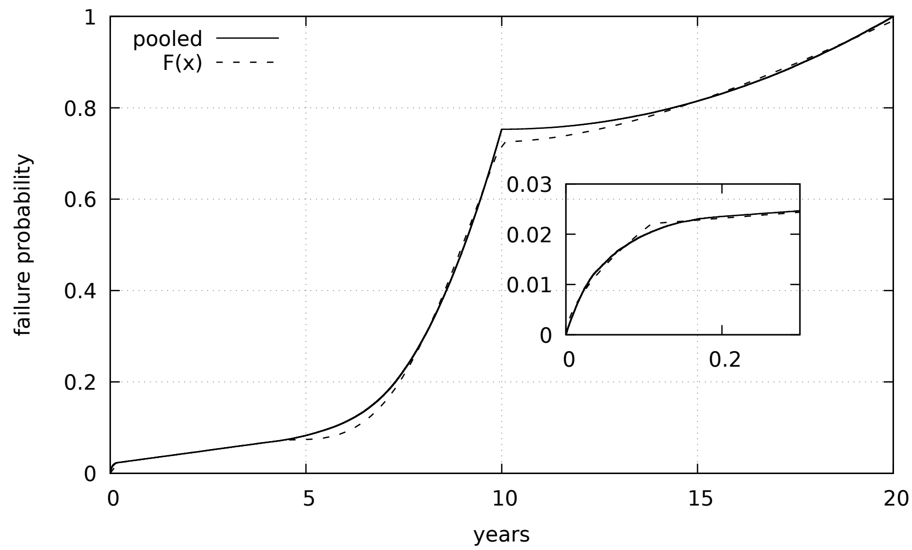

Figure 2: The failure curve produced by pooling simulations (solid line) and the fit of a piece-wise Weibull $F(x)$ given by Eqn. 20. The inset details the transition from early failures to random failures in the first few months (at about 0.2 years) of the drive lifetime. The transition to wear out occurs in years five to six.

For the data collected from 2013-2017, the model anticipates a failure rate within the range given by the data. However, our analysis clearly favors the higher rates of failure exhibited in the years 2013-2017. The data collected from 2018-2021 show much reduced rates of failure, and our analysis overstates the random and wear out failure rates with respect to those data.

7 PIECE-WISE WEIBULL FIT

The pooled failure function has a surprising semblance to a standard bathtub curve. There is a region of linear increase - the random failure region - where the straight line indicates a constant failure rate. The smoothly decreasing failure rates in the early failure region and increasing rates in the wear out region are also apparent. Indeed, the pooled curve is unlike the stylized curves from which it originates in these important aspects.

A good fit to a piece-wise Weibull function can be expected because in each piece-wise section we are considering only a finite span of time. For instance, in the region of random failure, the pooled simulations distribute some fraction p of failures uniformly over a span of length T. Looking at the data in this span, we perceive a constant failure rate of p/T. This can be fit closely to an exponential model over the same finite span. Similar arguments apply to the early and wear out regions where samples over the finite span will be denser at the start and ends, respectively, yielding the perception of declining or climbing failure rates.

We fit a four piece Weibull distribution to the curve produced by pooling when early failures are included. The form of cumulative distribution for this piece-wise probability function is

$$F(t) = \begin{cases} 1 - e^{-\lambda_i t^{g_i}} & t \leq t_i \\ F(t_i) + 1 - e^{-\lambda_r (t - t_i)} & t_i < t \leq t_r \\ F(t_r) + 1 - e^{-\lambda_w (t - t_r)^{g_w}} & t_r < t \leq t_f \\ F(t_w) + 1 - e^{-\lambda_f (t - t_w)^{g_f}} & t > t_w \end{cases} \tag{20}$$

The change points were selected by trial and error to produce a good fit. However, an automated fit could be obtained with a suitable heuristic optimization method (e.g., simulated annealing, a genetic

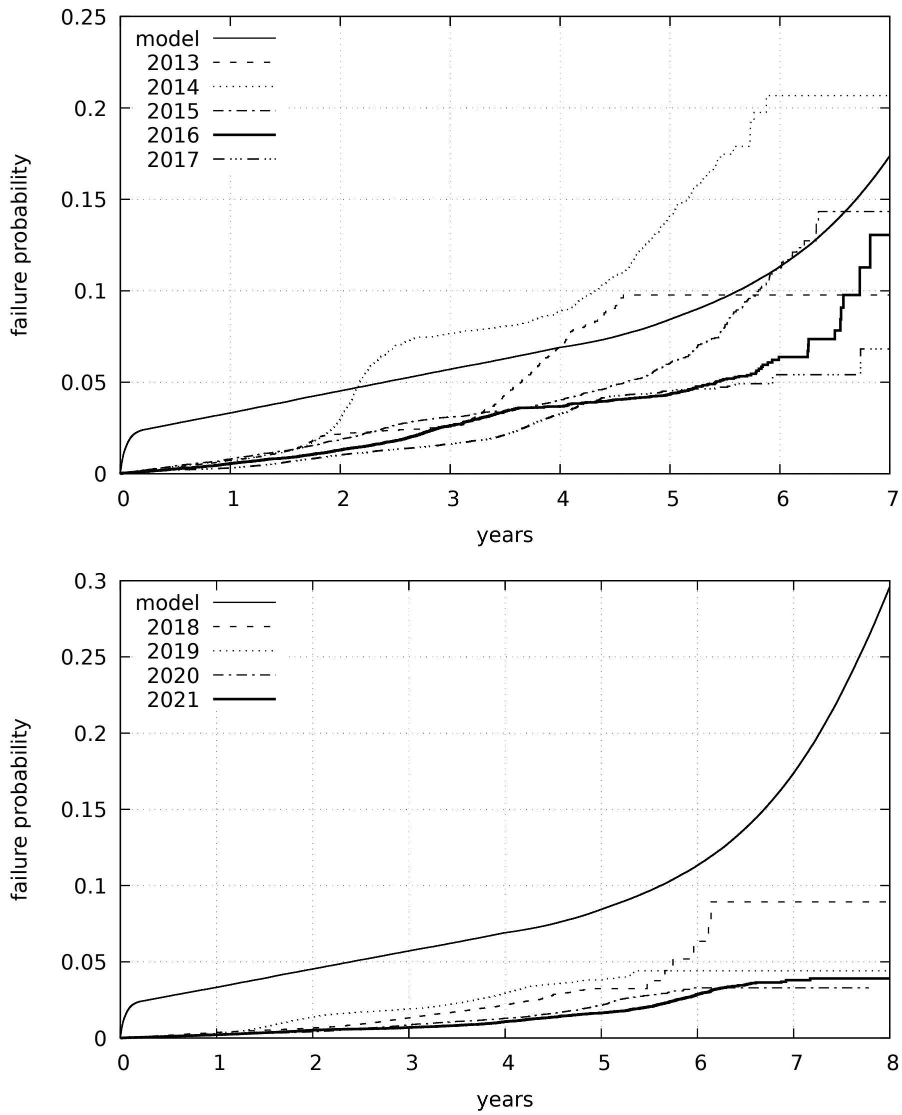

Figure 3: Comparison of model failure curve and Kaplan-Meier failure curves for data collected from 2013 to 2021.

algorithm, or something similar). A gradient descent method was used to find parameters in each section.

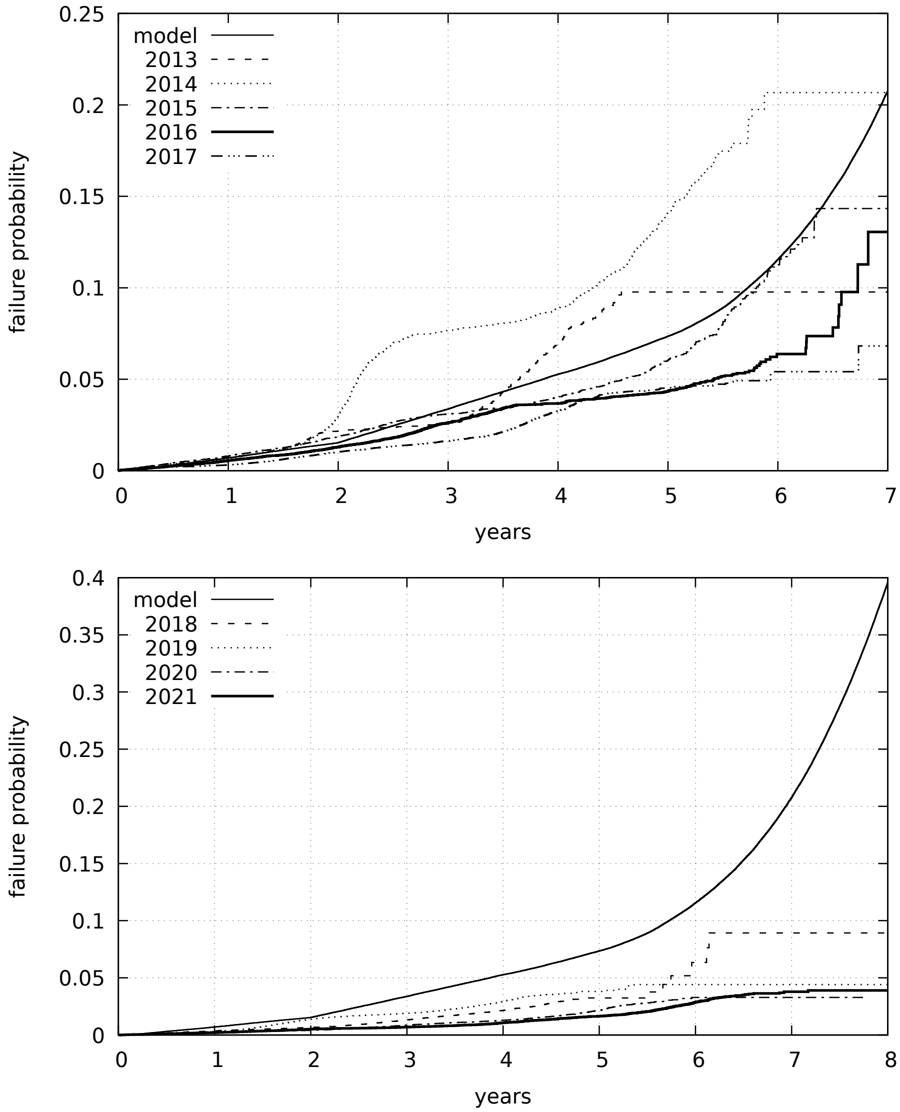

Figure 4: Comparison of model failure curve and Kaplan-Meier failure curves for data collected from 2013 to 2021 when early failures are removed from the model.

The fit curve is shown alongside the data in Figure 2. The parameters that produce this fit are

$$t_i = 0.110$$
$$\lambda_i = 0.0857$$
$$g_i = 0.609$$
$$t_r = 4.50$$
$$\lambda_r = 0.0120$$
$$t_w = 10.0$$
$$\lambda_w = 0.00454$$
$$g_w = 3.20$$
$$t_f = 20.0$$
$$\lambda_f = 0.00586$$
$$g_f = 1.72$$

8 CONCLUSIONS

In the numerical example, pooling best engineering judgement produces rates of failure in the random and wear our regimes that are near those seen in practice. However, there is a consistent tendency to over estimate the rate of early failures. At the same time, early failures appear to be quite rare, and we expect that few of our interview subjects have experienced the early failure of a disk drive (however, this was not one of our questions).

Estimates of useful lifetime and wear out appear to better reflect older drive technologies. The survey of opinion was taken in 2022. If we assume that a six year life is typical for a hard drive (this, at least, appears to be the approximate point where we see an accelerating rate of failure), then experience with drives produced prior to 2016 will have shaped the opinions reflected in our interviews. This is consistent with the improved agreement of model and data for the years 2013 to 2015.

These results suggest that the wisdom of crowds may be a useful tool for answering reliability questions early in the design of a system, before reliability data becomes available through testing and operating experience. However, the data presented here also suggests that known problems with estimating low probability events (in this case, the early failure of drives) and statistical outliers in an opinion pool could pose fundamental limits to the approach.

REFERENCES

Arshad, M. Z., M. Z. Iqbal, and A. A. Mutairi. 2021, May. "Statistical Properties of a New Bathtub Shaped Failure Rate Model With Applications in Survival and Failure Rate Data". *International Journal of Statistics and Probability* 10(3):49–68 https://doi.org/https://doi.org/10.5539/ijsp.v10n3p49.

Backblaze 2022, July. "Hard Drive Data and Stats". https://www.backblaze.com/b2/hard-drive-test-data.html.

Cai, M., Y. Lin, B. Han, C. Liu and W. Zhang. 2017. "On a Simple and Efficient Approach to Probability Distribution Function Aggregation". *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 47(9):2444–2453 https://doi.org/10.1109/TSMC.2016.2531647.

Cheng, Y. 2017. *Reliability prediction in early design stages*. Ph. D. thesis, Missouri University of Science and Technology https://doi.org/https://scholarsmine.mst.edu/doctoral_dissertations/2593.

He, B., W. Cui, and X. Du. 2016. "An additive modified Weibull distribution". *Reliability Engineering & System Safety* 145:28–37 https://doi.org/https://doi.org/10.1016/j.ress.2015.08.010.

Hjorth, U. 1980. "A Reliability Distribution with Increasing, Decreasing, Constant and Bathtub-Shaped Failure Rates". *Technometrics* 22(1):99–107 https://doi.org/https://doi.org/10.2307/1268388.

Ji, C., X. Lu, and W. Zhang. 2021. "A Biobjective Optimization Model for Expert Opinions Aggregation and Its Application in Group Decision Making". *IEEE Systems Journal* 15(2):2834–2844 https://doi.org/10.1109/JSYST.2020.3027716.

Kahneman, D. and A. Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk". *Econometrica* 47(2):263–291 https://doi.org/https://doi.org/10.2307/1914185.

Kaplan, E. L. and P. Meier. 1958, June. "Nonparametric Estimation from Incomplete Observations". *Journal of the American Statistical Association* 53(282):457–481 https://doi.org/https://doi.org/10.2307/2281868.

Kogan, V. I. 1988. "Polynomial Models of Generalized Bathtub Curves and Related Moments of the Order Statistics". *SIAM Journal on Applied Mathematics* 48(2):416–424 https://doi.org/https://doi.org/10.1137/0148022.

Larrick, R. P., A. Mannes, and J. Soll. 2012. "The social psychology of the wisdom of crowds". In *Social judgment and decision making*, 227–242. Psychology Press https://doi.org/https://doi.org/10.4324/9780203854150.

Leithead, G. S. 1970, May. "Model for Bathtub-Shaped Hazard Rate: Monte Carlo Study". Master's thesis, Utah State University, Logan, Utah https://doi.org/https://doi.org/10.26076/fb8a-c323.

Nutaro, J. 2023, May. "Prospect Theory and the Favorite Long-Shot Bias in Baseball". *Risks* 11(5):1–11 https://doi.org/https://doi.org/10.3390/risks11050095.

O'Hagan, A., C. E. Buck, A. Daneshkhah, J. R. Eiser, P. H. Garthwaite, D. J. Jenkinson, *et al*. 2006. *Uncertain Judgements: Eliciting Experts' Probabilities*. John Wiley & Sons, Ltd https://doi.org/https://doi.org/10.1002/0470033312.

Peng, C., G. Liu, and L. Wang. 2016. "Piecewise modelling and parameter estimation of repairable system failure rate". *Springerplus* 5 https://doi.org/https://doi.org/10.1186%2Fs40064-016-3122-4.

Pérez, I. J., F. J. Cabrerizo, S. Alonso, and E. Herrera-Viedma. 2014. "A New Consensus Model for Group Decision Making Problems With Non-Homogeneous Experts". *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 44(4):494–498 https://doi.org/10.1109/TSMC.2013.2259155.

Rajchel, E. T. 2020. “The conceptual design reliability prediction method: establishing funcational-physical reliability relationships for system reliability predictions during conceptual design”. Master’s thesis, Naval Postgraduate School https://doi.org/https://apps.dtic.mil/sti/citations/AD1126560.

Soares, M. O., L. Sharples, A. Morton, K. Claxton and L. Bojke. 2018. “Experiences of Structured Elicitation for Model-Based Cost-Effectiveness Analyses”. *Value in Health* 21(6):715–723 https://doi.org/https://doi.org/10.1016/j.jval.2018.01.019.

Wang, X., R. J. Hyndman, F. Li, and Y. Kang. 2022. “Forecast combinations: An over 50-year review”. *International Journal of Forecasting* 39(4):1518–1547 https://doi.org/https://doi.org/10.1016/j.ijforecast.2022.11.005.

Yong, T. 2004, September. *Extended Weibull distributions in reliability engineering*. Ph. D. thesis, National University of Singapore, Singapore https://doi.org/https://scholarbank.nus.edu.sg/handle/10635/14186.

ACKNOWLEDGEMENTS

This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

AUTHOR BIOGRAPHIES

JAMES NUTARO is a Distinguished Research & Development staff member in the Computational Sciences & Engineering directorate at Oak Ridge National Laboratory. His research interests include model methods, discrete event and hybrid dynamic systems, simulation algorithms, and systems & software engineering. He serves as an associated editor for the ACM Transactions on Modeling and Computer Simulator and the journal SIMULATION. His email address is nutarojj@ornl.gov.