Accelerated Tests as Hypothesis Tests for Reliability using the Probability of Test Success

Alexander Grundler[a*], **Martin Dazer**[b]
[a]Robert Bosch GmbH, Schwieberdingen, Germany
[b]Institute of Machine Components, University of Stuttgart, Stuttgart, Germany

Abstract:
This paper is concerned with the planning of accelerated reliability demonstration tests. The approach presented herein uses the hypothesis testing context of the Probability of Test Success for assessment of test configurations. The framework yields advantages in the planning of the tests and allows for an effortless selection of the correct test parameters e.g. load level height and number of load levels, sample sizes as well as a statistically sound test planning. Due to the integration of the Probability of Test Success in accelerated reliability demonstration tests as the statistical power of the test, the most efficient one can be selected which still holds a high probability of demonstrating the required reliability of the product.

Keywords: Probability of Test Success, Reliability Demonstration, Statistical Power, Test Planning.

1. INTRODUCTION

Products shall be developed in such a way that they are able to endure the loads they are subjected to during their use. This requirement is placed by customers, governments, and regulations as well as the company itself. In order to make sure such a reliability requirement is met, physical reliability demonstration tests are essential prior to market entry [1]. The engineers are then faced with the challenge of choosing the best suitable test in order to determine the actual reliability of the product [2]. The greater the sample size in the test, the more accurate the information gets. However, this is in contrast to the possibilities a company has in order to be successful on the market, since the budget should be kept as small as possible. To overcome this challenge, accelerated reliability demonstration tests can be used [3]–[5]. They allow for a shorter test time by increasing the load. To conduct such accelerated reliability demonstration tests, the test has to be planned in such a way to achieve a successful reliability demonstration in the shortest time possible and at the same time using the smallest sample size possible, while not exceeding the available budget. For this, several parameters must be chosen. For example, the number and height of load level needs to be defined, as well as the share of specimens amongst those [6].

Existing approaches usually deal with the distribution of specimens in terms of load level height, number of load levels as well as sample size in order to get a best possible estimate of the parameters of the lifetime model [7]. For example placing the specimen on only a few load levels is advised to do if the load limits are known [4]. On the other hand, if the variance of the parameter estimation is taken into account, several optimization criteria can be found, such as A-, D- or I-Optimality [8], [9]. However, they are not focusing on the reliability demonstration itself but rather on the lifetime model such plans are trying to estimate. In order to get an estimate about the probability of successfully demonstrating the reliability requirement the Probability of Test Success was introduced [10]–[12]. It can derive a required sample size to achieve the desired probability of a successful reliability demonstration. Even for failure-based tests and accelerated reliability demonstration tests required sample sizes can be calculated. This concept has been studied for non-accelerated tests [13]–[15]. For this, the effects of uncertainty [15], [16], interval censoring [14] as well as the additional use of prior knowledge by means of Bayes' theorem [5], [16], [17] have been studied. Also for demonstration of system reliability, analyses have been conducted [12], [18]–[20]. Herzig et al. showed that the Probability of Test Success is suited for an advanced planning of accelerated tests in terms of overall expenditure and a tradeoff between accuracy and resources [21], [22]. This approach also coincides with the optimal distribution of specimen between higher and lower load level according to Nelson [8], [9], but is able the take the assessment even further and take the actual reliability demonstration into account and thus allows for the identification of an optimal test. In addition, Benz et al. demonstrated the use of common load profiles in this concept and compared the most common strategies for placing the specimen on the load levels [23], [24]. However, none of the existing research makes use of the recent development of considering the Probability of Test Success as the statistical power of a reliability demonstration test, as introduced in [11], [18], [25]. This context helps in developing easy to implement algorithms and procedures for the identification of the optimal test in the individual case [11]. Recent studies show the applicability of this method to failure-free tests, as well as failure-based tests for systems with multiple failure mechanisms [12], [18], [20], [26]. Furthermore, the additional use of Bayes' theorem was introduced to cope with the usually high effort involved in non-accelerated tests [12], [16]. Although first studies for accelerated tests have been made [13], [21], [22], [27], the hypothesis testing framework still needs to be established for those type of tests.

2. PROBABILITY OF TEST SUCCESS: HYPOTHESIS TESTING FOR RELIABILITY

The development of a product has to ensure the required functionality is provided. This has to be validated by physical tests. Since the fulfillment of the requirements cannot be assumed without observing them, a hypothesis about the reliability requirement can be formulated. To reject or confirm such hypothesis, the reliability demonstration test is conducted. The reliability demonstration test can thus be regarded as a hypothesis test for reliability [11], [12]. The reliability requirement is defined by a required life t_{req} at required reliability R_{req}. Since a test can only reject a hypothesis about the absence of a phenomenon under investigation [28], the null hypothesis H_0 represents the non-fulfillment of the reliability requirement. The conducted test shall gather information so that the null hypothesis can be rejected. Since the statistical power of a test corresponds to the discovery of the alternative to the null hypothesis, the alternative hypothesis H_1 represents the fulfillment of the reliability requirement. Thus the hypothesis for a reliability demonstration test are [11], [12]:

$$H_0\colon\ t_R < t_{\mathrm{req}} \tag{1}$$

$$H_1\colon\ t_R \geq t_{\mathrm{req}} \tag{2}$$

With t_R being the estimated quantile of the lifetime of the product at required reliability. The required confidence C_{req} of a reliability demonstration test corresponds to the probability of rejecting the null hypothesis although it is actually true, i.e., the product does actually not achieve the reliability requirement. This is the type I statistical error [29]. The type II error, however, describes the probability of falsely accepting the null hypothesis. Its complement, the statistical power of a test, is the probability of correctly rejecting the null hypothesis. By using the hypothesis of eq. 1 and 2, the Probability of Test Success P_{ts} is the statistical power of the reliability demonstration test and equals the probability of a successful demonstration of the requirements [11], [12]. The P_{ts} can only be provided for a certain failure distribution. If the product endures the loads well during operation, high lifetimes are to be expected and therefore also greater failure times. These would result in an estimated failure distribution in the reliability test, which is more likely to fulfill the reliability requirement. For a test planning using the P_{ts}, prior knowledge about the expected failure behavior of the product needs to be available. If none can be estimated properly, a parameter study can be conducted instead. To describe the over fulfillment of the reliability requirement, the safety distance s was introduced [10]:

$$s = 1 - \frac{t_{\mathrm{req}}}{t_{\mathrm{p}}} \tag{3}$$

With t_{p} being the lifetime quantile at required reliability given by prior knowledge. If the product does exactly meet the requirement, the safety distance becomes zero. The distribution of the lifetime quantile under validity of the null hypothesis f_{H_0} is called null distribution and its location is t_{req} for the limit case of $\lim_{s\to 0-} t_{\mathrm{p}} = t_{\mathrm{req}}$. The alternative distribution under validity of the alternative hypothesis f_{H_1} has location of t_{p} for the case of $s = 1$. To calculate the Probability of Test Success P_{ts}, the following integrals need to be evaluated [11], [12], [15]:

$$P_{ts} = \int_{t_{\mathrm{crit}}}^{\infty} f_{H_1}(t_R)\, \mathrm{d}t_R \tag{4}$$

$$C = \int_{0}^{t_{\mathrm{crit}}} f_{H_0}(t_R)\, \mathrm{d}t_R \tag{5}$$

The value t_{crit} needs to be calculated according to C_{req}. To estimate the distributions f_{H_0} and f_{H_1} two methods are proposed in [11], [12]. The first one being a bootstrap algorithm, which samples failure times from prior knowledge according to the test sampling scheme. By estimating the failure distribution and calculating the lifetime quantile in each iteration, the integrals can be calculated similar to a percentile bootstrap confidence bound [30], [31]. The failure distribution used for f_{H_1} is the one from prior knowledge as far as $s > 0$ and the one used for f_{H_0} is shifted, so that $s = 0$ as in the limit case. The Second method is an analytic one, which is

based on the asymptotic properties of the maximum likelihood estimator and the central limit theorem [11], similar to Fisher confidence bounds [1].
By making use of the variance-covariance matrix V as the inverse of the Fisher Information matrix I [32]

$$V = \begin{bmatrix} \mathrm{Var}(P_1) & \mathrm{Cov}(P_1, P_2) & \dots & \mathrm{Cov}(P_1, P_K) \\ & \mathrm{Var}(P_2) & \dots & \mathrm{Cov}(P_2, P_K) \\ & & \ddots & \vdots \\ \mathrm{sym} & & & \mathrm{Var}(P_K) \end{bmatrix} = I^{-1} = \begin{bmatrix} -\frac{\partial^2 \Lambda}{\partial P_1^2} & -\frac{\partial^2 \Lambda}{\partial P_1 \partial P_2} & \dots & -\frac{\partial^2 \Lambda}{\partial P_1 \partial P_K} \\ & -\frac{\partial^2 \Lambda}{\partial P_2^2} & \dots & -\frac{\partial^2 \Lambda}{\partial P_2 \partial P_K} \\ & & \ddots & \vdots \\ \mathrm{sym} & & & -\frac{\partial^2 \Lambda}{\partial P_K^2} \end{bmatrix}^{-1} \tag{6}$$

the partial derivatives of the Log-Likelihood [32] Λ

$$\Lambda(P_1, P_2, \dots, P_K) = \ln\left(\Pi_{i,j,l}\, f(t_i) \cdot \left(1 - F(t_j)\right) \cdot F(t_l)\right) \tag{7}$$

of the failure distribution $F(t)$ (cdf with pdf $f(t)$ and parameters $P_1, P_2, \dots, P_K$) can be used together with the partial derivatives

$$\Psi = \left[\frac{\partial F^{-1}(q)}{\partial P_1}, \frac{\partial F^{-1}(q)}{\partial P_2}, \dots, \frac{\partial F^{-1}(q)}{\partial P_K}\right] \tag{8}$$

of the quantile function $F^{-1}(q)$ of the failure distribution, to calculate the variance of the lifetime quantile

$$\mathrm{Var}(t_q) = \Psi' V \Psi. \tag{9}$$

Using this variance and the central limit theorem, the confidence distribution of the lifetime quantile can be estimated as $t_q \sim \mathcal{N}\left(t_q; \sqrt{\mathrm{Var}(t_q)}\right)$ [11], [12] which is then used for f_{H_1} and f_{H_0}. The calculation of the P_{ts} results in [12], [16]

$$P_{\mathrm{ts}} = 1 - \Phi\left(\Phi^{-1}\left(C_{\mathrm{req}}; t_{\mathrm{req}}, (1-s) \cdot \sqrt{\mathrm{Var}(t_R)}\right); t_{\mathrm{p}}, \sqrt{\mathrm{Var}(t_R)}\right) \tag{11}$$

Whereas $\Phi(x; \mu, \sigma)$ is the cumulative distribution function (cdf) at x of the normal distribution with parameters μ and σ and $\Phi^{-1}(q; \mu, \sigma)$ is the quantile function of the normal distribution for the proportion q. The herein used failure times t_i and censored times t_j, t_l are calculated as expected times using the quantile function and respective censoring scheme of the test [11]. The advantages of the analytic approach are the short computation time, since no Monte-Carlo Simulation has to take place, and the good approximation. The bootstrap method on the other hand does allow for all sampling schemes and boundary conditions that may be present during the individual tests. It also allows for very accurate estimations if high iteration counts are used [11], [25]. For a more detailed explanation on the calculation of the Probability of Test Success as well as the effect of certain influencing parameters, refer to [11], [12], [15], [16], [18], [20], [25], [26]. The P_{ts} can also be calculated for failure free tests [11].

3. ACCELERATED RELIABILITY DEMONSTRATION TEST AS A HYPOTHESIS TEST

In order to know how much the failure times are reduced by increasing the load, a lifetime model is required [1], [3]. The most used lifetime models in reliability engineering are the SN-curve according to Wöhler and Basquin [4], [7] for mechanical loads as well as the Arrhenius equation for temperature loads [1], [3]. The most used failure distribution in this context are the Weibull distribution $\mathcal{W}(t; \tau, b)$ [1], [33] with parameters τ and b as well as the lognormal distribution $\mathcal{N}_{\log}(t; \mu, \sigma)$ with parameters μ and σ. We will use those in the following to establish the algorithms and equations needed for the calculation of the Probability of Test Success P_{ts}. For estimating the lifetime in actual field usage, it is crucial to have a proper and representative estimate of the relevant load profile in the field. This usually is also a quantile in the form of e.g. a 95 % customer usage [34]. It is better to have a load profile with several load steps rather than a block profile or even a block load profile derived e.g. from standards. In order to use SN-curves, it is important to use the correct equivalent stress (e.g. von Mises stress [35]) and also have properly calibrated models to calculate those stresses and their components as well as the best concept available to evaluate local stresses at the correct location of the product [4]. When using the Arrhenius model, it is equally important to make sure the temperature load at the relevant location on the product is used. For both models it is important, to use the appropriate counting algorithm, e.g. rainflow counting for SN-curve [34] and dwell time counting for the Arrhenius model. In order to do a reliability demonstration using those two models, specimen have to be subjected to load. In the case of the SN-curve, to alternating load and in the case of the Arrhenius model they have to be subjected to thermal load, usually constant. The failure times of the specimen are then used to do an estimate of the parameters of the respective models. The reciprocal value of the damage is the estimate of the lifetime quantile. Using a Fisher confidence bound and comparing it to the reliability requirement can result in a successful reliability demonstration [1]. Since the failure times scatter however, this will not always be the case. It depends on the actual lifetime model parameters, the safety distance, the scattering parameters, the sample sizes, the load profile and its shape as well as the load levels during testing. The P_{ts} does overcome this challenge.

Wöhler-Lognormal Model

The Wöhler-Lognormal model makes use of the Basquin equation and the lognormal distribution. The equation of the failure distribution results in [3]

$$f_{\text{WöL}}(t) = \frac{1}{t\cdot\sigma\sqrt{2\pi}} e^{-\frac{\left(\ln(t)-\ln\left(N_A\cdot\left(\frac{S}{S_A}\right)^{-k}\right)\right)^2}{2\sigma^2}}. \tag{12}$$

With parameters N_A, S_A, k of the Wöhler model and σ as the standard deviation of the logarithmic failure times of the lognormal distribution. The parameters N_A, S_A can be combined to a single parameter. However, the Wöhler model is usually defined by two parameters, which represent the load S_A (usually stress in MPa) and the number of load cycles N_A which can be endured at this load. $k$is the slope of the model. For the estimation of those parameters S_A will be a fixed in this paper. The damage induced by the load profile with load levels $S_{\text{load}} = [S_1, S_2, \dots, S_m]$ and cycles $t_{\text{load}} = [t_1, t_2, \dots, t_m]$ is calculated by

$$D_{\text{WöL}} = \Sigma_j^m \frac{t_j}{N_A\cdot\left(\frac{S_j}{S_A}\right)^{-k}}. \tag{13}$$

Since the parameters of the Wöhler model represent the $50\ \%$ quantile, the lifetime quantile for R_{req} is

$$t_{R,\text{WöL}} = \frac{D_{R_{\text{req}}}}{D_{\text{WöL}}}\cdot t_{\text{req}} = \frac{\mathcal{N}_{\log}^{-1}(R_{\text{req}}; 0,\sigma)}{\Sigma_j^m \frac{t_j}{N_A\cdot\left(\frac{S_j}{S_A}\right)^{-k}}}\cdot t_{\text{req}} \tag{14}$$

as far as the load profile is representant of the required lifetime t_{req}.

Wöhler-Weibull Model

The Wöhler-Weibull model makes use of the Basquin equation and the Weibull distribution. The equation of the failure distribution is

$$f_{\text{Wö}W}(t) = \frac{b}{N_A\cdot\left(\frac{S}{S_A}\right)^{-k}\cdot(-\ln(0.5))^{-\frac{1}{b}}}\left(\frac{t}{N_A\cdot\left(\frac{S}{S_A}\right)^{-k}\cdot(-\ln(0.5))^{-\frac{1}{b}}}\right)^{b-1}\cdot e^{-\left(\frac{t}{N_A\cdot\left(\frac{S}{S_A}\right)^{-k}\cdot(-\ln(0.5))^{-\frac{1}{b}}}\right)^{b}}. \tag{15}$$

b is the Weibull shape parameter of the Weibull distribution. Here a parametrization with $(-\ln(0.5))^{-\frac{1}{b}}$ is used. This is so that the parameters of the Wöhler model estimated using this equation, represent the $50\ \%$ quantile and not the $63.2\ \%$ quantile. The respective lifetime quantile using this model is calculated by

$$t_{R,\mathrm{W\ddot{o}}W} = \frac{D_{R_{\mathrm{req}}}}{D_{\mathrm{W\ddot{o}}W}} \cdot t_{\mathrm{req}} = \frac{\left(\frac{\ln\left(R_{\mathrm{req}}\right)}{\ln\left(0.5\right)}\right)^{1/b}}{\Sigma_j^m \frac{t_j}{N_A \cdot \left(\frac{S_j}{S_A}\right)^{-k}}} \cdot t_{\mathrm{req}}. \tag{16}$$

This equation is also valid for the load profile with S_{load} and t_{load} representative of t_{req}.

Arrhenius-Lognormal Model

The Arrhenius-Lognormal model makes use of the Arrhenius equation and the lognormal distribution. The resulting equation is [3]

$$f_{\mathrm{AL}}(t) = \frac{1}{t \cdot \sigma\sqrt{2\pi}} e^{-\frac{\left(\ln(t) - \ln\left(A \cdot e^{\frac{E_A}{k \cdot T}}\right)\right)^2}{2\sigma^2}}. \tag{17}$$

Here, the parameters of the Arrhenius equation are the activation energy E_A and A as a constant. The variable k represents the Boltzmann constant [36] as $8.6173 \cdot 10^{-5}$ eV/K. The damage which is induced by the load profile with m load levels S_{load} and respective dwell times t_{load} representative of t_{req} is

$$D_{\mathrm{AL}} = \Sigma_j^m \frac{t_j}{A \cdot e^{\frac{E_A}{k \cdot S_j}}}. \tag{18}$$

To calculate the lifetime quantile for R_{req} the equation is, similar to eq. 14 and 17:

$$t_{R,\mathrm{AL}} = \frac{D_{R_{\mathrm{req}}}}{D_{\mathrm{AL}}} \cdot t_{\mathrm{req}} = \frac{\mathcal{N}_{\log}^{-1}\left(R_{\mathrm{req}}; 0, \sigma\right)}{\Sigma_j^m \frac{t_j}{A \cdot e^{\frac{E_A}{k \cdot S_j}}}} \cdot t_{\mathrm{req}}. \tag{19}$$

Arrhenius-Weibull Model

The Arrhenius-Weibull model makes use of the Arrhenius equation and the Weibull distribution. The failure distribution is

$$f_{\mathrm{A}W}(t) = \frac{b}{A \cdot e^{\frac{E_A}{k \cdot T}} \cdot \left(-\ln(0.5)\right)^{-\frac{1}{b}}} \left(\frac{t}{A \cdot e^{\frac{E_A}{k \cdot T}} \cdot \left(-\ln(0.5)\right)^{-\frac{1}{b}}}\right)^{b-1} \cdot e^{-\left(\frac{t}{A \cdot e^{\frac{E_A}{k \cdot T}} \cdot \left(-\ln(0.5)\right)^{-\frac{1}{b}}}\right)^b}. \tag{20}$$

Again, the equation allows for the parameters of the lifetime model to be estimated as a 50 % quantile. The lifetime quantile using the load profile results in

$$t_{R,\mathrm{A}W} = \frac{D_{R_{\mathrm{req}}}}{D_{\mathrm{A}W}} \cdot t_{\mathrm{req}} = \frac{\left(\frac{\ln\left(R_{\mathrm{req}}\right)}{\ln\left(0.5\right)}\right)^{1/b}}{\Sigma_j^m \frac{t_j}{A \cdot e^{\frac{E_A}{k \cdot S_j}}}} \cdot t_{\mathrm{req}}. \tag{21}$$

Using these equations, lifetime quantiles can be calculated for the relevant load profiles. In order to calculate the Probability of Test Success for tests which are estimating the lifetime quantile through the use of these models, the confidence distribution of these lifetime quantiles under validity of both hypotheses f_{H0} and f_{H1} need to be estimated, see eq. 1, 2, 4 and 5. A bootstrap procedure as well as an analytic procedure will be presented in the following. They are able to estimate the two distributions by using above equations.

3.1. Bootstrap Calculation Procedure

The Bootstrap calculation procedure is similar to those presented in [11], [12]. However, the failure distribution now becomes dependent on the load S in the equations 12, 16, 19 and 22.

First, the failure times on the test load levels $S_{\mathrm{test}} = [S_{\mathrm{test},1}, S_{\mathrm{test},2}, \dots, S_{\mathrm{test},o}]$ with their respective sample size $n = [n_1, n_2, \dots, n_0]$, are generated using a pseudo random number generator and the quantile function of either the lognormal distribution or the Weibull distribution, alternatively, a rejection sampling method can be used [25]. However, the quantile function must take into account the load level height according to above equations. The failure times together with the load levels are then used to estimate the parameters of the Wöhler-Lognormal model, the Wöhler-Weibull model, the Arrhenius-Lognormal model or the Arrhenius-

Weibull model. The lifetime quantile is then calculated using the load profile S_{load} and t_{load} and the respective quantile functions stated above. By iterating this algorithm multiple times, e.g. 10,000 times, a sample from the distribution of t_{R,H_1} is obtained. To correct possible estimation bias (e.g. from MLE [37], [38]) the values of the sample should be corrected according to

$$t_{R,H_1,i} = \frac{t_{R,H_1,i} \cdot t_{\mathrm{req}}}{D_{\mathrm{p}} \cdot t_{R,H_1,50\%}}. \tag{22}$$

In order to correspond to the hypothesis of eq. 1 and 2. With D_{p} being either $D_{\mathrm{W\ddot{o}L}}$, $D_{\mathrm{W\ddot{o}}\mathcal{W}}$, D_{AL}or $D_{\mathrm{A}\mathcal{W}}$ and $t_{R,H_1,50\%}$ being the median of the generated values of the bootstrap procedure. Since the lifetime model only requires a linear shift, so that $s = 0$ of H_0 is fulfilled, the lifetime quantiles under validity of H_0 can be calculated using the values of $t_{R,H_1,i}$ as follows:

$$t_{R,H_0,i} = \frac{t_{R,H_1,i} \cdot t_{\mathrm{req}}}{t_{R,H_1,50\%}}. \tag{23}$$

Using those values, the P_{ts} calculates to

$$C_{\mathrm{req}} \overset{!}{=} \frac{\text{Number of } t_{R,H_0} \leq t_{\mathrm{crit}}}{\text{Total number of iterations}} \tag{24}$$

$$P_{\mathrm{ts}} = \frac{\text{Number of } t_{R,H_1} \geq t_{\mathrm{crit}}}{\text{Total number of iterations}}. \tag{25}$$

3.2. Analytic Calculation Procedure

For the analytic calculation procedure, the asymptotic behavior of the maximum likelihood estimate (MLE) regarding the parameters of the model according to the central limit theorem is used [11], [12]. The failure distributions models of eq. 12, 15, 17 and 20 need to be plugged into eq. 6 to 9. However, instead of using the quantile function in eq. 8, the logarithm of the quantile function shall be used, since the lifetime quantile is always greater than zero. The resulting equations without censoring are the following.

<u>Wöhler-Lognormal Model</u>

Partial derivatives of the Log-Likelihood for variance covariance matrix:

$$\frac{\partial^2 \Lambda}{\partial N_A^2} = -\frac{n}{N_A^2 \sigma^2} + \frac{1}{N_A^2 \sigma^2} \sum_i^n \left(\ln\left(N_A \left(\frac{S_{\mathrm{test},i}}{S_A} \right)^{-k} \right) - \ln(t_i) \right). \tag{26}$$

$$\frac{\partial^2 \Lambda}{\partial N_A \partial k} = \frac{1}{N_A \sigma^2} \sum_i^n \ln\left(\frac{S_{\mathrm{test},i}}{S_A} \right). \tag{27}$$

$$\frac{\partial^2 \Lambda}{\partial N_A \partial \sigma} = -\frac{2}{N_A \sigma^3} \sum_i^n \left(\ln(t_i) - \ln\left(N_A \left(\frac{S_{\mathrm{test},i}}{S_A} \right)^{-k} \right) \right). \tag{28}$$

$$\frac{\partial^2 \Lambda}{\partial k^2} = -\frac{1}{\sigma^2} \sum_i^n \ln\left(\frac{S_{\mathrm{test},i}}{S_A} \right)^2. \tag{29}$$

$$\frac{\partial^2 \Lambda}{\partial k \partial \sigma} = \frac{2}{\sigma^3} \sum_i^n \ln\left(\frac{S_{\mathrm{test},i}}{S_A} \right) \left(\ln(t_i) - \ln\left(N_A \left(\frac{S_{\mathrm{test},i}}{S_A} \right)^{-k} \right) \right). \tag{30}$$

$$\frac{\partial^2 \Lambda}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{3}{\sigma^4} \sum_i^n \left(\ln(t_i) - \ln\left(N_A \left(\frac{S_{\mathrm{test},i}}{S_A} \right)^{-k} \right) \right)^2. \tag{31}$$

Partial derivatives of the logarithmic quantile function:

$$\Psi = \left[\frac{1}{N_A}, -\frac{\Sigma_j^o t_j \left(\frac{S_{\mathrm{load},j}}{S_A} \right)^k \ln\left(\frac{S_{load,j}}{S_A} \right)}{\Sigma_j^o t_j \left(\frac{S_{\mathrm{load},j}}{S_A} \right)^k}, \sqrt{2}\, \mathrm{erf}^{-1}\left(2\left(R_{\mathrm{req}} - 1 \right) - 1 \right) \right] \tag{32}$$

Wöhler-Weibull Model

Partial derivatives of the Log-Likelihood for variance covariance matrix:

$$\frac{\partial^2\Lambda}{\partial N_A^2} = \frac{b}{N_A^2}\sum_i^n\left(1-(b+1)\left(\frac{t_i}{N_A\left(\frac{S_{\mathrm{test},i}}{S_A}\right)^{-k}(-\ln{(0.5)})^{-1/b}}\right)^b\right). \tag{33}$$

$$\frac{\partial^2\Lambda}{\partial N_A\partial k} = \frac{b^2}{N_A}\sum_i^n\ln\left(\frac{S_{test,i}}{S_A}\right)\left(\frac{t_i}{N_A\left(\frac{S_{\mathrm{test},i}}{S_A}\right)^{-k}(-\ln{(0.5)})^{-1/b}}\right)^b. \tag{34}$$

$$\frac{\partial^2\Lambda}{\partial N_A\partial b} = \frac{1}{N_A}\sum_i^n\left(\left(\left(\frac{t_i}{N_A\left(\frac{S_{\mathrm{test},i}}{S_A}\right)^{-k}(-\ln{(0.5)})^{-1/b}}\right)^b\cdot\left(1-\ln(-ln(0.5))+b\cdot ln\left(\frac{t_i}{N_A\left(\frac{S_{\mathrm{test},i}}{S_A}\right)^{-k}(-\ln{(0.5)})^{-1/b}}\right)\right)\right)-1\right) \tag{35}$$

$$\frac{\partial^2\Lambda}{\partial k^2} = -b^2\sum_i^n\ln\left(\frac{S_{\mathrm{test},i}}{S_A}\right)^2\left(\frac{t_i}{N_A\left(\frac{S_{\mathrm{test},i}}{S_A}\right)^{-k}(-\ln{(0.5)})^{-1/b}}\right)^b. \tag{36}$$

$$\frac{\partial^2\Lambda}{\partial k\partial b} = \sum_i^n\ln\left(\frac{S_{\mathrm{test},i}}{S_A}\right)\left(1-\left(\frac{t_i}{N_A\left(\frac{S_{\mathrm{test},i}}{S_A}\right)^{-k}(-\ln(0.5))^{-\frac{1}{b}}}\right)^b\left(1+b\ln\left(\frac{t_i}{N_A\left(\frac{S_{\mathrm{test},i}}{S_A}\right)^{-k}(-\ln(0.5))^{-\frac{1}{b}}}\right)-(\ln(-\ln(0.5)))\right)\right) \tag{37}$$

$$\frac{\partial^2\Lambda}{\partial b^2} = -\frac{n}{b^2}-\sum_i^n\left(\left(\frac{t_i}{N_A\left(\frac{S_{\mathrm{test},i}}{S_A}\right)^{-k}(-\ln(0.5))^{-\frac{1}{b}}}\right)^b\cdot\frac{\ln\left(\frac{t_i}{N_A\left(\frac{S_{\mathrm{test},i}}{S_A}\right)^{-k}(-\ln(0.5))^{-\frac{1}{b}}}\right)-\ln(-\ln(0.5))}{b}\right)^2. \tag{38}$$

Partial derivatives of the logarithmic quantile function:

$$\Psi = \left[\frac{1}{N_A},-\frac{\Sigma_j^o t_j\left(\frac{S_{\mathrm{load},j}}{S_A}\right)^k\ln\left(\frac{S_{load,j}}{S_A}\right)}{\Sigma_j^o t_j\left(\frac{S_{\mathrm{load},j}}{S_A}\right)^k},-\frac{1}{b^2}\ln\left(\frac{ln(R_{req})}{ln(0.5)}\right)\right] \tag{39}$$

Arrhenius-Lognormal Model

Partial derivatives of the Log-Likelihood for variance covariance matrix:

$$\frac{\partial^2\Lambda}{\partial A^2} = \frac{1}{A^2\sigma^2}\sum_i^n\left(\ln\left(A\cdot e^{\frac{E_A}{kS_i}}\right)-\ln(t_i)-1\right). \tag{40}$$

$$\frac{\partial^2 \Lambda}{\partial A \partial E_A} = -\frac{1}{Ak\sigma^2} \textstyle\sum_i^n 1/S_i. \tag{41}$$

$$\frac{\partial^2 \Lambda}{\partial A \partial \sigma} = -\frac{2}{A\sigma^3} \sum_i^n \left(\ln(t_i) - \ln\left(A \cdot e^{\frac{E_A}{kS_i}} \right) \right). \tag{42}$$

$$\frac{\partial^2 \Lambda}{\partial E_A^2} = -\sum_i^n \left(\frac{1}{k\sigma S_i} \right)^2. \tag{43}$$

$$\frac{\partial^2 \Lambda}{\partial E_A \partial \sigma} = -\frac{2}{k\sigma^3} \sum_i^n \frac{1}{t_i} \left(\ln(t_i) - \ln\left(A \cdot e^{\frac{E_A}{kS_i}} \right) \right). \tag{44}$$

$$\frac{\partial^2 \Lambda}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{3}{\sigma^4} \sum_i^n \left(\ln(t_i) - \ln\left(A \cdot e^{\frac{E_A}{kS_i}} \right) \right)^2. \tag{45}$$

Partial derivatives of the logarithmic quantile function:

$$\Psi = \left[\frac{1}{A}, \frac{\Sigma_i^o \frac{t_i}{AkS_i e^{\frac{E_A}{kS_i}}}}{\Sigma_i^o \frac{t_i}{Ae^{\frac{E_A}{kS_i}}}}, \sqrt{2}\,\mathrm{erf}^{-1}\left(2\left(R_{\mathrm{req}} - 1\right) - 1\right) \right]. \tag{46}$$

$\mathrm{erf}^{-1}(x)$ is the inverse of the error function $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \mathrm{d}t$ [39] and has to be solved numerically.

<u>Arrhenius-Weibull Model</u>

Partial derivatives of the Log-Likelihood for variance covariance matrix:

$$\frac{\partial^2 \Lambda}{\partial A^2} = \frac{b}{N_A^2} \sum_i^n \left(1 - (b+1) \left(\frac{t_i}{Ae^{E_A/(kS_i)}(-\ln(0.5))^{-1/b}} \right)^b \right). \tag{47}$$

$$\frac{\partial^2 \Lambda}{\partial A \partial E_A} = -\frac{b^2}{Ak} \sum_i^n 1/S_i \left(\frac{t_i}{Ae^{E_A/(kS_i)}(-\ln(0.5))^{-1/b}} \right)^b. \tag{48}$$

$$\frac{\partial^2 \Lambda}{\partial A \partial b} = \frac{1}{A} \sum_i^n \left(\left(\frac{t_i}{Ae^{E_A/(kS_i)}(-\ln(0.5))^{-1/b}} \right)^b \cdot \left(1 - \ln(-ln(0.5)) + b \ln\left(\frac{t_i}{Ae^{E_A/(kS_i)}(-\ln(0.5))^{-1/b}} \right) \right) - 1 \right). \tag{49}$$

$$\frac{\partial^2 \Lambda}{\partial E_A^2} = -\frac{b^2}{k^2} \textstyle\sum_i^n 1/{S_i}^2 \left(\frac{t_i}{Ae^{E_A/(kS_i)}(-\ln(0.5))^{-1/b}} \right)^b. \tag{50}$$

$$\frac{\partial^2 \Lambda}{\partial E_A \partial b} = 1/k \sum_i^n 1/S_i \left(1 + \left(\frac{t_i}{Ae^{E_A/(kS_i)}(-\ln(0.5))^{-1/b}} \right)^b \left(1 + b \ln\left(\frac{t_i}{Ae^{E_A/(kS_i)}(-\ln(0.5))^{-1/b}} \right) - (\ln(-\ln(0.5)))) \right). \tag{51}$$

$$\frac{\partial^2 \Lambda}{\partial b^2} = -n \left(\frac{1 + \ln(-\ln(0.5))(-\ln(0.5))^{-\frac{1}{b}}}{b^2} + \frac{\ln(-\ln(0.5))(-\ln(0.5))^{-\frac{1}{b}}(b - \ln(-\ln(0.5)))}{b^3} \right) - \sum_i^n \left(\left(\frac{t_i}{Ae^{E_A/(kS_i)}(-\ln(0.5))^{-1/b}} \right)^b \cdot (\ln\left(\frac{t_i}{Ae^{\frac{E_A}{kS_i}}(-\ln(0.5))^{-\frac{1}{b}}} \right) - \ln(-\ln(0.5)))/b \right)^2. \tag{52}$$

Partial derivatives of the logarithmic quantile function:

$$\Psi = \left[\frac{1}{A}, \frac{\Sigma_i^o \frac{t_i}{AkS_i e^{\frac{E_A}{kS_i}}}}{\Sigma_i^o \frac{t_i}{Ae^{\frac{E_A}{kS_i}}}}, -\frac{1}{b^2} \ln\left(\frac{ln(R_{req})}{ln(0.5)} \right) \right] \tag{53}$$

The P_{ts} can be calculated as

$$P_{\text{ts}} = 1 - \Phi_{\log}\left(\Phi_{\log}^{-1}\left(C_{\text{req}}; \ln\left(t_{req}\right), (1-s) \cdot \sqrt{\text{Var}(t_R)}\right); \ln\,(t_{\text{p}}), \sqrt{\text{Var}(t_R)}\right) \tag{54}$$

With $\Phi_{\log}(x;\ \mu, \sigma)$ as the cumulative distribution function of the lognormal distribution and $\Phi_{\log}^{-1}(q;\ \mu, \sigma)$ as its quantile function and $\text{Var}\left(\ln\left(t_q\right)\right) = \Psi' V \Psi$.

3.3. Comparison of Calculation Procedures

The two proposed calculation procedures based on the established hypothesis testing concept are compared exemplarily in the following using the fixed parameters shown in Tab. 1.

Table 1. Parameters for comparison

	Parameter Values
Wöhler	$N_A = 10^6; S_A = 60; k = 5$
Arrhenius	$A = 10^{-9}; E_A = 0.8$
Scattering	$\sigma = 0.6; b = 2.5$
Load profile	$S_{\text{load}}[180, 60, 30];\ t_{\text{load}}[10^3, 4 \cdot 10^4, 5 \cdot 10^6] \cdot \frac{D_{R\text{req}}}{D_{\text{p}}}(1-s)$
Test Configuration	$S_{\text{test}} = [200, 100]; n = [15, 5]$
Reliability Requirement	$R_{\text{req}} = 0.9; C_{\text{req}} = 0.9; s = 0.5; t_{\text{req}} = 1$ (normalized)

The safety distance of eq. 3 is varied in Fig. 1. It can be seen that the approximation of the analytic procedure is very good for all four models. If different scattering of the failure times is present, the approximation is also good, but deviates a bit from the values calculated using the Bootstrap procedure, as seen in Fig. 2. The Approximation of the Analytic Procedure is suitable for the usage of finding feasible solutions of test configuration, if a variation of the total sample size is analyzed, as in Fig. 3. The sample shares on the high and low load level stayed the same.

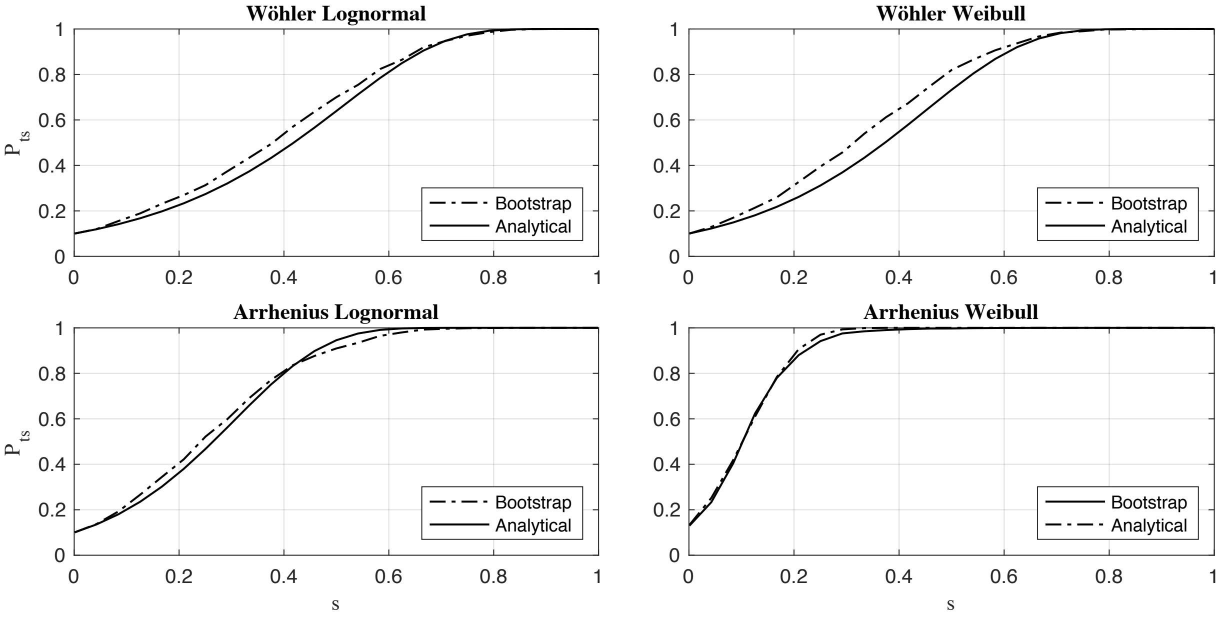

Figure 1. Comparison of Bootstrap and Analytic Calculation Procedure for varying safety distances s.

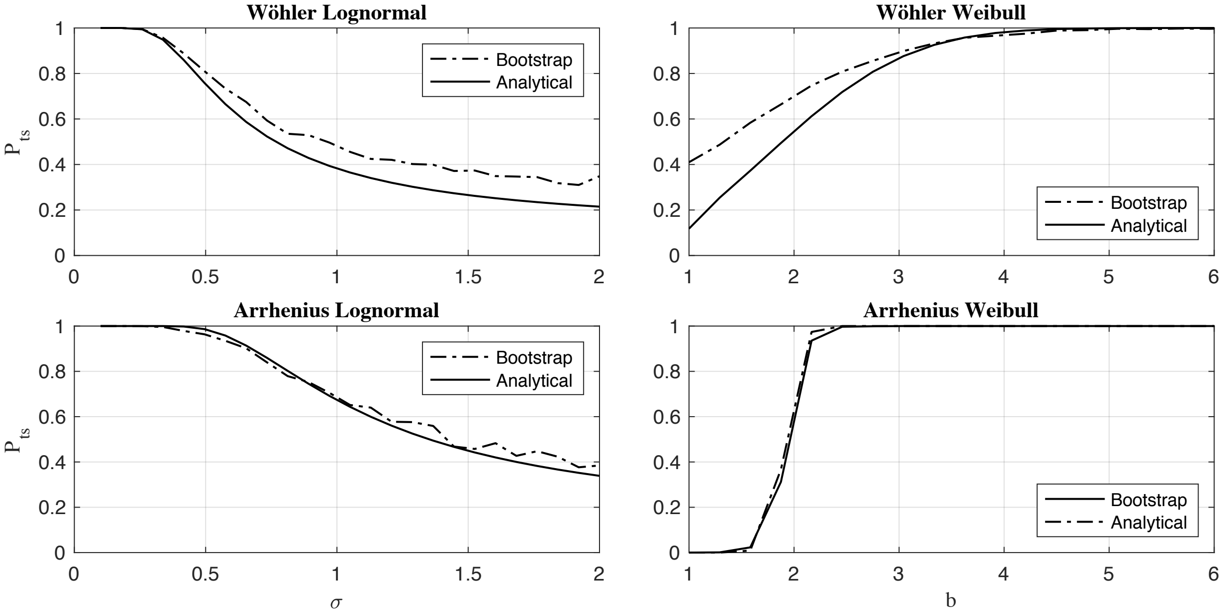

Figure 2. Comparison of Bootstrap and Analytic Calculation Procedure for varying parameters σ and b.

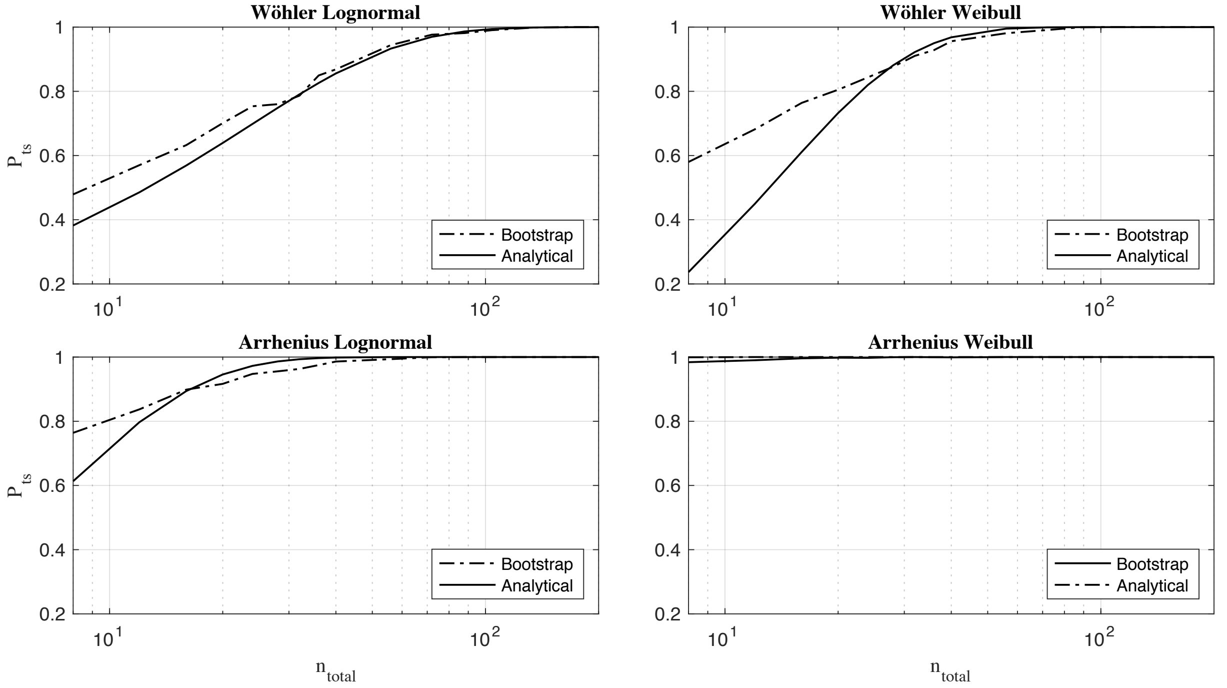

Figure 3. Comparison of Bootstrap and Analytic Calculation Procedure for varying sample sizes n_{total}.

3.4. Procedure for Planning of Efficient Accelerated Reliability Demonstration Tests

In order to identify the test configuration best suited to provide a successful reliability demonstration using one of the four lifetime models, a large parameter space must be analyzed. For the assessment of the tests, the P_{ts} has to be calculated. Due to the computational effort of the bootstrap approach, the analytic procedure can be used in order to narrow the feasible parameter space. Afterwards, the bootstrap procedure can be used for a more accurate calculation of the P_{ts}. By this, the question if more than one load level needs to be tested, how high the load levels shall be and how many specimens are to be tested on those load levels can be answered. Since the scattering of the lifetime model parameters are considered in this approach and are captured in the calculated lifetime quantiles which are used for the P_{ts} calculation, the shape of the load profile does matter and a reduction to a block profile is not advisable

4. CONCLUSION

The Probability of Test Success P_{ts} is a suitable metric to assess accelerated reliability demonstration tests. It is shown how the P_{ts} can be understood as the statistical power of such an accelerated test. This is done by using the hypothesis definitions of the P_{ts} and developing and establishing the required equations in order to make use of it for the accelerated tests. The lifetime models used in this paper are the Wöhler-Lognormal, Wöhler-Weibull, Arrhenius-Lognormal and Arrhenius-Weibull. For these, the equations are derived and two calculation procedures which make use of the hypothesis testing context are developed. The first one being a bootstrap approach and the second one is an analytic one, which allows for a very fast computation and thus enables the identification of the optimal test, since usually a large parameter space needs to be evaluated. A comparison of those calculation procedures shows good approximation of the analytic approach to the bootstrap approach for all four model. The approach and equations herein are a crucial step towards the planning of highly efficient accelerated reliability demonstration tests.

References

[1] B. Bertsche, *Reliability in Automotive and Mechanical Engineering*. Berlin: Springer, 2008.

[2] A. Birolini, *Reliability engineering*, 8th ed. Berlin: Springer-Verlag, 2017.

[3] W. B. Nelson, *Accelerated testing : statistical models, test plans and data analyses*. Hoboken: John Wiley & Sons, 2004.

[4] E. Haibach, *Betriebsfestigkeit*, 3rd ed. Wiesbaden: Springer-Verlag, 2006.

[5] A. Grundler, M. Bollmann, M. Obermayr, and B. Bertsche, “Berücksichtigung von Lebensdauerberechnungen als Vorkenntnis im Zuverlässigkeitsnachweis,” in *VDI-Fachtagung Technische Zuverlässigkeit 2019*, 2019.

[6] T. Herzig, “Anforderungsgerechte Produktauslegung durch Planung effizienter beschleunigter Zuverlässigkeitstests,” Universität Stuttgart, 2021.

[7] J. A. Bannantine, J. J. Comer, and J. L. Handrock, *Fundamentals of Metal Fatigue Analysis*, vol. 27, no. 05. New Jersey: Prentice Hall, 1990.

[8] W. B. Nelson, *Applied Life Data Analysis*. Hoboken: John Wiley & Sons, 2004.

[9] T. Herzig, M. Dazer, A. Grundler, and B. Bertsche, “Evaluation of optimality criteria for efficient reliability demonstration testing,” in *30th European Safety and Reliability Conference, ESREL 2020 and 15th Probabilistic Safety Assessment and Management Conference, PSAM 2020*, 2020.

[10] M. Dazer, D. Brautigam, T. Leopold, and B. Bertsche, “Optimal Planning of Reliability Life Tests Considering Prior Knowledge,” in *2018 Annual Reliability and Maintainability Symposium (RAMS)*, Jan. 2018, pp. 1–7, doi: 10.1109/RAM.2018.8463098.

[11] A. Grundler, M. Dazer, and T. Herzig, “Statistical Power Analysis in Reliability Demonstration Testing: The Probability of Test Success,” *Appl. Sci.*, vol. 12, no. 12, p. 6190, Jun. 2022, doi: 10.3390/app12126190.

[12] A. Grundler and M. Dazer, “Reliability Demonstration Test Planning for Systems Using Prior Knowledge,” *IEEE Access*, vol. 12, no. December 2023, pp. 3339–3355, 2024, doi: 10.1109/ACCESS.2023.3347035.

[13] M. Dazer, T. Herzig, A. Grundler, and B. Bertsche, “R-OPTIMA : Optimal Planning of Reliability Tets,” in *Proceedings IRF2020: 7th International Conference Integrity-Reliability-Failure*, 2020, pp. 695–702

[14] M. Dazer, A. Grundler, T. Herzig, D. Engert, and B. Bertsche, “Effect on Interval Censoring on the Probability of Test Success in Reliability,” in *2021Annual Reliability and Maintainability Symposium (RAMS)*, 2021.

[15] A. Grundler, M. Dazer, and B. Bertsche, “Effect of Uncertainty in Prior Knowledge on Test Planning for a Brake Caliper using the Probability of Test Success,” in *2021 Annual Reliability and Maintainability Symposium (RAMS)*, May 2021, pp. 1–9, doi: 10.1109/RAMS48097.2021.9605727.

[16] A. Grundler, M. Dazer, and B. Bertsche, “Efficient Reliability Demonstration using the Probability of Test Success and Bayes Theorem,” in *Proceedings of the Probabilistic Safety Assessment and Management Conference, PSAM 16*, 2022

[17] A. Grundler, M. Bartholdt, and B. Bertsche, “Statistical test planning using prior knowledge—advancing the approach of Beyer and Lauster,” in *Safety and Reliability – Safe Societies in a Changing World*, vol. 0, London: CRC Press, 2018, pp. 809–814.

[18] A. Grundler, M. Dazer, T. Herzig, and B. Bertsche, “Considering Multiple Failure Mechanisms in Optimal Test Design,” in *Proceedings IRF2020: 7th International Conference Integrity-Reliability-*

Failure, 2020, pp. 673–682

[19] M. Bartholdt, A. Grundler, M. Bollmann, and B. Bertsche, “Assurance of the System Reliability of a Gearbox Considering Prior Knowledge,” *Proc. Int. Des. Conf. Des.*, vol. 3, no. 1988, pp. 965–974, 2018, doi: 10.21278/idc.2018.0162.

[20] A. Grundler, M. Dazer, T. Herzig, and B. Bertsche, “Efficient System Reliability Demonstration Tests Using the Probability of Test Success,” in *Proceedings of the 31st European Safety and Reliability Conference ESREL 2021*, 2021, pp. 1654–1661, doi: 10.3850/978-981-18-2016-8.

[21] T. Herzig, M. Dazer, A. Grundler, and B. Bertsche, “Cost-And time-effective planning of accelerated reliability demonstration tests-A new approach of comparing the expenditure of success run and end-of-life tests,” in *Proceedings - Annual Reliability and Maintainability Symposium*, 2019, vol. 2019-Janua, doi: 10.1109/RAMS.2019.8769262.

[22] T. Herzig, M. Dazer, A. Grundler, and B. Bertsche, “Integrating Accelerated Life Tests into Optimal Test Planning,” in *Proceedings IRF2020: 7th International Conference Integrity-Reliability-Failure.*, 2020, pp. 665–672

[23] A. Benz, A. Grundler, M. Dazer, and B. Bertsche, “Deriving Prior Knowledge from Lifetime Simulations for Reliability Demonstration while Considering the Uncertainty of the Lifetime Model,” in *Proceedings of the 31st European Safety and Reliability Conference (ESREL 2021)*, 2021, vol. 1 PartF, pp. 466–472, doi: 10.3850/978-981-18-2016-8_228-cd.

[24] A. Benz, A. Grundler, T. Herzig, M. Dazer, and B. Bertsche, “Reliability Demonstration Test Planning for Field Load Spectra – an Approach for Identifying the Optimal Test Parameters considering Individual Cost and Time Constraints,” in *Annual Reliability and Maintainybility Symposium RAMS*, 2022.

[25] A. Grundler, “Zuverlässigkeitsnachweis von Systemen durch Komponenten- und Systemtests unter Berücksichtigung von Vorkenntnissen,” University of Stuttgart, 2024.

[26] A. Grundler, M. Dazer, and B. Bertsche, “Reliability-Test Planning Considering Multiple Failure Mechanisms and System Levels – an Approach for Identifying the Optimal System-Test Level, Type, and Configuration with Regard to Individual Cost and Time Constraints,” in *2020 Annual Reliability and Maintainability Symposium (RAMS)*, Jan. 2020, pp. 1–7, doi: 10.1109/RAMS48030.2020.9153678.

[27] T. Herzig, M. Dazer, and B. Bertsche, “Zuverlässigkeitsabsicherung ressourcenschonender Produkte durch effiziente Erprobungsplanung,” in *Stuttgarter Symposium für Produktentwicklung*, 2019.

[28] J. Neyman and E. S. Pearson, “On the Problem of the Most Efficient Tests of Statistical Hypotheses,” *Philos. Trans. R. Soc. London*, vol. 231, pp. 289–337, 1933, doi: 10.1007/978-1-4612-0919-5_6.

[29] R. A. Fisher, *The Design of Experiments*. Edinburgh: Oliver and Boyd, 1935.

[30] B. Efron and R. J. Tibshirani, “Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy,” *Stat. Sci.*, vol. 1, no. 1, pp. 54–77, 1986.

[31] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*. Chapman & Hall / CRC Press LLC, 1998.

[32] W. Q. Meeker, L. A. Escobar, and F. G. Pascual, *Statistical Methods for Reliability Data*, 2nd ed. New York, Chichester, Weinheim, Brisbane, Singapore, Toronto: Wiley, 2022.

[33] W. Weibull, “The New Weibull Handbook.”

[34] K. Lucan and B. Bertsche, “Standard Load Spectra for Commercial Vehicle Brakes,” *ATZ Worldw.*, vol. 121, no. 5, pp. 72–77, May 2019, doi: 10.1007/s38311-019-0043-1.

[35] R. von Mises, “Mechanik der Festen Korper im Plastisch Deformablen Zustand,” *Nachr. Ges. Wiss. Gottingen*, p. 582, 1913.

[36] L. Boltzmann, “Über die Beziehung zwischen dem zweiten Hauptsatz der mechnischen Warmtheorie und der Wahrscheinlichkeitsrechnung, respective den Sätzen über das Wärmegleichgewicht,” *Sitzungsberichte der Kais. Akad. der Wissenschaften. Math. Classe. Abt. 2, Math. Phys. Chemie, Mech. Meteorol. und Astron.*, vol. 76, 1877.

[37] T. Tevetoglu and B. Bertsche, “Bias corrected weibull parameter estimation and impact on confidence bounds,” in *30th European Safety and Reliability Conference, ESREL 2020 and 15th Probabilistic Safety Assessment and Management Conference, PSAM 2020*, 2020, doi: 10.3850/978-981-14-8593-0.

[38] T. Tevetoglu and B. Bertsche, “On the Coverage Probability of Bias-Corrected Confidence Bounds,” in *2020 Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling, APARM 2020*, 2020, doi: 10.1109/APARM49247.2020.9209464.

[39] J. W. L. Glaisher, “XXXII. On a class of definite integrals,” *London, Edinburgh, Dublin Philos. Mag. J. Sci.*, vol. 42, no. 280, pp. 294–302, Oct. 1871, doi: 10.1080/14786447108640568.