Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira,
Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria
 Published by Research Publishing, Singapore.
doi: 10.3850/ESREL2026061419_esrel26-p26203-cd

Bayesian Optimal Experimental Design for Planning Accelerated Life Testing Campaigns on Electronic Equipment

L. Riccardi
Department of Energy and Nuclear Science and Technology, Politecnico di Milano; Aramix s.r.l, Milano, Italy.
E-mail: luca.riccardi@polimi.it

L. Bellani
Aramix s.r.l, Milano, Italy.

M. Compare
Aramix s.r.l, Milano, Italy.

S. Medici
Aramix s.r.l, Milano, Italy.

A. Costa
Hitachi Rail, Napoli, Italy.

C. Maiorano
Hitachi Rail, Napoli, Italy.

E. Pascale
Hitachi Rail, Napoli, Italy.

E. Zio
Department of Energy and Nuclear Science and Technology, Politecnico di Milano; Mines-Paris, PSL University, CRC, Sophia Antipolis, France E-mail: enrico.zio@polimi.it

This paper presents a methodology for optimizing Accelerated Life Testing (ALT) to estimate the reliability of novel electronic equipment through Bayesian Optimal Experimental Design (BOED). The methodology consists of seven steps, from failure mode analysis, based on expert knowledge, to posterior parameter estimation, based on test data. Its practical application is demonstrated through a case study concerning a critical railway signaling component, the Balise Transmission Module (BTM). The results reveal that optimal resource allocation strategies are often non-intuitive trade-offs between test duration and the accuracy of reliability estimates.

Keywords: Accelerated Life Tests, Reliability testing, Electronic equipment, Bayesian Optimal Experimental Design, Bayesian Statistics, Design of Experiments

1. Introduction

The electronics industry is driven by technological advancements, with rapid innovations in materials, design architectures and manufacturing processes. This pace fuels a highly competitive market that continuously shortens the design-to-market timeline and accelerates obsolescence.

The fast-evolving market poses significant challenges to device reliability. Indeed, cornerstone standard-based approaches (e.g., International Electrotechnical Commission (2011), International Electrotechnical Commission (2003), FIDES Group (2009), Telcordia Technologies (2011)) might be inadequate for predicting the reliability of electronic components of new designs, which might be made up of different materials and manufactured with novel processes. This, in particular, questions the applicability of standards to guide critical design trade-offs and logistic support decisions, particularly for the Commercial-Off-The-Shelf (COTS) components that are now ubiquitous in high-reliability systems in all industrial sectors (e.g., Temsamani et al. (2017), Jang and Kim (2024)).

Furthermore, novel components do not have "Return from Experience" (ReX) data, so that reliability estimates based on data referring to previous designs of different technologies are affected by large uncertainties. Accelerated Life Test (ALT) campaigns need, then, to be performed: components are operated in laboratory-controlled environments under elevated stress conditions exceeding their nominal operational limits (e.g., high temperature, humidity, or vibration) to favor failure occurrence and gather statistical evidence within test durations compliant with the fast go-to-market requirements. ALT models are eventually used to translate results from tests to real-world operating conditions Indmeskine et al. (2023) Meeker and Escobar (1993).

Planning accelerated tests in an optimal way requires setting multiple parameters (e.g., the number of units to test, the stress factors, the stress levels, etc.). Many approaches have been proposed to optimally plan ALT campaigns (e.g., Nelson (2005a,b); Ryan et al. (2016); Escobar and Meeker (2006); Dodson and Schwab (2021)). In particular, Bayesian optimization of accelerated tests allows leveraging the manufacturer's know-how (e.g., past experiences on similar products) to set a priori distributions on model parameters. Bayesian setting was investigated in Zhang and Meeker (2006); Xu and Tang (2015), considering the Weibull distribution to model the device failure time, in Li et al. (2015), considering a Brownian process to model the degradation evolution, in Wang and Xu (2010); Li et al. (2017,?), considering degradation mechanisms obeying an inverse Gaussian process, and in Wang (2010); Li et al. (2016); Wan et al. (2014), considering Wiener degradation processes.

In this work, we propose a novel structured methodology for generating optimal, component-specific reliability evidence through ALT campaigns. This allows increasing the accuracy of reliability estimates for newly developed electronic devices prior to their market release, while minimizing the duration of testing. The proposed methodology builds on the Bayesian Optimal Experimental Design (BOED) theoretical framework Gaier and Ha (2019); Foster et al. (2020); Goda et al. (2022); Amzal et al. (2006), which provides the mathematical and statistical ground for selecting tests that are most likely to reduce the uncertainty in the estimates.

The paper is organized as follows. Section 2 provides an overview of the methodology, detailing its theoretical foundations and operational workflow. Section 3 presents a case study regarding the reliability assessment of a Balise Transmission Module (BTM). Section 4 offers a synthesis and discussion of the case study results, with an analysis of the key findings and their practical implications. Finally, Section 5 gives concluding remarks and insights for future research work.

2. A Structured methodology for Cost-Effective Reliability Testing

The proposed methodology provides a structured procedure to design an effective reliability ALT campaign. It consists of seven sequential steps, combining engineering expertise on the failure mechanisms affecting the components with the BOED theoretical framework and ALT models, for ensuring efficient allocation of resources while simultaneously enhancing understanding of the overall equipment reliability.

2.1. *Step 1: Component Failure Modes and Mechanisms Analysis*

The foundation stage of the proposed methodology is a comprehensive analysis of the equipment to identify the most critical components and their prevailing failure mechanisms, together with the stressing factors. The approach used integrates heterogeneous sources of knowledge, information and data (KID), such as:

(1) Reliability Handbooks and Technical Standards (e.g., International Electrotechnical Commission (2011, 2003); FIDES Group (2009); Telcordia Technologies (2011); U.S. Department of Defense (1995)), which provide methodological frameworks to identify both the components that are expected to be less reliable and those that are more sensitive to stressing operating conditions.

(2) A-priori knowledge and expert elicitation, which is leveraged to capture insights regarding component weaknesses and failure modes that are expected to affect the new technology.

(3) ReX data from as-similar-as possible operational equipment, which is analyzed to identify components that experience frequent failures in field operating conditions.

2.2. *Step 2: Selection of Reliability and Acceleration Models*

Based on Step 1, an appropriate mathematical model of the component failure behavior under stress is selected Pan (2009). Formally, the ALT model is defined by the probability density function (pdf) $f(t|\boldsymbol{\theta}, \mathbf{d})$ of the failure time t, where $\boldsymbol{\theta}$ is the vector of unknown model parameters, and $\mathbf{d}$ is the vector representing the test conditions. This pdf encodes two parts: (i) lifetime distribution (e.g., Weibull, Lognormal) to describe the random nature of failures at a given stress level, and (ii) life-stress relationship to adjust the failure behavior across different stress levels. This latter model heavily depends on the findings of Step 1. Specifically, if a critical weakness is associated with the intrinsic reliability of certain components accelerated by temperature stress, then the Arrhenius model can be applied Escobar and Meeker (2006). Yet, if the main concern lies in the production process and solder joints, then the Coffin-Manson model may be more appropriate IPC-SM-785 (1992).

2.3. *Step 3: Characterization of Experimental Censoring*

In an ideal ALT experiment, the exact failure time for every unit under test would be observed; in ALT practice, however, this is rarely achievable and a censoring mechanism is used. This yields test data with incomplete observation of failure times, which significantly affects the statistical model and limits the amount of information that can be extracted from a test.

The censoring mechanisms commonly used in ALT practice are:

- Right Censoring: due to the limited time and budget allocated for the test campaign, the experiment is conducted for a predetermined duration T. If a unit fails before the end of the test, its exact failure time is known. Otherwise, the true lifetime is only known to exceed the test duration. In this case, the likelihood for a single component reads:

$$p(t|\boldsymbol{\theta}, \mathbf{d}) = \begin{cases} f(t|\boldsymbol{\theta}, \mathbf{d}) & \text{if } t \leq T \\ 1 - \int_0^T f(\tau|\boldsymbol{\theta}, \mathbf{d})\, d\tau & \text{if } t > T \end{cases} \tag{1}$$

- Interval Censoring: due to technical constraints such as the inability to continuously power and monitor devices inside the testing chambers, units are inspected periodically every Δt units of time, at times $t_i = i \cdot \Delta t, i \in \{0, ...I\},\ I \cdot \Delta t \leq T$. Then, a failure is only known to have occurred within the interval between two inspections, so that the exact failure time remains unknown. The width Δt of the inspection interval represents a critical decision variable that affects the accuracy of the reliability estimate. In this case, the likelihood reads

$$p(t|\boldsymbol{\theta}, \mathbf{d}) = \begin{cases} \int_{t_i}^{t_{i+1}} f(\tau|\boldsymbol{\theta}, \mathbf{d})\, d\tau & \text{if } t_i < t \leq t_{i+1} \\ 1 - \int_0^T f(\tau|\boldsymbol{\theta}, \mathbf{d})\, d\tau & \text{if } t > T \end{cases} \tag{2}$$

- Complete Censoring: in case of more strict test facility limitations, inspection occurs only once, at the end of the test. This represents the least informative scenario, where the information is binary, indicating only whether the unit failed during the test period. In this case, the likelihood reads

$$p(t|\boldsymbol{\theta}, \mathbf{d}) = \begin{cases} \int_0^T f(\tau|\boldsymbol{\theta}, \mathbf{d})\, d\tau & \text{if } t \leq T \\ 1 - \int_0^T f(\tau|\boldsymbol{\theta}, \mathbf{d})\, d\tau & \text{if } t > T \end{cases} \tag{3}$$

The choice of the censoring scheme is the result of a trade-off between the cost and complexity of the ALT facility and the statistical precision of the resulting reliability estimates. Notice that experiment outcomes of the different components are always supposed to be statistically independent.

2.4. *Step 4: Elicitation of Prior Knowledge for Model Parameters*

Bayesian approaches incorporate existing knowledge into the analysis, through prior probability distributions van de Schoot et al. (2021); Jackman (2009); Cowles (2013). Before any new experiment is run, there is typically some existing information about the values of the unknown parameters θ in the acceleration models. This information can be sourced from the scientific literature, previous experiments on similar components, or with structured elicitation of expert knowledge, and is framed in terms of a prior distribution $p(\boldsymbol{\theta})$ for the vector of unknown parameters $\boldsymbol{\theta}$.

This step allows leveraging all available knowledge to make more informed decisions on the test design.

2.5. *Step 5: Formulation of the Constrained Optimization Problem*

In this step, the mathematical optimization problem is formulated. The objective is to identify the optimal experimental design that maximizes the expected information gain from the experiment, while satisfying real-world operational constraints. Let F denote the number of stressors to be considered, determined on the basis of the results of Step 1. The optimization problem can, then, be expressed as follows:

- **Decision Variables**: These correspond to the controllable parameters defining the experimental design, and include:
  (1) $L_j \in \mathbb{N}_+,\ j = 1, \ldots, F$: number of distinct stress levels for the *j-th* stressor (e.g., three temperature levels).
  (2) $s_{j,l} \in \mathbb{R}_+,\ j = 1, \ldots, F,\ l = 1, \ldots, L_j$: setting of the *l-th* stress level for the *j-th* stressor (e.g., 105 °C).
  (3) $n_{j,l} \in \mathbb{N}_+,\ j = 1, \ldots, F,\ l = 1, \ldots, L_j$: number of components to be tested under a specific stress condition.

  These variables are organized in the design vector

  $\mathbf{d} = (L_1, ...L_F, s_{1,1}, ..., s_{F,L_F}, n_{1,1}, ..., n_{F,L_F})$

- **Objective Function**: The design selection criterion is the maximization of the Expected Information Gain (EIG) Ryan et al. (2014); Chaloner and Verdinelli (1995). Specifically, the Information Gain, denoted as $U_{\mathcal{I}}$, quantifies the reduction in entropy (i.e., a measure of uncertainty) between the prior and posterior parameter distributions resulting from the collection of experimental time-to-failure data t:

$$U_{\mathcal{I}}(\mathbf{d}, t) = \mathbb{E}_{p(\boldsymbol{\theta}|\mathbf{d},t)}[\log p(\boldsymbol{\theta}|\mathbf{d}, t)] - \mathbb{E}_{p(\boldsymbol{\theta})}[\log p(\boldsymbol{\theta})] \tag{4}$$

  The EIG, denoted as $\mathcal{I}(\mathbf{d})$, associated with a design $\mathbf{d}$, quantifies the expected amount of information ($U_{\mathcal{I}}$) obtained by collecting experimental evidence under the design hypotheses $\mathbf{d}$. Formally, it is defined as the Kullback-Leibler (KL) Bernardo (1979); Lindley (1956) divergence between the posterior distribution $p(\boldsymbol{\theta}|t, \mathbf{d})$ to the prior distribution $p(\boldsymbol{\theta})$, averaged over all possible data outcomes t that the experiment could produce:

$$\mathcal{I}(\mathbf{d}) = \mathbb{E}_{p(t|\mathbf{d})}\left[\mathbb{E}_{p(\boldsymbol{\theta}|\mathbf{d},t)}[\log p(\boldsymbol{\theta}|\mathbf{d}, t)] - \mathbb{E}_{p(\boldsymbol{\theta})}[\log p(\boldsymbol{\theta})]\right] \tag{5}$$

  Intuitively, a test protocol with a large EIG is expected to yield highly informative data, resulting in a posterior distribution that is substantially more concentrated (and therefore less uncertain) than the prior. Formally, the objective function is expressed as:

$$\mathbf{d}^* = \underset{\mathbf{d} \in D}{\text{argmax}}\ \mathcal{I}(\mathbf{d})$$
$$D = \mathbb{N}_+^F \times \mathbb{R}_+^S \times \mathbb{N}_+^S$$

  where:

$$S = \sum_{j=1}^{F} L_j$$

- **Fixed parameters**: Additional to the decision variables above, the following parameters can influence the ALT campaign optimization:
  (1) $T \in \mathbb{R}_+$: Test duration.
  (2) $N_{max} \in \mathbb{N}_+$: maximum number of components to be tested.

(3) C: number of environmental chambers.

However, including these as decision variables would result in an ill-posed EIG optimization problem, with the obvious result of an unbounded solution, as EIG would increase indefinitely with these parameters. Indeed, since T, C and N_{max} heavily influence the test campaign costs, a multi-objective function would be needed to balance EIG against cost. This will be done in future research work. Therefore, we treat T, C and N_{max} as given parameters. This also reflects common industrial practice, where such variables are often predetermined by practical constraints. For instance, the test time T is typically dictated by the strict schedules and availability of testing facilities.

- **Constraints**: The optimization of the experimental design is subject to a set of practical constraints:

(1) Testing Capacity Constraint: the total number of unique test conditions (e.g., combinations of stress factors such as temperature and humidity) cannot exceed the number of available environmental chambers, C:

$$\sum_{j=1}^{F} \sum_{l=1}^{L_j} \mathbb{I}_{n_{j,l}>0} \leq C$$

where $\mathbb{I}_x$ is a Boolean indicator function, equal to 1 if x is true and 0 otherwise.

(2) Component Availability Constraint: the total number of components allocated across all test conditions must not exceed the maximum available, N_{max}:

$$\sum_{j=1}^{F} \sum_{l=1}^{L_j} n_{j,l} = N_{max}$$

(3) Stress-Condition Constraint: the applied stress levels, $s_{j,l}$ must remain within their feasible operational range, $R_{j,l}$. Specifically, the maximum allowed stress $s_{j,max}$. must avoid the activation of failure mechanisms that are unrealistic under normal operating conditions, ensuring valid extrapolation of results. The minimum stress must exceed the nominal design operating stress to guarantee meaningful acceleration degradation.

2.6. *Step 6: BOED solution algorithm*

The problem introduced in Step 5 can be formulated as an optimization task with a nested-expectation objective function (Eq.(5)), which is commonly addressed using stochastic gradient-based optimization methods Foster et al. (2020). Such methods are typically well-suited for identifying optimal design in high-dimensional and continuous design domains. In contrast, the optimization problem considered here is low-dimensional, under stringent constraints, and involving discrete decision variables. Given these conditions, a grid search over a predefined set of plausible candidate designs is a pragmatic and effective optimization strategy. The approach entails evaluating the EIG for each candidate design and selecting the one that maximizes it. Although computationally demanding, the method is straightforward to implement and guarantees identification of the optimal design within the candidate set, thereby providing a defensible compromise between theoretical sophistication and practical feasibility. Given that the associated computational cost remains manageable, we adopt the Nested Monte Carlo (NMC) estimator Foster et al. (2020), which ensures consistency provided that the number of inner and outer samples, M and K, is sufficiently large:

$$\hat{\mu}_{\text{NMC}}(\mathbf{d}) \triangleq \frac{1}{K} \sum_{k=1}^{K} \log \frac{p(t_k|\boldsymbol{\theta}_{k,0}, \mathbf{d})}{\frac{1}{M} \sum_{m=1}^{M} p(t_k|\boldsymbol{\theta}_{k,m}, \mathbf{d})}$$

where $\boldsymbol{\theta}_{k,m} \overset{\text{i.i.d.}}{\sim} p(\boldsymbol{\theta}),\ t_k \sim p(t|\boldsymbol{\theta} = \boldsymbol{\theta}_{k,0}, \mathbf{d})$.

2.7. *Step 7: Posterior Distribution Estimation*

The final step of the methodology occurs after the optimal experiment, as determined in Step 6, has been conducted and the data collected. The experimental evidence is used to update the state of knowledge about the model parameters $\boldsymbol{\theta}$, as described by the prior $p(\boldsymbol{\theta})$ elicited in Step 4. This is achieved using Bayes' formula Jackman (2009) to get the posterior distribution:

$$p(\boldsymbol{\theta}|\mathbf{d}^*, t) \propto p(t|\mathbf{d}^*, \boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}) \tag{6}$$

This posterior distribution provides the basis for all subsequent reliability predictions, such as calculating the mean-time-to-failure (MTTF) or failure rate under normal use conditions, now with significantly reduced uncertainty.

3. Case Study: Application to BTM

The proposed framework was applied to a Balise Transmission Module (BTM), which is a safety-critical on-board transmission unit of the European Railway Traffic Management System/European Train Control System (ERTMS/ETCS). BTM contributes directly to train traffic control, operational management and protection functions, with the BTM's primary role of processing uplink signals and telegrams exchanged with trackside infrastructure transmission devices (balises).

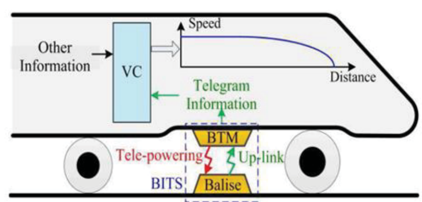

Fig. 1. Balise Transmission Module (BTM)

The BTM adopts a modular architecture, consisting of a limited set of Printed Circuit Boards (PCBs) configured as Line Replacement Units (LRUs) to enhance effective maintenance:

- redundant backplane unit (BLRB)
- transceiver (BLTX)
- signal receiver (BLRX), which integrates a Communication Processor & Radio Frequency Board (CPRF), and a CPU & Watchdog Board (BCPW).
- BTM power supply (BLPS).

For the present case study, the BTM is modeled as a series of non-redundant LRUs, where the failure of any individual LRU results in the failure of the entire equipment.

3.1. *Component Failure Modes and Mechanisms Analysis*

A preliminary reliability assessment was performed to identify the most critical LRUs, based on failure rates specified in the relevant IEC 62380:2004 International Electrotechnical Commission (2003). This investigation identified CPRF and BCPW as the LRUs with the highest failure rates.

Then, expert-based assessments indicated that LRU failures are predominantly driven by ceramic capacitor failures. Thus, BLRX unit and ceramic capacitors were selected for an ALT campaign.

Based on an analysis of the failure mechanisms of the components, temperature and humidity were identified as the primary environmental stressing factors. Elevated temperature accelerates degradation through mechanisms such as oxidation, ionic migration, dielectric creep and electromigration to metallization, whereas humidity promotes metallization corrosion, water ingress into dielectric or contacts, dendrite formation and dielectric breakdown Tucker et al. (2018); Reynolds (1977). Although the analysis also identified other potential stressing factors such as vibrations, thermal cycling, radiation, electromagnetic fields, chemical contamination and electrostatic discharge, the work presented here focuses on temperature and humidity for the sensitivity of electronic components to these stressing factors.

3.2. *Reliability ALT model*

The time to failure for each component was assumed to follow an exponential distribution $f(t|\boldsymbol{\theta}, \mathbf{d})$, parameterized by its failure rate $\lambda = \lambda(\boldsymbol{\theta}, \mathbf{d})$.

$$f(t|\boldsymbol{\theta}, \mathbf{d}) = \lambda(\boldsymbol{\theta}, \mathbf{d}) \cdot e^{-\lambda(\boldsymbol{\theta}, \mathbf{d}) \cdot t} \tag{7}$$

This distribution is commonly used in the reliability modeling of electronic components, based on the assumption that these devices satisfy the memory-less property for a long time period (Zio (2007)).

According to the findings of the previous step of analysis, the Peck ALT Peck (1986) model was selected to evaluate the reliability of components simultaneously subjected to humidity and temperature stresses. This model extends the Arrhenius model Escobar and Meeker (2006), which is commonly used for failure mechanisms driven by chemical and physical reactions that are strongly temperature dependent. The accelerated failure rate, $\lambda_i(\boldsymbol{\theta}, \mathbf{d})$, is estimated as the product of the baseline failure rate at reference conditions, θ_0, and the acceleration factor (ACC):

$$\begin{aligned}\lambda_i(\boldsymbol{\theta}, \mathbf{d}) &= \theta_0 \cdot ACC \\ &= \theta_0 \cdot \left[\left(\sum_{l=1}^{L_1} \frac{S_{1,l} \cdot \mathbb{I}_{i,l}(n_{1,l})}{H_{\text{ref}}} \right)^{\theta_1} \right] \cdot \\ &\quad \left[e^{\theta_2 \cdot \left(\sum_{l=1}^{L_2} \frac{1}{S_{2,l} \cdot \mathbb{I}_{i,l}(n_{2,l})} - \frac{1}{\rho_{ref}} \right)} \right]\end{aligned} \tag{8}$$

where

- $\boldsymbol{\theta} = [\theta_0, \theta_1, \theta_2]$;
- θ_1: $n \approx 2.7$, humidity acceleration exponent;
- θ_2: $\frac{E_a}{K} \approx 1.277 \times 10^4$;
- E_a is the effective activation energy [eV];
- K is the Boltzmann constant (8.62×10^{-5});
- $L_1 = 2$;
- $S_{1,\{1,2\}} = [0.5,\ 0.9]$;
- $H_{\text{ref}} = 0.5$: Humidity level during accelerated testing and under normal operating conditions.
- $L_2 = 3$,
- $S_{2,\{1,2,3\}} = [328,\ 353,\ 378]K$;
- $\rho_{\text{ref}} = 303K$;
- $\mathbb{I}_{i,l}(n_{1,l}),\ \mathbb{I}_{i,l}(n_{2,l})$ denote the setting condition according to the design. To break symmetry and simplify, we assume that items are ordered: the first $n_{1,1}$ items are set to $S_{1,1}$, and so on.

3.3. *Experimental Censoring*

In this case study, technical constraints defining the censoring scheme are not considered, and all three censoring schemes were evaluated to assess the impact of this choice on the results.

For the considered ALT model, Eqs. 1, 2, 3 become, respectively:

- Right Censoring:

$$p(t|\boldsymbol{\theta}, \mathbf{d}) = \begin{cases} \lambda(\boldsymbol{\theta}, \mathbf{d}) \cdot e^{-\lambda(\boldsymbol{\theta}, \mathbf{d}) \cdot t} & \text{if } t \leq T \\ e^{-\lambda(\boldsymbol{\theta}, \mathbf{d}) \cdot t} & \text{if } t > T \end{cases} \tag{9}$$

- Interval Censoring, where $t_i = i \cdot \Delta t,\ i \in \{0, \ldots, I\},\ I \cdot \Delta t \leq T$

$$p(t|\boldsymbol{\theta}, \mathbf{d}) = \begin{cases} \lambda(\boldsymbol{\theta}, \mathbf{d}) \cdot (e^{-\lambda(\boldsymbol{\theta}, \mathbf{d}) \cdot i\Delta t} \\ -e^{-\lambda(\boldsymbol{\theta}, \mathbf{d}) \cdot (i+1)i\Delta t}) & \text{if } t_i < t \leq t_{i+1} \\ e^{-\lambda(\boldsymbol{\theta}, \mathbf{d}) \cdot t} & \text{if } t > T \end{cases} \tag{10}$$

- Complete Censoring

$$p(t|\boldsymbol{\theta}, \mathbf{d}) = \begin{cases} 1 - e^{\lambda(\boldsymbol{\theta}, \mathbf{d}) \cdot t} & \text{if } t \leq T \\ e^{-\lambda(\boldsymbol{\theta}, \mathbf{d}) \cdot t} & \text{if } t > T \end{cases} \tag{11}$$

3.4. *Model parameters elicitation*

For the prior distributions of the parameter vector $\boldsymbol{\theta}$, the baseline failure rate of the capacitor ($\theta_{0,\text{res}}$) and the BLRX unit ($\theta_{0,\text{blrx}}$), as well as the uncertain Peck model parameters θ_1 and θ_2, are modeled by a Gamma distribution:

$$\Gamma(x; \alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha - 1} e^{-\beta x}, \quad x > 0 \tag{12}$$

The choice of the Gamma distribution is motivated by its strictly positive support and its flexibility to capture unimodal behaviors, particularly for large values of the shape parameter. The prior mean of θ_0 is set equal to the expected component failure rate, as prescribed by the reliability IEC-62380 standard International Electrotechnical Commission (2003), whereas the prior means of θ_1 and θ_2 are determined according to established values reported in the Peck–Arrhenius literature (Hernández-López et al. (2018); Escobar and Meeker (2006); Peck (1986)).

A shape parameter $\alpha = 10$ is assigned to both θ_0 and θ_2 to encode a moderate degree of prior confidence, which ensures a coefficient of variation of approximately 30% of the mean, thus quantifying the uncertainty around the expected failure rate. For θ_1, a larger prior uncertainty is assumed, with a standard deviation equal to 50% of the mean, reflecting the limited confidence in the corresponding literature-based estimate.

Then, $p(\theta_{\text{res}}) = p(\theta_{0,\text{res}}) \cdot p(\theta_1) \cdot p(\theta_2)$ and $p(\theta_{\text{BLRX}}) = p(\theta_{0,\text{BLRX}}) \cdot p(\theta_1) \cdot p(\theta_2)$.

$$\theta_{0,\text{res}} \sim \Gamma\Big(\alpha = 10,\ \beta = \frac{10}{3.19 \times 10^{-8}}\Big) \tag{13a}$$

$$\theta_{0,\text{BLRX}} \sim \Gamma\Big(\alpha = 10,\ \beta = \frac{10}{1.16 \times 10^{-6}}\Big) \tag{13b}$$

$$\theta_2 \sim \Gamma\Big(\alpha = 10, \beta = \frac{10}{1.277 \times 10^{4}}\Big) \tag{14}$$

$$\theta_1 \sim \Gamma\Big(\alpha = 5, \beta = \frac{5}{2.7}\Big) \tag{15}$$

3.5. *Optimal Test Design - Peck Model*

The constraint decision variables and the fixed parameters of the optimization problem are listed in Table 1.

Table 1. Optimization Problem constraints.

COMPONENT	C	N_{max}	$T[h]$
CAPACITOR	3	100	1000
BLRX unit	3	5	1000

3.6. *Nested Monte Carlo estimate*

The EIG was evaluated for 10 candidate designs, which varied in the allocation of test units across different stress levels (Table 2). For the EIG estimation, the inner (K) and outer (M) Monte Carlo sample sizes were set to 20.000 and 2.000, respectively. The computational cost amounted to a few minutes per candidate design on a standard laptop equipped with an 11th-generation Intel i7 processor.

3.7. *Posterior distribution estimation*

Data from the experimental results to be used for Bayesian updating to obtain the parameters' posterior distributions were not available at the time of the writing of the paper.

4. Results

The analysis reveals several key findings. First, a clear hierarchy of informativeness emerges across censoring schemes: as expected, right-censoring provides the largest information (highest EIG), followed by frequent interval censoring, whereas complete censoring provides the least. Future research work will focus on developing Value of Information cost models to justify the value of investing in upgrading the technology to switch from one monitoring technology to one allowing for more precise information. Second, the optimal design identified across the investigated test configurations is not the trivial allocation in which all units are placed at the highest stress level ($n_{2,3} = 100\%$). Instead, the ALT design strategy $n_{1,3} = 40\%,\ n_{2,3} = 60\%$ of N_{max} outperforms the trivial design. These findings suggest that (i) maximizing the number of component failures within the test duration is not necessarily the most effective strategy, and (ii) distributing units across two high-stress levels may yield more information than focusing all experimental units at a single extreme condition. This can be explained by the fact that allocating a subset of units to a second high-stress level enables comparison across samples with distinct failure rates induced by different stress conditions, thereby improving the estimation of the parameters of the ALT model. This effect is more evident under the complete censored scheme, where the optimal ALT design recommends allocating all components across multiple stress levels to mitigate information loss induced by censoring. Another notable finding is that while the highest temperature level is consistently prioritized in optimal designs, the optimal humidity setting appears less clear-cut. This can be attributed to the comparatively lower prior confidence associated with the humidity coefficient relative to the temperature coefficient. Overall, these outcomes highlight a more complex interaction between the stressors, which the BOED framework effectively exploits to identify the most informative test plan.

5. Conclusion & direction for future research

This paper presented a methodology for the strategic planning of ALT for electronic components, responding the industry need of optimally designing the tests by explicitly considering relevant degradation mechanisms and operational constraints (e.g., number of devices and test chambers). By prioritizing designs that maximize the information gained from ALT campaigns, EIG offers a straightforward ranking of candidate test plans.

The case study on the BTM illustrated the feasibility of application of the methodology under a specific life–stress model. Notably, the results highlight that the optimal test design is not general: it depends on the failure mechanisms under investigation, the prior uncertainty in the model parameters and operational limitations. The implementation of the recommended optimal tests facilitates updating the prior distributions of the ALT model, which in turn informs the selection of subsequent optimal designs. Changes in the underlying failure mechanism require the adoption of an alternative ALT model and, consequently, the corresponding EIG evaluation. For instance, in the assessment of solder joint reliability in electronic components and circuit boards, the Engelmaier–Wild model — capturing thermo-mechanical fatigue — would be more appropriate IPC-SM-785 (1992). This need of adaptability underscores the limitations of standardized "one-size-fits-all" testing protocols and highlights the advantages of the proposed methodology.

The analysis performed in the case study quantitatively

Table 2. EIG Results for Peck Model Test Designs

Design configuration	**EIG**				
	Right Censoring	**Complete Censoring**	**Δt: 24h**	**Δt: 72h**	**Δt: 168h**
capacitor screening test					
$n_{2,3} = 100$	2.83	1.96	2.69	2.62	2.51
$n_{1,3} = 20,\ n_{2,3} = 80$	3.32	2.37	3.18	3.08	2.98
$n_{2,2} = 10,\ n_{2,3} = 90$	2.81	2.16	2.74	2.67	2.62
$n_{2,2} = 20,\ n_{1,3} = 20,\ n_{2,3} = 60$	3.30	**2.57**	3.14	3.09	3.03
$n_{1,3} = 40,\ n_{2,3} = 60$	**3.42**	2.46	**3.26**	**3.18**	**3.08**
$n_{2,2} = 60,\ n_{2,3} = 40$	2.75	2.18	2.59	2.59	2.51
$n_{2,1} = 20,\ n_{2,2} = 40,\ n_{1,3} = 40$	2.73	2.28	2.64	2.63	2.57
$n_{2,2} = 30,\ n_{1,3} = 20,\ n_{2,3} = 50$	3.24	**2.57**	3.11	3.08	3.01
$n_{1,2} = 30,\ n_{1,3} = 20,\ n_{2,3} = 50$	3.17	2.42	3.03	2.96	2.89
$n_{1,2} = 30,\ n_{2,2} = 20,\ n_{1,3} = 50$	2.60	2.26	2.57	2.52	2.50
BLRX unit screening test					
$n_{2,3} = 5$	1.75	0.32	1.350	1.150	0.960
$n_{2,2} = 1,\ n_{1,3} = 1,\ n_{2,3} = 3$	1.80	0.55	1.540	1.380	1.210
$n_{2,2} = 1,\ n_{1,3} = 2,\ n_{2,3} = 2$	1.83	0.61	**1.600**	**1.440**	1.290
$n_{2,2} = 2,\ n_{1,3} = 1,\ n_{2,3} = 2$	1.76	0.62	1.540	1.420	1.270
$n_{1,2} = 2, n_{1,3} = 1, n_{2,3} = 2$	1.67	0.72	1.480	1.380	1.260
$n_{2,1} = 1, n_{1,3} = 2, n_{2,3} = 2$	1.77	0.63	1.530	1.390	1.240
$n_{1,1} = 1, n_{1,3} = 2, n_{2,3} = 2$	1.73	0.54	1.460	1.300	1.140
$n_{1,2} = 1, n_{2,2} = 1, n_{2,3} = 3$	1.67	0.62	1.450	1.330	1.190
$n_{1,3} = 2, n_{2,3} = 3$	**1.88**	0.52	1.560	1.380	1.190
$n_{2,1} = 1, n_{2,2} = 1, n_{1,3} = 3$	1.66	**0.74**	1.520	1.420	**1.300**

Note: For the Peck Model, the $n_{i,j}$ indexes refers respectively to the humidity ($i \in \{1,2\} \longleftrightarrow S_{1,\{1,2\}}$) and temperature ($j \in \{1,2,3\} \longleftrightarrow S_{2,\{1,2,3\}}$) stress levels.

demonstrates the significant information loss incurred under data-censoring schemes driven by facilities equipped with limited monitoring capabilities. This highlights the importance of enhanced test monitoring for reducing uncertainty in reliability estimates.

Future research directions include extending the methodology to address extrinsic failures arising from complex system-level interactions: while the current framework is primarily concerned with intrinsic component-level failures, subsequent work will address thermal, electrical and mechanical interactions among components.

Additionally, the case study employed a grid search for optimization, suitable for low-dimensional, discrete design spaces. For more complex scenarios involving continuous design variables (e.g., optimizing a continuous temperature ramp profile), an unbiased stochastic gradient-based optimization approach Foster et al. (2020) could be integrated. Finally, the present methodology considers only the maximization of EIG as an objective. A further extension will involve formulating a multi-objective optimization problem Guo et al. (2025) that explicitly accounts for test cost and duration, along with EIG. This will enable the generation of Pareto-optimal solutions, providing decision-makers with a spectrum of designs that optimally balance information acquisition, cost and time.

References

Amzal, B., F. Y. Bois, E. Parent, and C. P. Robert (2006). Bayesian-optimal design via interacting particle systems. Journal of the American Statistical Association 101(474), 773–785.

Bernardo, J. M. (1979). Expected information as expected utility. The Annals of Statistics 7(3), 686–690.

Chaloner, K. and I. Verdinelli (1995). Bayesian experimental design: A review. Statistical Science 10(3), 273–304.

Cowles, M. (2013). Applied Bayesian Statistics: With R and OpenBUGS Examples. New York, NY: Springer.

Dodson, B. and H. Schwab (2021). Accelerated testing: a practitioner's guide to accelerated and reliability testing. SAE international.

Escobar, L. A. and W. Q. Meeker (2006, November). A review of accelerated test models. Statistical Science 21(4).

FIDES Group (2009). FIDES Guide 2009: A Reliability Methodology for Electronic Systems. Updated in 2022.

Foster, A., M. Jankowiak, M. O'Meara, Y. W. Teh, and T. Rainforth (2020). A unified stochastic gradient approach to designing bayesian-optimal experiments. In Proceedings of the Twenty-Third International Conference on Artificial Intelligence and Statistics, Volume 108 of Proceedings of Machine Learning Research, pp. 2959–2969. PMLR.

Gaier, A. and D. Ha (2019). Weight agnostic neural networks. In Advances in Neural Information Processing Systems, Volume 32, pp. 5364–5378. Curran Associates, Inc.

Goda, T., T. Hironaka, W. Kitade, and A. Foster (2022). Unbiased mlmc stochastic gradient-based optimization of bayesian experimental designs. SIAM Journal on Scientific Computing 44(1), A286–A311.

Guo, J., Z. Han, S. Wu, J. Long, C. Chen, and Z. Liu (2025, apr). Multi-objective optimization design for accelerated degradation test based on game theory. Scientific Reports 15(1), 14707.

Hernández-López, A. M., J. A. Aguilar-Garib, S. Guillemet-Fritsch, R. Nava-Quintero, P. Dufour, C. Tenailleau, B. Durand, and Z. Valdez-Nava (2018). Reliability of x7r multilayer ceramic capacitors during high accelerated life testing (halt). Materials 11(10).

Indmeskine, F.-E., L. Saintis, and A. Kobi (2023). Review on accelerated life testing plan to develop predictive reliability models for electronic components based on design-of-experiments. Quality and Reliability Engineering International 39(6), 2594–2607.

International Electrotechnical Commission (2003). IEC 62380: Reliability data handbook - Universal model for reliability prediction of electronics components, PCBs and equipment. Also published as UTE C 80810.

International Electrotechnical Commission (2011). IEC 61709: Electric components - Reliability - Reference conditions for failure rates and stress models for conversion. Edition 2.0.

IPC-SM-785 (1992). Guidelines for accelerated reliability testing of surface mount solder attachments. IPC Northbrook, Illinois.

Jackman, S. (2009). Bayesian Analysis for the Social Sciences. Wiley.

Jang, I. and C. H. Kim (2024). Prediction of remaining useful life (rul) of electronic components in the posafe-q plc platform under npp dynamic stress conditions. Nuclear Engineering and Technology 56(5), 1863–1873.

Li, X., Y. Hu, R. Li, and L. Liu (2016). Bayesian optimal design of sequential stress accelerated degradation testing. In 2016 Prognostics and System Health Management Conference (PHM-Chengdu), pp. 1–6.

Li, X., Y. Hu, F. Sun, and R. Kang (2017). A bayesian optimal design for sequential accelerated degradation testing. Entropy 19.

Li, X., Y. Hu, E. Zio, and R. Kang (2017). A bayesian optimal design for accelerated degradation testing based on the inverse gaussian process. IEEE Access 5, 5690–5701.

Li, X., M. Rezvanizaniani, Z. Ge, M. Abuali, and J. Lee (2015). Bayesian optimal design of step stress accelerated degradation testing. Journal of Systems Engineering and Electronics 26(3), 502–513.

Lindley, D. V. (1956). On a measure of the information provided by an experiment. The Annals of Mathematical Statistics 27(4), 986–1005.

Meeker, W. Q. and L. A. Escobar (1993). A review of recent research and current issues in accelerated testing. International Statistical Review / Revue Internationale de Statistique 61(1), 147–168.

Nelson, W. B. (2005a). A bibliography of accelerated test plans. IEEE Transactions on Reliability 54(2), 194–197.

Nelson, W. B. (2005b, sep). A bibliography of accelerated test plans part ii—references. IEEE Transactions on Reliability 54(3), 370–373.

Pan, R. (2009). A bayes approach to reliability prediction utilizing data from accelerated life test sand field failure observations. Quality and Reliability Engineering International 25(2), 229–240.

Peck, S. (1986). Comprehensive model for humidity testing correlation. 24th International Reliability Physics Symposium, 44–50.

Reynolds, F. H. (1977). Accelerated-test procedures for semiconductor components. In 15th International Reliability Physics Symposium, pp. 166–178.

Ryan, E. G., C. C. Drovandi, J. M. McGree, and A. N. Pettitt (2016). A review of modern computational algorithms for bayesian optimal design. International Statistical Review 84(1), 128–154.

Ryan, E. G., C. C. Drovandi, M. H. Thompson, and A. N. Pettitt (2014). Towards bayesian experimental design for nonlinear models that require a large number of sampling times. Computational Statistics & Data Analysis 70, 45–60.

Telcordia Technologies (2011). SR-332: Reliability Prediction Procedure for Electronic Equipment. Issue 3.

Temsamani, A., S. Kauffmann, Y. Descas, B. Vandevelde, F. Zanon, and G. Willems (2017). Improved and accurate physics-of-failure (pof) methodology for qualification and lifetime assessment of electronic systems. Microelectronics Reliability 76-77, 42–46.

Tucker, J., R. Dhakal, G. Thiel, and V. Jadhav (2018). A monte carlo approach to predicting failure across multiple temperature and humidity field environments. In 2018 IEEE 68th Electronic Components and Technology Conference (ECTC), pp. 144–149.

U.S. Department of Defense (1995). MIL-HDBK-217F: Reliability Prediction of Electronic Equipment. Notice 2, Canceled and superseded by an ANSI standard in 2006, but still widely used.

van de Schoot, R., S. Depaoli, R. King, B. Kramer, K. Märtens, M. G. Tadesse, M. Vannucci, A. Gelman, D. Veen, J. Willemsen, and C. Yau (2021). Bayesian statistics and modelling. Nature Reviews Methods Primers 1(1), 1.

Wan, F., C. Zhang, Y.Tan, and X. Chen (2014). Data analysis and reliability estimation of step-down stress accelerated degradation test based on wiener process. In 2014 Prognostics and System Health Management Conference (PHM-Hunan).

Wang, X. (2010). Wiener processes with random effects for degradation data. Journal of Multivariate Analysis 101(2), 340–351.

Wang, X. and D. Xu (2010). An inverse gaussian process model for degradation data. Technometrics 52(2), 188–197.

Xu, A. and Y. Tang (2015). A bayesian method for planning accelerated life testing. IEEE Transactions on Reliability 64(4), 1383–1392.

Zhang, Y. and W. Q. Meeker (2006). Bayesian methods for planning accelerated life tests. Technometrics 48(1), 49–60.

Zio, E. (2007). An Introduction to the Basics of Reliability and Risk Analysis. Singapore: World Scientific.