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-p26322-cd

Reliability and Quality assurance methods for power electronics converters

Edoardo Martino[1 *], Arooj Akbar[1], Vadym Denysenko[1], Hua Shen[1], Adrien Maillet-Gonzalez[1 3], Lucas Radon[1,4], Gernot J. Riedel[2], Jürgen Schuderer[2]

1 Hitachi Energy Research, 5405 Baden -Dättwil, Switzerland
2 Hitachi Energy Ltd Switzerland, 5300 Turgi, Switzerland
3 École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
4 Institute for Power Electronics and Electrical Drives, RWTH Aachen University, 52074 Aachen, Germany

The strong acceleration of energy system electrification and the growing requirements for a resilient power grid, place power electronics as a fundamental enabling technology. The grid is evolving to become a digital system of systems where the critical function of transmission, control and protection is shifting from passive components to active Electronic Components and Systems (ECSs). Prominent examples in power transmission are the massive introduction of HVDC interconnections and STATCOMs, based on the Modular Multilevel Converter topology. These power electronics designs are built from a large quantity of electrically interconnected identical power electronics modules. ECSs in such equipment are critical parts of the power grid infrastructure and expected to satisfy the strict reliability and availability requirements of grid utilities. This is challenged by the complex supply chain of microelectronic components used in ECSs, where short product lifecycles and material changes for cost or regulatory motivation frequently occur. We present the foundational work for the implementation of Quality and Reliability assurance methods for the building blocks of power electronics converters, using a combination of design-for-reliability, reliability growth, validation and lifetime estimation by (highly) accelerated testing, stress screening inspection and digital product tracking designed to enable advanced data analytics.
This contribution is intended to outline quality and reliability process implementation for mission critical power electronics components, where novel methods for physics-informed deep learning and advanced data analytics can fundamentally enhance the knowledge creation out of continuous product testing and lead to optimized product lifecycle management. The discussion will also cover the challenges of quantitative knowledge generation out of highly accelerated testing, conceptually designed to only be a qualitative source of product validation.

Keywords: Accelerated Testing; Reliability Assurance; Quality Assurance; Physics-Informed Deep Learning

1. Introduction

The rapid electrification of energy systems and the need for a resilient, controllable grid are accelerating the deployment of power electronics across transmission and distribution. The grid is evolving into a digital “system of systems,” where critical functions such as transmission control, protection, and reactive power control increasingly depend on active Electronic Components and Systems (ECSs). In electrical power transmission, High-Voltage Direct Current (HVDC) interconnections are being built to enable exchange of power between geographic distant locations in ways that AC transmission cannot economically provide. Flexible AC Transmission Systems (FACTS) support the integration of renewable energy sources by mitigating the stability challenges that arise from replacing conventional rotating generators. Meanwhile, Battery Energy Storage Systems (BESS) help balance temporal mismatches between electricity production and demand. At the grid’s receiving end, applications such as e-mobility charging infrastructure and data-center grid-to-rack distribution systems are emerging as major growth areas for modular ECS solutions within the electrical power domain. All those power electronics systems share the fundamental design of being constituted by a large number of identical and electrically interconnected building blocks where the weakest components would define the overall system reliability (Hosseinabadi, et al. 2024).

To provide a concrete example we will take the Modular Multilevel Converters (MMC) topology as a reference we will investigate in this article. The modularity enabled flexibility in the creation converters with different rating and functionalities, while using the same building block and intrinsically enhancing system reliability thanks to redundancies (Nami, et al. 2014). The MMC converter simplified diagram is shown in Figure 1, where the building block, the MMC cell, is highlighted. The converter arm is made of a large number of series connected cells, whose number depends on the operating condition requirements. A certain number of additional cells can be added for redundancy to increase system reliability.

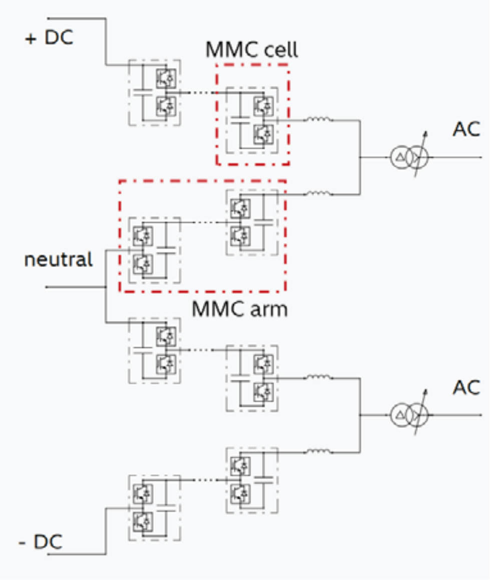

Figure 1. Simplified line diagram of a Modular Multilevel Converter (MMC) for a bipolar AC to DC converter.

In the discussion of power electronics reliability together with the system-level functional description that can be understood by converter level line diagram, an effective physics-based understanding requires a description of the key components (Edoardo Martino 2023). The example of MMC in Figure 2 schematically represents the key cell components: power semiconductor module, DC link capacitors, control electronics and heat dissipation system. Each component is identically repeated per cell, except for the heat dissipation that can be a shared system level system. The electronic components include microelectronics and circuit boards mounted near the power semiconductors modules, and are used for converter control, measurement, and communication. The different functions and nature of each of these components results in different ageing and failure mechanisms, which necessitate different strategies for reliability investigation and expected knowledge to be gained.

Low-voltage ECSs—gate drivers, auxiliary power supplies, measurement and control boards—are mission-critical within these converter cells, and their reliability is being challenged by complex, supply chains with short lifecycles and common material/process changes.

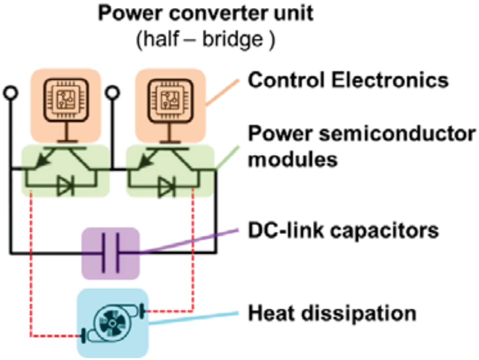

Figure 2. Power electronics building block, using a half-bridge converter as example. Conceptual representation of an MMC converter cell.

This article outlines a pragmatic, physics-informed framework for Quality and Reliability (Q&R) assurance of the converter building blocks, combining highly accelerated testing, stress screening, and digital product tracking to enable advanced analytics. It focuses on modular power electronics devices, where redundancies can be built-in the converter topology, and each cell reliability performance significantly impacts availability, and it argues that continuous, instrumented testing to failure, coupled with physics-of-failure (PoF) analysis and data-driven methods can optimize lifecycle management while addressing the well-known difficulty of extracting quantitative lifetime knowledge from highly accelerated tests (Pandey, et al. 2016). As power-electronics-dominated energy systems continue to expand, ensuring the long-term reliability of converter building blocks becomes essential for grid resilience, economic sustainability, and safe operation. This paper responds to this challenge by presenting an integrated reliability framework that spans the full spectrum of modern practices: from classical reliability methods to advanced physics-informed approaches. Specifically, it examines conventional and accelerated testing methodologies, provides a PoF-based understanding of ageing mechanisms, and discusses Design for Reliability (DfR) strategies that enhance robustness at both component and system levels.

2. Power electronics reliability: state of the art

Predicting the lifecycle of power electronic systems presents significant challenges due to their demanding operational requirements. These systems are often designed for long operational lives exceeding 40 years, which makes lifetime estimation from accelerated testing and modelling inherently complex. Furthermore, they operate under high duty cycles and stress conditions, where the maximum achievable acceleration factor during accelerated life testing is typically limited to 2–5. This relatively low acceleration factor restricts the ability to replicate decades of ageing within practical test durations, thereby increasing uncertainty in lifetime predictions and reliability modelling.

Historically, maintenance started as corrective maintenance, where the system was serviced and repaired after a failure occurred. This approach, while straightforward, leads to unexpected downtimes and increased costs. Subsequently, time-based maintenance emerged, where maintenance is performed at scheduled intervals regardless of the system condition. Although this method reduced unexpected failures, it was still not optimal because it did not account for the actual wear and tear of the components. With the advancement of sensor technologies, condition-based maintenance (CBM) became the new paradigm. CBM involves collecting sensor data from the system, deriving health indicators, and performing maintenance actions based on these indicators. Despite its effectiveness, CBM does not predict future failures, leading to the development of Prognostics and Health Management (PHM). One feature of PHM is the prediction of Remaining Useful Lifetime (RUL) of components by analysing the health indicators and forecasting future degradation (Kirill Ivanov 2025). For power electronics, two developing reliability strategies can be considered: PHM and zero-failures tolerance.

In PHM the components may degrade over time, but the system continuously monitors their health so that any developing problem is detected early and corrected before it leads to failure. This requires active effort—power, cost, and data processing—and therefore increases operational expenses (OPEX). PHM is useful when the system must keep operating even as degradation occurs.

The second strategy is zero-failures tolerance, where the system is designed so robustly that failures are not expected to occur during normal operation. High-quality components, strong design margins and effective redundancies allow the system to operate with very high availability and without additional OPEX for active monitoring. Data-driven methods are used during product design, development and production testing to achieve this “zero-failures” behavior from the start.

2.1 Design for Reliability

Design for Reliability (DfR) is a systematic approach that integrates reliability considerations into the product development lifecycle (Raheja n.d.). Reliability and quality assurance play a critical role throughout the product lifecycle, from initial design to final production. The product lifecycle spans the design phase, prototype phase and production phase. Each phase should incorporate specific activities aimed at improving the reliability of the product.

The product design phase involves conceptualizing and developing detailed plans and specifications to meet user needs and business goals, ensuring functionality, mechanical design and manufacturability. The prototype phase is focused on validating design assumptions and identifying potential reliability issues in a physical model. The functional prototypes are tested to verify that the design conforms to specifications. During this phase, reliability testing methods are used to identify weaknesses and improve design robustness. This is also the final opportunity to identify issues in design, quality, reliability, manufacturing, test, and supplier issues before the design is released for production.

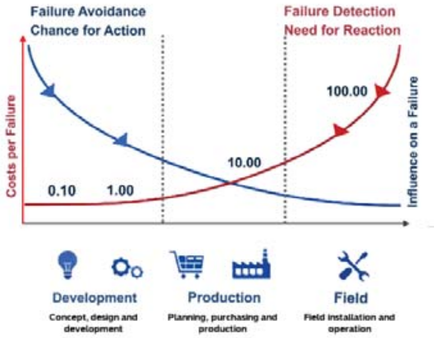

Figure 3. Relationship between cost for failure and product lifetime. Late identification of reliability issues has significant cost penalty, with an approximate scale by factor 10 and every consequent phase in the product lifecycle.

In power electronics, DfR aims to ensure that components such as gate drivers, power supplies, and control boards meet stringent operational requirements under diverse environmental and electrical stresses (Huai Wang 2012). The concept revolves around anticipating failure mechanisms early in the design phase, thereby reducing costly redesigns and field failures. The challenges in implementing DfR for power electronics are multifaceted. A complex supply chain, characterized by material changes and short component lifecycles, introduces variability that complicates reliability predictions. Accurate lifetime modelling further demands a deep understanding of physics-of-failure principles, including degradation signatures and failure precursors, which are often difficult to capture comprehensively. Data scarcity adds another layer of complexity, as limited structured reliability data hampers predictive modelling and quantification efforts.

Predictive modelling and accelerated testing help overcome the complexity of supply chains and variability in materials by providing early insights into potential failure modes using Highly Accelerated Lifetime Testing (HALT). By combining predictive analytics with physics-informed models, engineers can estimate the time-to-failure and establish reliability confidence intervals with greater accuracy.

Reliability Growth Management (RGM) provides a tool to address DfR on the development and prototyping of power electronics devices. RGM transforms DfR from a one-time design activity into a dynamic, data-driven process, that can be applied to an operational and repairable demonstrator. Product improvement is achieved by identifying failures during testing and performing repairs to resume operation until the next failure event is triggered. Consecutive repairs and change in materials are done until a new design release is necessary to continue to cycle of testing and repair. RGM supports early identification of reliability issues and demonstration of product reliability and maturity level.

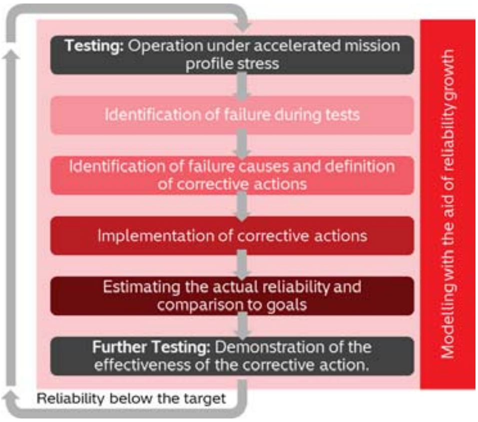

Figure 4. Iterative principle of reliability growth management.

Mathematical models are developed to provide an estimated value for the product reliability at every step of the iteration loop of testing and improvement. The reliability figure used is the estimated Mean Time Between Failures (MTBF) computed as ratio between operational time and number of failures, is used to assess if the product has reached the final reliability target. Because such methodology was originally developed for mechanical systems such as vehicles (Crow 2008), in this case the operational time is either engine running hours or converted in kilometers driven. The application of such concept to power electronics prototype development requires to review the "time" variable most suitable for estimation of MTBF to relate with a functional reliability target for the final product.

To estimate overall reliability and availability of the complete operating power converter, the system must be represented in a modeling framework that captures how reliability flows through the architecture. Reliability Block Diagrams (RBDs) are commonly used for this purpose (Gissler and Shrivastava 2015), also defined in the IEC 61078 reliability block diagram standard. More advanced techniques such as Monte-Carlo simulation, Markov chains, and Petri Nets can model systems where the failure probability distribution is considered and to include in the model repair, switching behavior, or complex state transitions.

System level modelling of the entire converter can guide identifying critical components, intended as those whose failure has the greatest effect on system reliability. Once identified, these components' reliability can be studied in more detail to quantify their unique failure probability distributions and life-stress relationships. Such relationships are derived from laboratory ageing tests or long-term field data. By combining component-level behavior and system-level modeling, engineers can simulate overall system lifetime with greater accuracy. As the reliability of the most critical component will strongly affect the overall system failure probability, is useful to consider the three primary categories of failure modes a component might show on how they affect the overall system reliability. For generalization and simplification, we will refer to them as: *the Bad*, *the Ugly* and *the Good*.

"*The Bad*" refers to components failure mechanism that appears randomly distributed in time, unpredictable and with no identifiable degradation mechanism. This failure mode does not allow redundancies, such as the case of destructive events like explosions, where component failure equals system failure. For modular systems built by N identical components, such failure mode would imply that the more components the higher the probability of system failure, and no effective redundancy can be created.

"*The Ugly*" describes a situation where is not possible to predict failures or track degradation, as in the case of "The Bad", but the single component failure does not cause a system failure allowing for redundancy. For a system with N identical components, the redundancy can be described by k components required to operate the system ($k < n$), resulting in a *k-out-of-n* redundancy.

"*The Good*" represents the ideal case, where a physics-based understanding of the failure mechanism has been developed. By uncovering the root causes and stress-dependent behaviors that drive failure, engineers can describe the resulting failure probability distribution, refine predictive lifetime models, and guide more effective design-for-reliability and qualification strategies. Thus, transitioning from "Bad/Ugly" randomness to "Good" physics-informed reliability enables more accurate lifetime prediction, better economic decisions, and improved system robustness. As the case above, a single failure has not caused a system level failure, allowing for a *k-out-of-n* redundancy.

Table 1. Failure description categories per critical components.

	The Bad	***The Ugly***	***The Good***
Degradation	Non evident	Non evident	Measurable
Failure prediction	Not possible	Not possible	Possible
Failure probability	Random / Early failures	Random / Early failures	Growing in time (ageing)
Redundancy possible	No	Yes	Yes

3. Quality and reliability testing methods: a structured summary

A practical reliability program integrates tests for qualitative overstress discovery, quantitative lifetime modelling, and production screening (Edoardo Martino 2023), with data capture and Failure Reporting, Analysis, and Corrective Action System (FRACAS) for documenting the learning loops and quantify reliability during product lifecycle. Fundamental success for a systematic reliability growth in product development is the rigorous failure reporting, root-cause tracking, and corrective actions linked to product history, enable reliability growth across design–prototype–production phases (IEEE 2012).

3.1. Reliability Quantification

Reliability quantification aims to establish measurable confidence in a system’s ability to perform its intended function over its expected lifetime. This section outlines the methods used to determine reliability levels, relying on accelerated stress testing, probabilistic lifetime modeling, and statistical approaches for rare-event characterization. Quantification process begins with the identification and strengthening of design margins through Highly Accelerated Lifetime Testing (HALT), as graphically shown in Figure 5 and Figure 6, where products are subjected to stresses far beyond expected operating conditions to rapidly precipitate failures and expose weak links in the design. HALT functions as a discovery tool rather than a lifetime predictor, revealing the true operational and destructive limits of a system by incrementally increasing thermal, vibrational, or electrical stresses until failure modes become observable. This process enables engineers to understand the mechanisms that govern unreliability in early life and during ageing, and to implement corrective actions that enhance robustness.

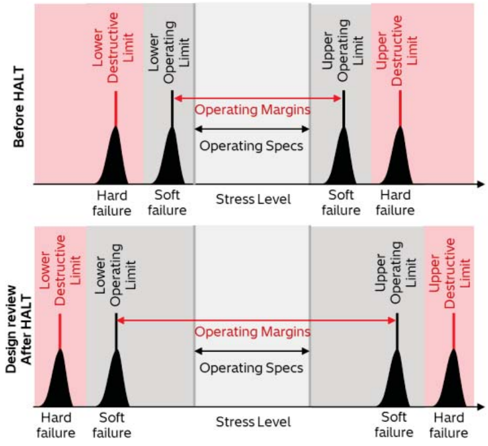

Figure 5. Change of product operating and destructive limits after HALT testing and design improvements implementation.

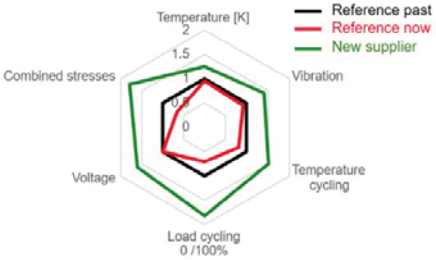

Figure 6. Results of product margins improvements verified along the different stress profiles carried out in HALT testing.

Once the design weaknesses are addressed, reliability continues at the production level through Highly Accelerated Stress Screening (HASS) where a validated “proof-of-screen” profile, Figure 7, is applied that precipitates latent defects without harming good units; combined modulated excitations (temperature + multi-axis vibration + power cycling) shorten screen time and reduce “No Defect Found” (NDF).

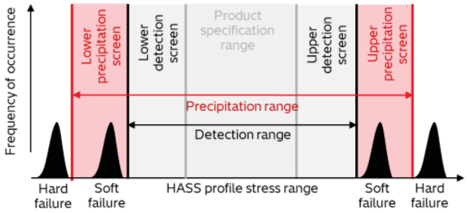

Figure 7. Application ranges of product screening (HASS) using knowledge of product margins estimated from HALT.

Recurring HALT (Re-HALT) is also conducted periodically, not only to verify that previously identified weaknesses have been effectively resolved and that the design margin has demonstrably improved, but also to ensure continued reliability after product release. This aim at mitigating risks due to production variation, changes in the BOM (Bill of Materials) or component changes over the product life cycle. Occasional re-HALT after production is performed to confirm that these supplier-driven modifications have not introduced new vulnerabilities.

The implementation of those reliability tests in the product lifecycle is summarized in Figure 8.

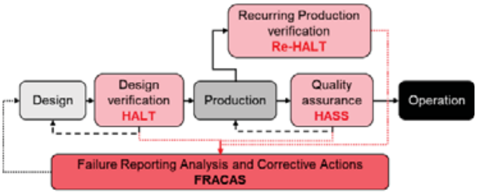

Figure 8. Process implementation of reliability testing processes.

ALT exposes components to elevated stresses—such as temperature, humidity, voltage, or vibration—that speed up the failure mechanisms, allowing failures to occur much faster than they would under normal operating conditions. By running tests at multiple stress levels and fitting life-stress acceleration models, ALT enables the collection of meaningful failure data and the extrapolation of lifetime behavior at nominal operating conditions (Collins 2017). This "time compression" provides the failure time distributions required for statistical modeling without waiting months or years for natural failures to occur. Lifetime in reliability engineering refers to the statistically distributed duration that components or systems can operate before failing. Because individual units do not all fail at the same moment, lifetime is described using probability distributions such as Weibull, lognormal, or exponential, each capturing different behaviors across early-life, useful-life, and wear-out phases (Meeker 2009). These distributions allow engineers to estimate metrics such as mean time to failure (MTTF), B5 life (the time by which the component has a 5% failure probability), or mission-time reliability. However, any lifetime estimate derived from ALT or normal-stress testing is only as credible as the data and modeling assumptions behind it, which is why confidence intervals are essential. Confidence intervals quantify how certain we are that the estimated lifetime represents the true underlying behavior of the product population (Wang, et al. 2025). For example, stating that a component has a B5 life of 50,000 hours at 90% confidence means that, based on accelerated test data, fitted distributions, and stress-life models, we are 90% certain that 5% of units will fail before 50,000 hours. Confidence intervals combine the epistemic uncertainty caused by limited sample sizes, that can be reduced at the cost of more testing, and the aleatoric uncertainty that is inherent to the variability of failure events. Lifetime quantification and confidence interval description are required for turning lifetime predictions into decision-ready reliability statements.

In the modelling of low-probability events, as failures that occur extremely rarely but can have severe consequences, are a central challenge in reliability engineering. These events often lie in the tail of failure probability distributions and

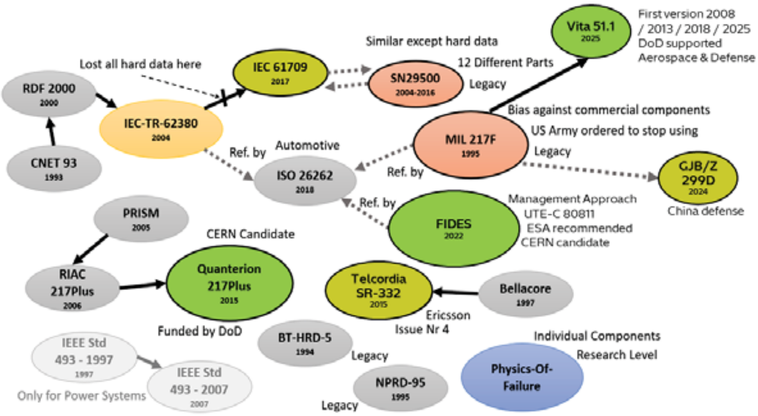

Figure 9. Known reliability standards and their evolution over time

cannot be directly observed with traditional testing because failures are too infrequent. Therefore, engineers use specialized statistical and computational methods to estimate their probability with acceptable accuracy (Hua, et al. 2014). There are different standards for reliability prediction (Mou, et al. 2013), for example MIL-HDBK-217F, Telcordia SR-332, FIDES, Quanterion 217Plus, IEC 61709or GJB 299D. Each of these standards had been developed for different industries, considering different stress factors and different failure rates. These determine the standard's severity and applicability. For example, MIL-HDBK-217F is one of the most conservative standards designed for military industry, using more pessimistic failure rates and characterized by higher complexity (considering more stress factors), where Telcordia SR-332 had been developed for commercial application, using more optimistic failure rates (in comparison with MIL-HDBK-217F) and characterized moderate complexity. Figure 9 shows known standards and their evolution over time.

4. Reliability testing economics and optimization

Reliability assurance testing strategy is ultimately an economic choice made under uncertainty. Practically speaking, the trade-off is between reducing uncertainty in lifetime estimates and failure models via laboratory testing budget and, ultimately, time.

For modular systems, such as power converters, the economic consequence of uncertainty is amplified as submodule level events propagate through a *k-out-of-n* system structure – even smallest underestimate in failure probability can translate into earlier consumption of system redundancy and, as the result, higher expected costs.

A compact yet effective way to formalize this is to separate the decision made "downstream" and information generated "upstream". Let d denote an operational decision such as an inspection interval, and let g denote a chosen ALT design, such as sample size, censoring time, stress levels, monitoring etc. Then the expected total cost can be expressed in the following form:

$$\mathbb{E}[C_{tot}(\theta,g)] = \mathbb{E}[C_{LC}(d|\theta)] + C_{test}(g),$$

where C_{LC} is the expected life cycle cost of operation and maintenance and θ are uncertain lifetime-model parameters. Planning guidance such as IEC 60300-3-5 (Commission, IEC 60300-3-5:2001 2001) is useful here because it presents reliability tests as a statistical tool to facilitate decision making. The testing cost term needs to be explicitly shown as it is rarely a simple function of number of datapoints. A cost model is a function of many parameters, such as equipment, staff-hours, lab costs, and others. It can be generalized as:

$$\begin{aligned} C_{test}(g) = C_{setup} + n_g C_{unit} \\ + t_{test}(g)\big(C_{equipment} \\ + C_{personnel}\big) + C_{data}, \end{aligned}$$

where n_g is the number of tested units (or datapoints), C_{unit} includes build, preparation and instrumentation costs, t_{test} is the test time for which the setup is occupied, $C_{equipment}$ and $C_{personnel}$ represent running costs of the equipment and hourly rate of personnel involved, and C_{data} covers all the expenses related to the data analysis and any other data related aspects. The parallelization of testing efforts, as well as censoring, influences the economics – with m parallel channels and a censoring time τ, a simple approximation is $t_{test} \approx \tau(n_g/m)$. This simple idea captures the common batch effect – it is nearly free to add a datapoint within the existing batch, but costly if it forces an additional batch.

Although, when detailed costs are unavailable, it is still useful to test sensitivity with simple surrogates that reflect common understandings: a linear form that would correspond to sequential testing, a logarithmic form to reflect efficiency gains from shared setup and parallel testing, a power-law form to represent bottlenecks in form of queuing effects that increase cost as the testing campaign grows. These forms are not the final model, but they are good enough to show when economics will force certain limitations for performed tests.

While talking about costs, when resources are tight, the question "is another test worth it" should be answered explicitly. Bayesian decision methods provide a natural framework because they treat uncertain parameters, failure rates, and model choices as part of the decision (Antonios Kamariotis 2022). In ALT specifically, these methods show how prior information and test design interact, and why the same budget can yield very different information depending on decisions taken. This perspective connects directly to Value of Information (VoI), where additional testing is justified when the financial gain expected by reduction in epistemic uncertainty exceeds the cost of test. For power electronics, this usually means prioritizing information that affects maintenance strategy and the implemented redundancy, as these are key drivers of lifecycle cost.

Inspection and sampling provide the first economic control point. Sampling standards such as ANSI/ASQ Z1.4 formalize the trade-off between inspection effort and acceptance risk through Acceptance Quality Levels (AQLs). Rather than applying fixed inspection levels, tightened inspection following changes, followed by relaxation once stability is demonstrated, minimizes cost while preserving confidence.

Reliability demonstration tests (RDTs) are success-run and binomial-based sampling to demonstrate target reliability at specified confidence, with Larson nomographs (G. Calot - E. Morice - F. Dousset 1967) (Defense. 1996 (originally issued 1987)) guiding sample-size vs. observed failures trade-offs.

For binomial RDTs, confidence depends on sample size, test duration, and the number of observed failures. Nomographs and closed-form expressions show that, as reliability targets increase, the required population grows rapidly, especially when none or one failure is allowed. For long-life power electronics, where observation windows are short relative to mission life, demonstration cost can escalate beyond practical limits. This reinforces the distinction between evidence gathered to support compliance statements and evidence gathered to understand failure mechanisms.

Accelerated life testing and degradation testing address this gap, but they must be planned with care. It is common, for a power electronics system, to achieve rather modest acceleration factors, limiting the benefit of simply increasing stress when combined with realistic limits of the stress applied. Under these conditions, what is measured becomes as important as how long the test runs. Well-chosen degradation indicators can provide usable information earlier than hard failures. Investment in instrumentation and monitoring can therefore reduce the required test duration and improve the information gained per test hour.

Monte Carlo based test planning allows for evaluation candidate ALT design choices under realistic constraints. Instead of optimizing a single test in isolation, the approach identifies combinations of stress levels, sample allocations, etc. that meet accuracy requirements at lower expected cost. This is particularly relevant when multiple stress mechanisms compete for the same infrastructure.

Production screening sits at the boundary between factory cost and field exposure. Analytical cost models show that partial screening is often economically unfavoured: depending on defect probability and field consequence, either minimal screening or near-100% screening is optimal. For the case of "*the Bad*" type of failure mode, where redundancy is not sustainable due to failure mode, the breakeven point frequently shifts toward higher screening coverage. These models support rational selection of HASS fractions and highlight the need for proof-of-screen validation to avoid introducing damage mechanisms unrelated to field stress.

The largest economic gains typically occur at prototype stage, where RGM can be applied. Models such as the Duane formulation and the Crow/AMSAA approach quantify improvement achieved through test–analyze–fix cycles. Separating instantaneous from cumulative MTBF helps teams distinguish current design capability from previously established performance. This supports explicit maturity targets and helps align resource allocation with the remaining reliability gap.

Once again, across all stages, the guiding question is whether additional testing is worth its cost. Framing reliability assurance in terms of accuracy, time, and cost makes this question explicit. Tests that reduce uncertainty in parameters affecting redundancy consumption, maintenance strategy, or operational availability usually deliver higher value than tests that only marginally refine secondary estimates. For mission-critical power electronics, careful and well thought through planning of inspection, testing, screening, and growth activities is therefore as important as the technical execution of the tests themselves.

5. Physics-informed deep learning and advanced data analytics

Physics-informed machine learning (PIML) is a practical compromise between purely empirical, data-driven, models and purely physics-based models. The central idea is to train a model against measured data while enforcing consistency with known dynamics or constraints. In its most common form, this can be expressed through a reduced physical model with unknown parameters θ:

$$\dot{x}(t) = f(x(t), u(t), \theta),$$
$$y(t) = h(x(t), u(t), \theta) + \epsilon(t),$$

where x is the true (physical) state vector, u represents known inputs (load, duty cycle, ambient conditions…), and y would be the representation of measured signals. PIML methods typically optimize a loss that balances measurement fit against physics consistency:

$$\mathcal{L}(\omega,\theta) = \sum_{i=1}^{N} \|y_i - \hat{y}(t_i;\omega,\theta)\|^2 + \lambda \sum_{i=1}^{N} \left\|\dot{\hat{x}}(t_i;\omega,\theta) - f(\hat{x}(t_i;\omega,\theta), u(t_i), \theta)\right\|^2 + \Omega(\omega,\theta).$$

In here, $\hat{x}(t)$ acts as a counterpart to the physical state $x(t)$, representing the model's inferred state vector learned from outputs $y(t)$. The second term discourages solutions that violate the model structure and is the main reason these methods can remain stable when data are sparse or collected under operating regimes different from those seen in training. In the equation, ω marks trainable weights, λ sets the strength of the physical constraints, and Ω captures regularization and practical constraints such as bounded parameters.

This philosophy is aligned with physics-informed neural network formulation, which uses differential-equation residuals as part of training rather than relying only on data (M. Raissi 2019).

Within reliability engineering, the most useful PIML outputs are interpretable quantities – parameter estimation is a great example: instead of treating the system as a black box, the model estimates effective parameters (θ) from measured trajectories. Estimated parameters can then be tracked over time and across different stress conditions as system health indicators, and, for obvious reasons, they are easier to connect to failure analysis than opaque latent embeddings. The IEEE work by Zhao, Peng, Zhang, and Wang provides a concrete example, illustrating how converter dynamics can reduce dependence on large training datasets and yields parameters that remain interpretable for failure analysis (S. Zhao 2022).

When a physics-constrained model predicts expected behavior at the observed operating point, the deviation between measurement and prediction can be treated as an anomaly signal. This approach is usually more robust than simple fixed limits as it is less affected by changes in operating conditions. It is especially helpful when early warning signs are small slow drifts or occasional dropouts that can be hidden by normal operating changes. Recent reviews of PIML in condition monitoring emphasize residual-based detection and *interpretability* as key advantages in industrial diagnostics and anomaly detection (Yuandi Wu 2024).

For prognostics, the hybrid modeling approach can be shortly expressed through an internal damage state $D(t)$ that accumulates under stress:

$$D(t + \Delta t) = D(t) + g(S(t), T(t), \phi)\Delta t,$$

where $S(t)$ represents stress metrics, $T(t)$ is temperature profile, and ϕ are uncertain damage parameters. Following this parametrization, failure is defined as $D(t) \geq 1$. A crucial consequence here is that a physics-informed model constrains $g()$ to follow known behavior. The key requirement for reliability-based decision making is explicit quantification of uncertainty, as it is as important as the estimated point value of failure probability – maintenance planning depends on whether risk is clearly increasing or remains within its expected variability. Exactly when diagnostics and prognostics are used operationally, the PHM standards surveys conducted by NIST are useful references (Gregory W. Vogl 2014).

A coherent implementation requires more than a model. A practical stack follows a clear pipeline from generated data to decision making. Continuous data acquisition during HALT, Re-HALT, HASS and ALT should be planned as part of the test itself, with consistent metadata so results remain useful for the entire lifetime of a product and beyond. Feature definition should be guided by the FMMEA (Failure Mechanism, Mode & Effects Analysis), so that each measured signal is mapped to plausible failure mode and the features remain meaningful also after tests. On the modelling side, Physics-Informed Deep Learning (PIDL) can embed reduced thermal-electrical surrogate behaviour and cumulative stress or damage accounting so that learned predictions stay consistent with known dynamics while still capturing effects that are difficult to parameterize. For monitoring functions that must react quickly in operation, a subset of indicators can run as edge analytics in the converter control hardware, supporting latency-sensitive monitoring and, possible selection and prioritization for deeper offline analysis. Finally, the same data structure should feed a FRACAS, where field events are used to update models, ensuring that learning remains valid throughout the lifetime of the product.

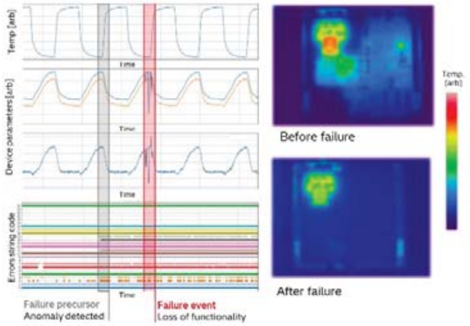

Figure 10. Example of data collected during reliability testing and failure analysis where PIML can be applied to accelerate the reliability growth methods. The data example is a combination of time series of analogue quantities and errors text string with thermal imaging before and after test.

Finally, it is important to acknowledge that a significant portion of reliability information is recorded and preserved in a text format: failure reports, notes, logs, etc. Natural language processing can reduce this bottleneck by extracting valuable information in structured form. Recent work introducing the Failed Components Extraction from Text (FaCET) shows how large language models (LLMs) can support FRACAS-driven learning loops that Chapter 3 motivates: continuous testing produces signals, physics-informed models turn signals into interpretable indicators, and FRACAS turns outcomes into systematic improvement across the product lifecycle.

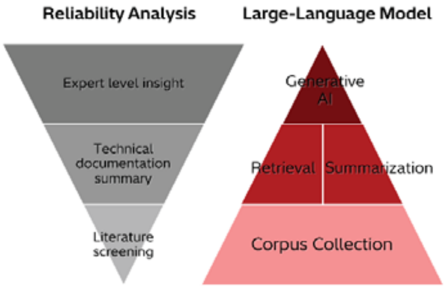

Figure 11. Intersection between duties of a domain expert in Reliability analysis and the potential overlap with a Large Language Model.

5.1. Autonomous Reliability Laboratory

The fast development of automation, robotics and AI can find an application in the domain of reliability engineering for the development of autonomous laboratory, that can be understood as a reliability-engineering concept that inspire and extends from the current association for "materials discovery" (Szymanski 2023). While there is justified skepticism regarding AI-driven systems making true scientific "discoveries" in the sense of uncovering new fundamental laws of nature, these systems are far better suited to products failure-mode and root cause identification. In this context, AI can identify analogue patterns across large libraries of failure signatures, guiding expert judgement in ways that would otherwise be limited by the need to focus only on the most urgent investigations (Alexander V. Tobias 2025).

Autonomous reliability laboratory frameworks that integrate HALT, Re-HALT, HASS and Accelerated Lifetime Testing (ALT) can reveal distinct and otherwise hard-to-observe weaknesses in power electronics or help explore a dense parameter space of operation states and stresses that would be economically unfeasible with standard testing procedures. For example, gate-drive insulation DC/DC converters may show completely different failure signatures depending on the stress regime: thermal cycling may produce transformer wire breaks, whereas high-temperature–humidity exposure can lead to electrochemical insulation degradation. Similarly, clock-oscillator solder joints can exhibit intermittent failures as temperature-cycling ranges broaden—behavior that becomes visible only through instrumented, run-to-failure testing. Together, these capabilities demonstrate how autonomous reliability laboratories can accelerate and deepen reliability insights in ways that traditional workflows cannot.

6. Conclusion and outlook

This work presented an integrated framework for Quality and Reliability (Q&R) assurance tailored to mission-critical power electronics, in particular emphasis high-power converters of modular topologies built by many series and parallel connected cells. By combining Design-for-Reliability (DfR), Reliability Growth Management (RGM), accelerated testing methodologies (HALT, HASS, ALT), and structured data-centric processes such as FRACAS and digital traceability, the paper outlined how reliability knowledge can be systematically generated, preserved, and applied throughout the full product lifecycle.

Advancing reliability assurance also demands improved toolchains: with physics-informed and data-driven approaches emerged as a central theme. Highly accelerated tests uncover design weaknesses, while physics-of-failure modelling contextualizes the observed behaviour and enables quantitative lifetime estimation despite the inherent limitations of accelerated testing. Physics-informed deep learning was shown to provide a bridge between sparse laboratory data and operational variability, enabling interpretable parameter estimation, anomaly detection, and degradation forecasting without relying on impractically large datasets.

Complementing these developments, the notion of autonomous reliability laboratories illustrates a paradigm shift: automated, AI-supported test workflows can explore stress spaces, detect precursors, and accelerate the identification of failure mechanisms beyond human-driven methods alone.

Finally, experts capable of linking physics of failure at component level to systems-level converter reliability, remain crucial. Automated Root Cause Analysis (RCA) engines can accelerate investigations, but expert oversight remains important for interpreting complex interactions and validating predictive models. Building this combined human-digital expertise will be critical for ensuring that future converter systems achieve robustness, predictability, and maintainability.

Acknowledgement

We thank Magnus x. Jobs and Marcin Wentrys for the composition of Figure 9

References

Alexander V. Tobias, Adam Wahab. 2025. "Autonomous ‘self-driving’ laboratories: a review of technology and policy implications." *Royal Society Open Science* 12 (7).

Antonios Kamariotis, Eleni Chatzi, Daniel Straub. 2022. "Value of information from vibration-based structural health monitoring extracted via Bayesian model updating." *Mechanical Systems and Signal Processing* 166.

Choi, Ui-Min, and Frede Blaabjerg. 2018. "Separation of Wear-Out Failure Modes of IGBT Modules in Grid-Connected Inverter Systems." *IEEE Transactions on Power Electronics* 33.

Collins, D. H., Freels, J. K., Huzurbazar, A. V., Warr, R. L., & Weaver, B. P. 2017. "Accelerated Test Methods for Reliability Prediction." *Journal of Quality Technology* 45 (3).

Commission, International Electrotechnical. 2001. "IEC 60300-3-5:2001." *Dependability management - Part 3-5: Application guide - Reliability test conditions and statistical test principles.* 07 03.

Accessed 02 20, 2026. https://webstore.iec.ch/en/publication/1304.

—. 2016. "IEC 61078:2016 Reliability Block Diagrams." Geneva, Switzerland, aug.

Crow, Larry H. 2008. "A methodology for managing reliability growth during operational mission profile testing." *Annual Reliability and Maintainability Symposium.*

Defense., U.S. Department of. 1996 (originally issued 1987). "MIL-HDBK-781A: Handbook for Reliability Test Methods, Plans, and Environments for Engineering, Development, Qualification, and Production."

Edoardo Martino, Andrew Fairbrother, Riddhi Ghosh, Jurgen Shuderer. 2023. "Data-driven and physics-based reliability tests to failure of a power electronics converter." *Microelectronics Reliability* 150.

G. Calot - E. Morice - F. Dousset, R. Casamitjana, J.L. Groboillot, A. Warnier de Wailly - J. Wasilevski - P. Thionet - Harry H. Larson - Georges Roux, Maurice Roux - J. Mouy - G. Sergeant - G. Maillet - J.R. Barra - G. 1967. *Revue de Statistique Appliquée.* Centre de Formations aux Applications Industrielles de la Statistique.

Gissler, Bob, and Pankaj Shrivastava. 2015. "A system for design decisions based on reliability block diagrams." *Annual Reliability and Maintainability Symposium (RAMS).*

Gregory W. Vogl, Brian A. Weiss, and M. Alkan Donmez. 2014. "Standards for Prognostics and Health Management (PHM) Techniques within Manufacturing Operations, ." *Annual Conference Of The Prognostics And Health Management Society,NIST publications.*

Hosseinabadi, Farzad, Sajib Chakraborty, Sachin Kumar Bhoi, Guenter Prochart, Dino Hrvanovic, and Omar Hegazy. 2024. "A Comprehensive Overview of Reliability Assessment Strategies and Testing of Power Electronics Converters." *IEEE Open Journal of Power Electronics* 5.

Hua, Bowen, Zhaohong Bie, Siu-Kui Au, Wenyuan Li, and Xifan Wang. 2014. "Extracting Rare Failure Events in Composite System Reliability Evaluation Via Subset Simulation." *IEEE Transactions on Power Systems* 30.

Huai Wang, Ke Ma, Frede Blaabjerg. 2012. "Design for Reliability of Power Electronic Systems ." *8th Annual Conference of the IEEE Industrial Electronics Society (IECON 2012).*

IEEE, Publisher:. 2012. "Beyond FRACAS: Integrating field failures and field data." *Proceedings Annual Reliability and Maintainability Symposium.*

Kirill Ivanov, Horst Lewitschnig. 2025. "RUL prediction using Bayesian polynomial regression." *Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference.*

M. Raissi, P. Perdikaris, G.E. Karniadakis,. 2019. "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations." *Journal of Computational Physics* 378.

Meeker, W. Q., Escobar, L. A., & Hong, Y. 2009. "Using Accelerated Life Tests Results to Predict Product Field Reliability." *Technometrics* 51 (2).

Mou, Haowen, Weiwei Hu, Yufeng Sun, and Guangyan Zhao. 2013. "A comparison and case studies of electronic product reliability prediction methods based on handbooks." *International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE).*

Nami, Alireza, Jiaqi Liang, Frans Dijkhuizen, and Georgios D. Demetriades. 2014. "Modular multilevel converters for HVDC applications: Review on converter cells and functionalities." *IEEE Transactions on Power Electronics* .

Pandey, Vikas, Utpal Rabha, Gayatri Pandit, and Swapnil Sabde. 2016. "Uncovering the latent defects in dimming control devices by introducing application specific stress profile in HALT — A case study." *International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO).*

Raheja, Dev. n.d. *Design for Reliability.* John Wiley & Sons, Inc.

S. Zhao, Y. Peng, Y. Zhang and H. Wang,. 2022. "Parameter Estimation of Power Electronic Converters With Physics-Informed Machine Learning." *IEEE Transactions on Power Electronics* 37.

Szymanski, Nathan J., Bernardus Rendy, Yuxing Fei, Rishi E. Kumar, Tanjin He, David Milsted, Matthew J. McDermott et al. 2023. "An autonomous laboratory for the accelerated synthesis of novel materials." *Nature 624 no. 7990* 86-91.

Wang, Yanfang, Ying Chen, Yingyi Li, and Rui Kang. 2025. "Reliability Modeling Analysis and Uncertainty Quantification Method for Electronic Systems With Multireliability Dependency." *IEEE Transactions on Reliability* 74 (4).

Yuandi Wu, Brett Sicard, Stephen Andrew Gadsden. 2024. "Physics-informed machine learning: A comprehensive review on applications in anomaly detection and condition monitoring." *Expert Systems with Applications* 255.