From Prediction to Discovery: The Changing Paradigm in Electronic Reliability

Kirk A. Gray
Accelerated Reliability Solutions, L.L.C.
Revised September 2026

Abstract

Electronic reliability engineering developed during an era when vacuum tubes, discrete components, and early solid-state devices exhibited prominent intrinsic failure and wear-out mechanisms. Reliability programs therefore emphasized component failure-rate prediction, burn-in, and qualification to specified environmental limits. Modern electronic systems are different. Highly integrated devices, complex assemblies, global supply chains, rapid component turnover, firmware-software interactions, and manufacturing variation have shifted much of the practical reliability challenge from estimating a single failure rate toward discovering design weaknesses, process-induced defects, and insufficient operating margins. This article argues for a discovery-based reliability strategy centered on stress-to-limit testing, root-cause analysis, physics of failure, Highly Accelerated Life Testing (HALT), and Highly Accelerated Stress Screening (HASS). The objective is not to discard quantitative reliability analysis, but to use prediction only where its assumptions and evidence are appropriate, while placing greater engineering emphasis on discovering how a product actually fails and increasing the distance between normal use stresses and the product's true limits. This distinction is important: empirically validated life models can be useful when a specific physical failure mechanism is known, the relevant stresses can be measured, and the model is supported by representative empirical data. That is fundamentally different from assigning statistically derived component failure rates to a complex solid-state electronic system and treating the resulting number as a prediction of future field reliability.

Keywords: electronic reliability, HALT, HASS, reliability prediction, design margin, physics of failure, manufacturing variation, accelerated testing

1. Why the Reliability Paradigm Is Changing

For many decades, electronic hardware was comparatively fragile. Vacuum tubes contained glass structures and filaments with recognizable wear-out mechanisms. Early solid-state devices were vulnerable to contamination, material instability, packaging limitations, and other mechanisms that could produce time-dependent failures. Because many early defects were accelerated by elevated temperature, burn-in became an effective method for removing some early-life failures. Large discrete components also imposed mechanical limitations that made vibration and shock survival a major design concern.

In that environment, reliability testing was often considered successful when equipment survived its specified shipping, storage, and operating environments. Reliability engineering also developed a strong quantitative tradition: component failure rates were estimated, summed, and converted into system-level metrics such as failure rate and MTBF.

Those methods provided a common language for comparison and acquisition decisions, but the electronics industry has changed dramatically. Today's systems contain highly integrated devices, dense interconnects, high-speed interfaces, complex power-management networks, firmware, software, and components produced through rapidly changing global manufacturing processes. The resulting field reliability is influenced not only by intrinsic component mechanisms, but also by materials, geometry, assembly, process control, mission profile, environmental loading, design margin, and interactions among subsystems.

The important question for the practicing reliability engineer is therefore no longer simply, "What failure rate does a handbook predict?" It is also, and often more importantly, "Where are the weak links, what stresses expose them, and how much margin does the design actually have?"

2. Statistical Failure-Rate Prediction Is Not Demonstrated Field Reliability

For complex solid-state electronic systems, the central issue is not whether a prediction method is standardized or consistently applied, but whether its calculated result has demonstrated correlation with actual reliability. Jais, Werner, and Das reported that MIL-HDBK-217 and related handbook methods produced highly inaccurate results when compared with demonstrated test data from military programs [10]. In an AMSAA survey of 15 systems, the ratio of predicted to demonstrated reliability ranged from 1.2:1 to 218:1; their Spearman rank-order analysis further found that MIL-HDBK-217-based predictions could not support comparisons between systems [10]. A calculation can therefore be internally consistent without being a valid predictor - or even a valid ranking - of field reliability.

This distinction is important because handbook predictions are often treated as engineering information about a product. Yet the same component population can produce very different reliability when board support, component placement, thermal paths, assembly quality, manufacturing variation, or other system-level factors change. Jais, Werner, and Das illustrate this with two circuit boards containing the same components but different mounting support: the mechanical design substantially changed vibration response and reliability, while MIL-HDBK-217 produced the same reliability prediction for both [10]. The National Academies has likewise documented limitations of MIL-HDBK-217-type approaches, including their lack of explicit treatment of materials, geometry, failure mechanisms, and realistic loading histories [3].

MIL-HDBK-217F was validated by the U.S. Department of Defense for continued use in acquisition in June 2022 [5]. However, the 2022 action was a validation notice, not a substantive technical revision. The current technical methodology remains MIL-HDBK-217F, Notice 2, dated February 28, 1995. More than 30 years of changes in semiconductor technology, packaging, materials, interconnects, manufacturing, power density, operating speed, and hardware-software interaction are therefore not represented by a substantive modernization of the handbook. Jais, Werner, and Das make the additional point that simply updating the handbook database would not remove its underlying technical limitations [10].

This criticism is directed at generic statistical component failure-rate prediction of complex solid-state electronic systems, not at all mathematical life modeling. When a physical degradation or wear-out mechanism is known, the controlling stresses can be measured, and the stress-damage or stress-life relationship is supported by representative empirical data, a physical model can provide useful life information. Mechanical fatigue and wear are obvious examples; batteries, solder-joint fatigue, bond-wire and die-attach fatigue, corrosion, creep, and other identifiable mechanisms may also be modeled when the mechanism, material behavior, loading, and application are sufficiently characterized. That is fundamentally different from assigning generic constant failure rates to electronic parts and summing them to predict system MTBF.

3. A Discovery-Based Reliability Strategy

A discovery-based reliability program starts with a different objective. Instead of asking only whether a product passes a specified environmental test, the engineer deliberately explores the product's actual operating and destruct limits. Stress is increased in controlled steps, failures are investigated, and the physics behind each weak link is understood. Corrective action is then directed at increasing margin rather than merely demonstrating compliance.

This philosophy is consistent with the qualitative purpose described in IEST-RP-PR003, which characterizes HALT and HASS as methods for improving product reliability and quality rather than as quantitative methods for calculating field life [6]. HALT can reveal whether a design requires corrective action to become more

robust; HASS can then use knowledge gained during HALT to detect adverse process or component changes in production.

The distinction is fundamental: qualification asks whether a product meets a requirement; HALT asks how far the product can be pushed, what fails first, why it fails, and whether that limit should be improved.

4. Margin, Variation, and the Weakest Units

Every volume-manufactured electronic product has a distribution of inherent strength. Component tolerances, solder joints, connector engagement, material properties, assembly processes, cooling effectiveness, firmware settings, and supplier variation all contribute to unit-to-unit differences. The stresses experienced in the field also form a distribution. Reliability problems occur when the low-strength tail of the product population overlaps the high-stress tail of the use environment.

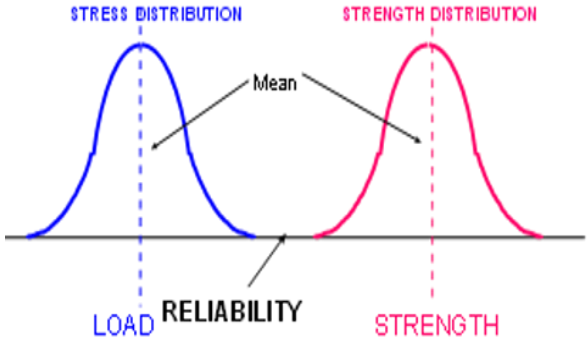

Figure 1. Conceptual separation between fatigue-life consumption and the inherent strength distribution of electronic assemblies. Reproduced from the author's original article.

A robust design maintains substantial separation between these distributions. Manufacturing changes can reduce that separation without producing an obvious immediate failure. An engineering change, operator change, new supplier lot, process drift, altered material, or several small changes acting together may shift the strength distribution downward. The result can be a population of latent defects that passes ordinary production tests but fails under real-world stress.

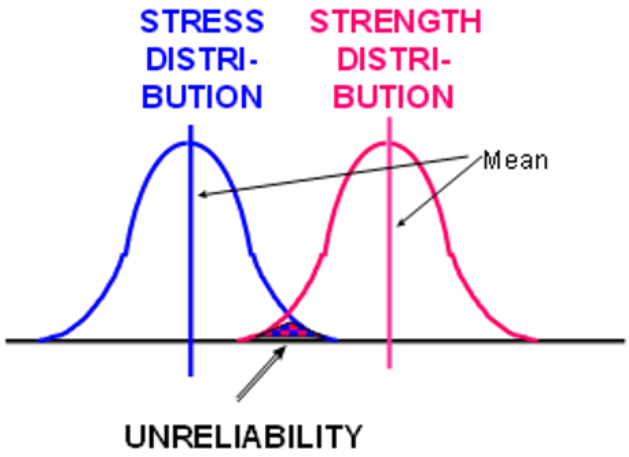

Figure 2. Conceptual field-failure region created when the low-strength tail of the product population overlaps the applied-stress distribution. Reproduced from the author's original article.

The important engineering response is not to attempt to predict every future combination of events. It is to discover the weak links, understand what controls their strength, increase the available margin, and establish production controls that can detect significant shifts before affected units reach customers.

5. Case History 1 - From Burn-In to a One-Hour HASS

One client organization, referred to here as Company A, had a product with an annualized field return rate of approximately 5%. Its existing reliability screen relied on an elevated-temperature burn-in room operating at approximately 90 °F. A HASS process was developed using knowledge of the product's margins and failure behavior.

Six months after implementation of a one-hour HASS process, the reported annualized field return rate had fallen to approximately 0.5%. The important lesson is not that every HASS program will produce the same numerical improvement. Rather, the case illustrates what can happen when a production screen is based on experimentally discovered product capability and is designed to expose latent manufacturing defects rapidly instead of simply accumulating time at a moderately elevated temperature.

This type of result also explains why detailed industrial HALT/HASS results are not commonly published. The combination of lower warranty cost, faster screening, reduced rework, and improved customer experience can create a meaningful competitive advantage.

6. HALT: Finding the Real Limits

HALT is conceptually straightforward, although effective execution requires disciplined instrumentation, functional monitoring, failure analysis, and engineering judgment. Stress is increased beyond normal specification limits in steps while the unit under test remains functionally monitored. The objective is to identify operational limits, determine whether recovery occurs when stress is reduced, continue toward destruct limits when appropriate, and investigate the root cause of each failure.

A practical sequence often begins with low temperature, followed by high temperature, random multi-axis vibration, and then combinations of stresses. The exact sequence should be adapted to the product and the information sought. Individual stresses can simplify root-cause isolation, while combined stresses can expose interactions and are especially valuable when developing production screens.

Three to five samples can provide substantial engineering information during an exploratory HALT, although more units may be required when the objective is to characterize the distribution of a particular failure mechanism. Statistical analysis remains valuable, but it should not displace the central HALT objective: identify the mechanism, understand the physics, and increase the margin.

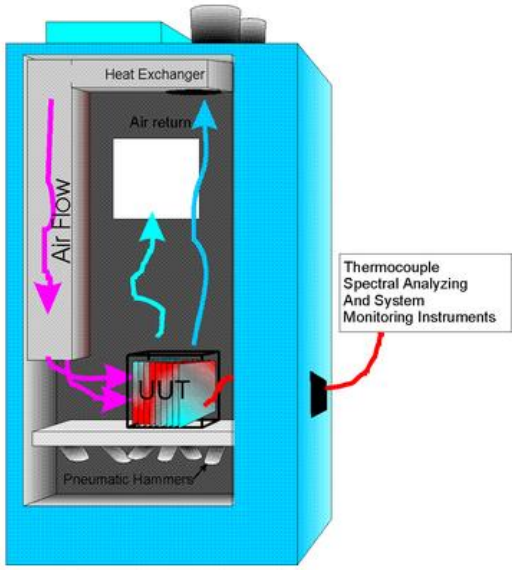

Figure 3. Typical HALT/HASS chamber configuration showing rapid thermal airflow, pneumatic vibration excitation, and functional monitoring. Reproduced from the author's original article.

7. Case History 2 - A 30 °C Margin Improvement from a Small Design Change

A second case demonstrates why failures beyond the product specification can be valuable engineering information. During step-stress testing of an RF power supply, the unit under test reached an upper operating limit of approximately 60 °C. Investigation traced the limitation to a small regulating diode in the +15 Vdc auxiliary power supply. The diode was positioned next to a heat sink but did not make thermal contact with it.

The component was repositioned so that it contacted the heat sink. That small, essentially no-cost change increased the observed upper operating limit by approximately 30 °C, from 60 °C to 90 °C.

The original product already met its specified operating requirement. A conventional pass/fail qualification test therefore would have provided little reason to investigate this component. Step-stress testing exposed the true limit, root-cause analysis identified the mechanism, and a minor design change produced a major increase in thermal margin.

8. Small Physical Changes Can Produce Large Reliability Gains

The same principle applies to mechanical design. Component mass, lead length, support, orientation, mounting method, and local board stiffness can have a large effect on vibration response. Relatively inexpensive layout or attachment changes made early in development can therefore provide substantial improvement in mechanical stress margin.

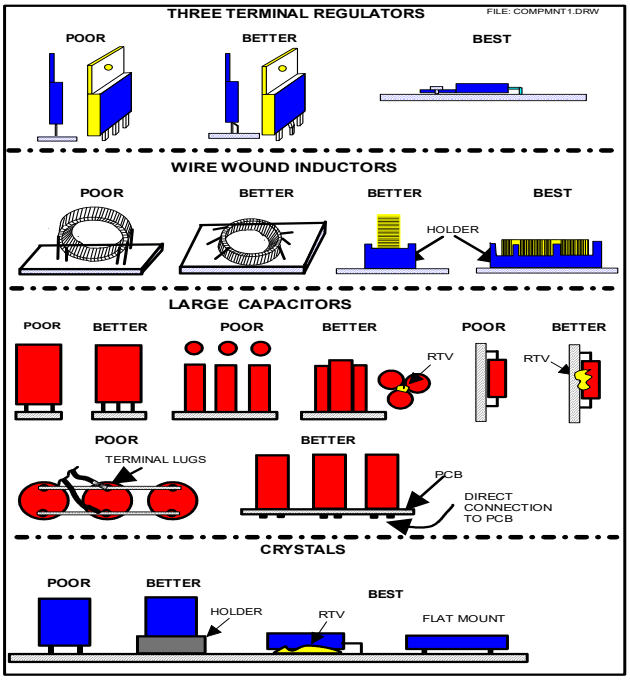

Figure 4. Examples of component-placement and support practices that can improve vibration margin. Original figure courtesy of C. Felkins, Storage Technology; reproduced from the author's original article.

The value of stress-to-limit testing is that these weaknesses become visible while changes are still inexpensive. Reliability improvement is usually far more economical during development than after field returns, redesign, requalification, production disruption, and customer dissatisfaction have begun.

9. HASS: Using Demonstrated Margin to Control Manufacturing

Once HALT has established meaningful operating and destruct margins, that knowledge can be used to develop a HASS process. HASS is not simply 'HALT on production units.' Its purpose is to expose manufacturing defects and process shifts quickly without consuming an unacceptable fraction of the useful life of conforming hardware.

The screen must therefore be both effective and safe. Gregg Hobbs emphasized the importance of Safety of HASS because an overly severe screen can itself reduce usable field life [7]. Appropriate HASS development requires knowledge of the product margins, monitoring of relevant functions, verification that the screen precipitates defects, and evidence that good units are not being damaged.

When properly developed, a higher-stress combined-environment screen can often reveal latent defects much faster than a long-duration burn-in. The engineering advantage is not simply shorter test time; it is that the screen is tied to the actual capability of the product rather than to an arbitrary time-temperature recipe.

10. Physical Life Modeling, Failure Mechanisms, and Stress-to-Limit Testing

A discovery-based reliability strategy is not an argument against mathematical modeling. It is an argument that a model should have a demonstrated physical and empirical basis. When a specific degradation or wear-out mechanism is known, the controlling material properties and stresses are measurable, and an experimentally established stress-life relationship is valid for the application, modeling can provide valuable estimates of damage accumulation or life. Jais, Werner, and Das make this same distinction for power electronics: wire-bond and die-attach failures depend on usage and loading and cannot properly be represented by a constant failure rate, yet the mechanisms and their time to failure can be characterized with established physical models [10].

The appropriate sequence is therefore mechanism first, model second. Testing, field-return analysis, or HALT can identify the weak mechanism; physics-of-failure analysis can explain its sensitivity to materials, geometry, loading, and design variables; empirical data can establish or validate the stress-life relationship; and subsequent testing can verify that corrective action increased margin. A model used in this way describes a particular physical mechanism under defined conditions. It should not be confused with a generic statistical failure rate assigned to a solid-state component or with a summed component-rate prediction of complex system reliability.

10.1 The Mission-Profile Problem: You Cannot Model Stresses You Do Not Know

More sophisticated reliability-prediction methodologies attempt to address some limitations of traditional component-count approaches by incorporating mission profiles, environmental stresses, supplier quality, and manufacturing-process factors. But adding more variables does not solve the central evidentiary problem: the prediction requires representative knowledge of the stresses actually experienced by the relevant failure sites throughout the product population and life cycle. If those inputs are assumed rather than measured and validated, additional mathematical sophistication can create an appearance of precision without establishing predictive accuracy.

The age of MIL-HDBK-217 compounds this problem. The current technical content has not received a substantive update since Notice 2 in February 1995 [5]. In the more than 30 years since then, electronic materials, semiconductor technologies, packaging, interconnect density, printed-circuit-board construction, power density, operating speeds, manufacturing processes, software-hardware interactions, and global supply chains have changed dramatically. Jais, Werner, and Das identified the inability of handbook databases to keep pace with rapid advances in electronics as a major source of prediction error and concluded that simply updating the database and models would still exclude emerging technology and would not cure the underlying methodological limitations [10].

For the practicing engineer, relying on such an old methodology without demonstrated correlation to relevant modern field data can be analogous to using a medical textbook from a distant era to diagnose a modern patient. Some underlying principles may remain valid, but the materials, technologies, diagnostic knowledge, and available evidence have changed. The concern is not merely imprecision. Jais, Werner, and Das found that prediction errors can promote poor programmatic and design decisions and can divert effort from reliability-growth and Design-for-Reliability activities [10]. An authoritative-looking number can therefore do more harm than good when it creates false confidence or redirects engineering resources away from discovering actual weak links and failure mechanisms.

Even a valid physical life model faces a second and separate problem: the mission profile. A fatigue, wear, battery-degradation, solder-joint, or other physics-based model requires representative loading at the location where damage occurs. For many complex electronic products, manufacturers do not possess population-wide distributions of the stresses that customers will impose over the complete life cycle. A smartphone may experience widely different combinations of ambient temperature, solar heating in a parked vehicle, internal self-heating, charging temperature and frequency, mechanical shock, bending, vibration, humidity, contamination, software workload, and duty cycle. An automobile presents an even broader problem: electronics behind the windshield, in the passenger compartment, in a door, near a wheel, or under the hood can experience radically different local histories.

The external environment is only the first level of uncertainty. Reliability-relevant stress must ultimately be known at the potential failure site. A system-level temperature or vibration profile must be translated into the thermal cycles, strain, electrical loading, humidity, or other stresses experienced by a particular circuit board, solder joint, connector, semiconductor package, capacitor, battery, or interconnect. Thermal paths, structural resonances, shielding, airflow, mounting, neighboring components, and operating state can attenuate or amplify external stresses. Multiple stresses can also interact. Jais, Werner, and Das specifically note that product life-cycle loads include manufacturing, storage, handling, operating, and non-operating

conditions and criticize handbook approaches for failing to account adequately for temperature cycling, humidity cycling, vibration, mechanical shock, and their combinations [10].

This is why even physics-based modeling must be qualified by the quality of its mission-profile evidence. FIDES explicitly requires mission-profile characterization and acknowledges that local equipment conditions must be derived from system-level environments [9]. But a sophisticated model cannot recover information that was never measured. If the distributions of field stresses, their local transfer to the failure site, their sequencing, and their interactions are not known with sufficient fidelity, the resulting life estimate necessarily inherits those assumptions and uncertainties.

The empirical evidence also argues against assuming that handbook predictions become useful merely because they are applied consistently. In the AMSAA study reported by Jais, Werner, and Das, MIL-HDBK-217-based predictions not only differed greatly from demonstrated results but failed a rank-order correlation test for comparing systems [10]. Earlier SINCGARS data in the same paper showed similarly dramatic reversals between predicted and demonstrated MTBF. These results challenge the idea that an inaccurate absolute prediction can nevertheless be relied upon as a comparative metric.

Stress-to-limit testing approaches the unknowns from the opposite direction. Rather than assuming that every important future stress distribution and interaction can be characterized accurately, HALT experimentally explores the system until weak links and operating limits become observable. Failure analysis then identifies the mechanism and local conditions that produced the failure. Where a known wear-out mechanism is involved and representative loading data exist, a validated physical life model can then be applied. Where the mechanism or field loading is not sufficiently known, the engineering response should be to acquire evidence, discover margins, and reduce uncertainty rather than substitute a generic statistical failure-rate calculation.

11. What Reliability Engineers Should Measure

For a discovery-based program, the most actionable measures are often not predicted MTBF values but engineering margins and demonstrated process capability. Useful questions include:

- What are the lower and upper operating limits of the complete system?
- What are the destruct limits, and what physical mechanisms define them?
- How much margin exists between normal use and those limits?
- Which design parameters move the limits most strongly?
- What manufacturing variations reduce the margin?
- Can the relevant functions be monitored continuously while stress is applied?
- What production screen will reveal significant defects quickly without damaging good hardware?
- Do field-return root causes match the weaknesses discovered during development and production testing?

12. Conclusion - Reliability as Discovery and Margin Improvement

The central challenge of modern electronic reliability is not simply to calculate a number that represents future failure. It is to discover the mechanisms that can cause failure, determine how close the design is to those mechanisms, and increase the margin before customers encounter them.

For complex solid-state electronic systems, published evidence does not justify assuming that MIL-HDBK-217-type statistical component failure-rate predictions represent future field reliability or even provide a dependable comparative ranking. Jais, Werner, and Das reported predicted-to-demonstrated reliability ratios ranging from 1.2:1 to 218:1 across DoD systems and found no statistically useful rank-order correlation for comparing systems [10]. Their results also show why a standardized calculation should not be confused with an empirically validated reliability measure.

This conclusion is not a rejection of mathematical life modeling. Mechanical systems and structures can experience identifiable fatigue, wear, corrosion, creep, fracture, and other cumulative damage mechanisms. Electronic products can likewise contain batteries, solder joints, bond wires, die attach, connectors, capacitors, and other elements with identifiable degradation or wear-out mechanisms. When the physical mechanism is known, the controlling stresses and material behavior can be measured, the stress-life relationship is supported by representative empirical data, and the actual application loading is sufficiently characterized, mathematical models can provide valuable life estimates. The objection is to assigning generic statistical failure rates to complex solid-state electronics and treating the resulting system MTBF as predictive without demonstrated physical and empirical correlation.

Even a physically valid model cannot overcome an unknown mission profile. The reliability engineer must know not only the external environment but the stress history at the relevant failure site and how thermal, mechanical, electrical, humidity, duty-cycle, storage, handling, and other loads interact over time. When those distributions are unavailable, the uncertainty should be exposed rather than hidden behind a precise numerical prediction. HALT, root-cause analysis, physics-of-failure investigation, field data, and targeted life testing provide a path for converting unknowns into engineering evidence.

The most useful reliability question is therefore not merely, "What failure rate do we predict?" It is, "What is the weakest link, what physical mechanism makes it weak, what stresses actually act on it, and how much margin can we create before the product reaches the customer?"

References

[1] IEEE Std 1413-2010, IEEE Standard Framework for Reliability Prediction of Hardware. Source

[2] IEC 61709:2017, Electric components - Reliability - Reference conditions for failure rates and stress models for conversion. Source

[3] National Research Council, Reliability Growth: Enhancing Defense System Reliability, National Academies Press, 2015. See discussion comparing MIL-HDBK-217 with physics-of-failure approaches. Source

[4] A. Gaonkar, R. B. Patil, D. Das, M. H. Azarian, B. Sood, and M. Pecht, "Assessment of the FIDES Guide 2022 Electrical, Electronic, and Electromechanical Reliability Prediction Methodology," e-Prime - Advances in Electrical Engineering, Electronics and Energy, 2023, 100353. Source

[5] U.S. Department of Defense ASSIST, MIL-HDBK-217, Reliability Prediction of Electronic Equipment, current status listing. Source

[6] Institute of Environmental Sciences and Technology (IEST), IEST-RP-PR003, HALT and HASS. Source

[7] G. K. Hobbs, "Pitfalls to Avoid in HALT and HASS," Hobbs Engineering. Source

[8] M. Pecht et al., Physics-of-Failure approaches to electronics reliability and prognostics; overview in Prognostics and Health Management of Electronics, Wiley. Source

[9] FIDES, FIDES Guide 2022 Edition A (July 2023), sections on mission profiles and reliability prediction methodology. Source

[10] C. N. Jais, B. Werner, and D. Das, "Reliability Predictions - Continued Reliance on a Misleading Approach," U.S. Army Materiel Systems Analysis Activity and Center for Advanced Life Cycle Engineering, University of Maryland. The paper compares MIL-HDBK-217-based predictions with demonstrated military-system reliability and discusses technical limitations and reliability-assessment alternatives.

Author's Note

The two case histories and the original figures in this article are based on the author's prior engineering and consulting experience and were preserved from the May 7, 2026 version of this article. Published references are provided to distinguish the broader industry and academic literature from the author's experience-based observations.