The Merits and Limitations of Reliability Predictions

Manthos Economou, Brocade Communications Systems, San Jose

Key Words: Reliability Prediction, Reliability Models, Product Development, Failure Rate

SUMMARY AND CONCLUSIONS

A lot has been said and published about the limitations of reliability predictions based on the models first introduced by MIL-STD-217 (now MIL-HDBK-217) and commercialized by many. Some of the information in the literature denounces these methods as inaccurate and unreliable and promotes qualitative methods of ensuring reliability such as HALT/HASS, or quantitative methods such as Physics of Failure, Accelerated Life Models, etc. This paper will present the merits and limitations of reliability predictions as contrasted to reliability testing and assurance techniques from a product development standpoint. It will also attempt to answer questions such as: are MIL-based reliability prediction methods useful? At what stages of the product development process? Which elements of the prediction can be practically used, and which should be discounted? How can the accuracy of reliability predictions be improved? Every method offers a certain benefit at a certain cost, is limited by a time element. No single answer exists in accurately predicting and demonstrating reliability. Balancing cost, benefit and time, the essential elements of a new product reliability & quality assurance program, provide a framework for selecting the methods. Specific, theoretical and practical examples will be used to demonstrate the concepts and illustrate the methods that have been successfully used with encouraging results. In addition, useful interpretations of reliability predictions will be presented, since it appears many popular misconceptions exist in the electronics industry.

1. INTRODUCTION

Reliability prediction methods and data, based on MIL-STD-217 (now MIL-HDBK-217) have been used for decades as a means to consistently produce an estimate figure for the predicted reliability of a product. However, these methods and data have been denounced by some reliability professionals as inappropriate due to their lack of accuracy. Other methods involving qualitative and quantitative techniques have been promoted to predict and improve the reliability of a product. Since one of the main reasons to predict reliability is to ultimately improve reliability, the alternatives to U.S. military-based (MIL-based) reliability predictions make a lot of sense. An entire industry of reliability consultants and companies has been created to fulfill demand for products and services towards reliability improvement and improved reliability predictions. Are these people bashing the MIL-based reliability prediction methods to promote their own interests or is there a merit to their allegations? Perhaps both are true to one extent or the other.

Many studies indeed indicate that predictions based on MIL-STD-217 or derivative products do not agree with each

other, and are far from accurately predicting product reliability. It has been shown that results using these reliability prediction methods are usually conservative, and in many cases, the actual product reliability is several times better than the one predicted. However, the MIL-based reliability predictions are still extensively used and, more often than not, required by many customers.

If a model or process is not utilized correctly, then the popular cliché "garbage in, garbage out," becomes a reality. Many examples can be cited in various industries where mathematical models are used to predict physical behavior. Finite Element Analysis (FEA) and Computational Fluid Dynamics (CFD) are some of the tools used extensively in the electronics and other industries to properly design and predict behavior of physical elements. The success of modeling using any tool or method will greatly depend on the experience and skill of the user, and how the model is used. The more flexible or powerful the model, the worst the blunder, if not used properly.

Models can only approximate reality. As such, they are inherently inaccurate (the extend of inaccuracy depends on one's definition of accuracy) and their value is higher in comparative studies than exact results. Human-related factors, such as design, test, skill, training, quality, service & maintenance are directly coupled to reliability and extremely difficult to forecast, quantify or model properly, regardless of the method used. In addition, the comparison of prediction results and field failures can be completely characterized only when all products have failed and true times to failure have been recorded.

Every method and process has advantages and limitations. The reliability professional should capitalize on the advantages while observing the limitations. No single method is a remedy for all problems. Several methods should be employed at different points in time, or during the product lifecycle, to properly predict, assess, characterize and improve product reliability.

2. THE CONSTANT FAILURE RATE ASSUMPTION

MIL-based reliability predictions based on constant failure rates, first introduced by MIL-STD-217, have been used for many years to estimate product and system failure rates and Mean Time Between Failures (MTBF). Numerous other commercial standards have spawn from MIL-STD-217 (Table A). The main premise is that reliability depends on a Parts-Count and Parts-Stress approach, where the reliability of individual components determines the reliability of the system or product. This is represented by simply adding the individual component failure rates to derive the total product failure rate. In addition, the main assumption, and the one that is challenged the most, is that failure rates of components are constant or failures are exponentially distributed.

Model	Description		
MIL-HDBK-217	Original worldwide standard (MIL-STD-217) for commercial & military electronic components		
Telcordia SR-	Original Bellcore standard for commercial grade		
332	electronic components		
PRISM	Originally developed by the Reliability Analysis		
	Center (RAC), incorporates process grading factors		
CNET 93	Developed by France Telecom		
RDF-2000	Newer Version of CNET 93 developed by UTE		
HRD-5	Developed by British Telecommunications plc		
GJB/z 299B	Chinese Standard		

Table A: Reliability Prediction Models/Standards

Although intuitively the constant failure rate assumption may seem unrealistic, consider the following: The typical electronic product has a lifecycle (from production release to end-of-life) between 2 to 5 years. In this time period of interest, and after the initial "early failure" stage, the constant failure rate (FR) assumption may not be unreasonable because most of the components in a product have not reached their wear-out stage. Even though in reality the FR may be slightly decreasing (depending on the length of the "early failure" stage), or increasing (beginning of wear out stage), it can be assumed to be constant for most practical purposes. The assumption of constant FR also greatly simplifies the mathematics involved in reliability calculations. If distributions other than the exponential are used, the accuracy of the prediction may improve, but the mathematics complexity will increase by several orders of magnitude. If the constant FR methods are used and interpreted properly, the penalty of a slight inaccuracy may worth the benefit of mathematical simplicity.

3. THE MEANING OF MEAN TIME BETWEEN FAILURES

Based on these assumptions the system failure rate (λ_{System}) and Mean Time Between Failures (MTBF) or Mean Time to Failure (MTTF) are represented by:

$$\lambda_{System} = \sum_{i=1}^{l=n} \lambda_i \, , \, n = \text{number of system components}$$
(1)

$$MTBF_{System} = MTTF_{System} = \frac{1}{\lambda_{System}} , \qquad (2)$$

based on the exponential distribution with,

a Probability Density Function (pdf),
$$f(t) = \lambda e^{-\lambda t}$$
, (3)

a Reliability Function,
$$R(t) = e^{-\lambda t}$$
, and (4)

a Failure Rate,
$$\lambda = \frac{f(t)}{R(t)} = \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}}$$
 independent of time (5)

Many reliability professionals have denounced the constant FR assumption as unrealistic and inappropriate in estimating or predicting reliability. At first glance this statement makes a lot of sense. Since constant FR does not vary with time, the implication is that a component or a system does not age or wears out. From life experience, it should be evident to everyone that time progression does indeed increase failure rate. Products do age and wear out

displaying much higher failure rates as time passes.

Even if the constant FR assumption sounds unreasonable at the beginning, what is the physical meaning of MTBF in terms of actual product failures? Based on the mathematical representation of the exponential distribution, the probability of failure F(t) for the exponential distribution (constant failure rate) is,

$$F(t) = 1 - e^{-\lambda t}$$
 or $F(t) = 1 - e^{-MTBF}$ and if, $t = MTBF$, (6)
 $F(MTBF) = 1 - e^{-\frac{MTBF}{MTBF}} = 1 - e^{-1} = 0.632 \text{ or } 63.2\%$

This result indicates that by the time the MTBF of a product is reached, 63.2% of the products in the field would have failed at least once. Contrary to some beliefs and mathematical treatment, MTBF should not necessarily be interpreted as that, on the average a product will fail at the time indicated by the MTBF, or that an average product is supposed to be operational when the time designated by the MTBF is reached.

As an example, consider the MTBF of a product stated at 500,000 hours or with the corresponding failure rate of 2,000 FIT. The 500,000 hours can be converted to years (assuming 24-hour operation of a product),

500,000 hours
$$\times \frac{1}{8,766 \frac{\text{hours}}{\text{year}}} = 57 \text{ years}$$

thus, 63.2% of the products will experience a failure in 57 years, or the probability that one product will fail in 57 years is 63.2%. Although mathematics indicate a mean life of 57 years, in reality the product would have failed long before 57 years due to wear out. Perhaps for some reliability professionals these numbers are useful, but for the average practitioner in the electronics industry these numbers are still cryptic and provide little insight. If the same number is viewed from another angle, such as expected failures per year, it will start making better sense. In this example, if the FR of the product is 2000 FIT, then the average failure rate (AFR), sometimes called the cumulative FR is defined as,

$$AFR = \lambda \times t = \frac{Number \ of \ Failures}{Power \ On \ Hours} \times t \tag{7}$$

and, for one year, the AFR is,

$$2,000 \times 10^{-9} \frac{\text{failures}}{\text{hour}} \times 8,766 \frac{\text{hours}}{\text{year}} = 0.0175 \frac{\text{failures}}{\text{year}}$$
 or 1.75%

In other words, an MTBF = 500,000 hours can be viewed as a 1.75% average yearly failure rate, and if 1,000 units are in the field, it should be expected that ~18 units will fail in the course of one year. This is probably one of the most useful interpretations of MTBF for most people. The MTBF can be simply considered as a mathematical figure used to produce an average yearly failure rate. Note that equation (7) does not presuppose an underlying distribution and MTTF (Mean Time To Failure) can be used instead of MTBF, which is only applicable to the exponential failure distribution.

In the same fashion, the probability of failure, F(t) can be considered as an indication of failure rate, and for one year using the exponential failure distribution can be calculated as,

$$F(t) = 1 - R(t) = 1 - e^{-\frac{t}{MTBF}}$$

$$F(t = 8,766) = 1 - R(t = 8,766) = 1 - e^{-\frac{8,766}{500,000}} = 0.0174 \text{ or } 1.74\%$$
(8)

From the results of equations (7) and (8), can be seen that the AFR and F(t) are practically identical, but not exactly the same. In fact, for low MTBFs, the AFR and F(t) results will diverge even further as seen in Figure 1.



Figure 1: Effect of MTBF on AFR & F(t)

Only for MTBFs $>\sim$ 50,000 hours the results converge to become identical for practical purposes. Caution should be exercised when using averages such as equation (7), since the results may be overly conservative for low MTBFs.

The reason for this difference is that F(t) considers the original population of units in the field, while AFR considers an ongoing population with additional infusion of units and repaired or replaced units that failed.

4. RELEVANCE OF CONSTANT FAILURE RATE

The constant FR assumption can be a good approximation of product field failures depending on the product lifecycle and nature of product. Figure 2 shows the typical constituent graphs of the "bathtub" curve: an early failure region (decreasing FR), a constant FR region (useful, normal, or service life) and a wear-out region (increasing FR). The FR of an electronic product with a maximum lifecycle of five years can be approximated well with a constant FR as shown in Case I. During the product's lifecycle, the FR is first underestimated (early failures) and then over-estimated by the constant FR. The early failures can be attributed to production ramp-up issues and/or due to TTM pressures of a product that has been introduced into the market prematurely. Releasing a product before all the early failure root causes have been identified and corrected, will increase the FR for the initial time period. On the average, the constant FR assumption seems to be fairly valid in this example due to error averaging. However, in Case II, where the product lifecycle is ten years, the constant FR may not be a good assumption since the failure rate is grossly underestimated in the last five years. As a result, spares over-purchased (due to FR overestimation) in

the first five years may have exhausted their shelf-life, before they are actually needed in the wear-out period.



Figure 2: Constituent Curves of the "Bathtub" Curve

5. MIL-BASED VS. OTHER METHODS

Reliability methods can be classified in four major categories: reliability predictions, qualitative methods, quantitative methods, and analytical methods.

Reliability predictions are based on database tools, such as MIL-HDBK-217, Telcordia SR-332, etc. Qualitative methods involve aggressive testing such as HALT/HASS, ESS, HAST, etc. Quantitative methods employ techniques such as Finite Element Analysis (FEA), Physics of Failure (PoF), etc. Analytical methods are a blend of reliability prediction tools and quantitative methods (Weibull analysis, life stress distributions, etc.).

MIL-based reliability prediction methods are mainly used to establish a baseline reliability figure while the design is still on paper. The components that comprise the product along with projected stresses and end-use environmental conditions are needed to derive a reliability prediction. Databases provide failure rates for different types of components. The databases are constructed by supplier-provided "field" failure data. Since field failures depend mainly on design and application, these data are not representative of all cases. By including as much data as possible, these databases tend to provide over conservative failure rates. No single model covers all components, and a combination of models may provide the best coverage. These methods are likely to provide more accurate results for a system containing more parts rather than for a small system due to variance averaging.

Qualitative methods are mainly employed to improve the reliability of a product rather than measure or derive it. These methods involve some type of accelerated testing environment, where the product is subjected to elevated stresses to precipitate latent failures or design weaknesses. Actual products and specialized equipment are needed to employ these methods. The benefit realized is considerable, but the cost and time demands may also be considerable depending on the type and extend of testing. The techniques used, occasionally precipitate failures due to the testing environment or method rather than the field or use environment. In addition, results may not be consistent across identical products, due to inconsistencies in the testing environment.

Quantitative methods are computationally intensive since the reliability of the product is derived mostly via computer simulation analysis. They require a design that can be modeled in a computer and a plethora of supporting data. Results can be fairly accurate, depending on the skill level and data availability of the modeling. However, using quantitative methods can be tedious, time-consuming, challenging, and may conflict with the aggressive TTM requirements of today's electronic products.

Analytical methods are a mixture of prediction and quantitative methods and require data obtained using qualitative techniques. These methods can provide quite accurate results in terms of deriving, measuring or proving reliability. They are versatile in terms of modeling without being limited to the exponential distribution. However, some of the required data are obtained by testing actual product.

Table B provides a comparison of the different methods in terms of their suitability for a particular task.

	Reliability Predictions	Qualitative Methods	Quantitativ e Methods	Analytical Methods	
Steady State Reliability Prediction	Yes (H), C(L)	No (L)	No (M)	Yes (H), C(M)	
Lifecycle Reliability Prediction	No (L)	No (M)	No (M)	Yes (H), C(M)	
System Availability Downtime, Outage	Yes (H), C(L)	No (L)	No (M)	Yes (H), C(L)	
Reliability Improvement	Yes (M), C(L)	Yes (H), C(H)	Yes (M), C(H)	Yes (M), C(L)	
Competitive Analysis	Yes (H), C(L)	Yes (M), C(H)	Yes (M), C(H)	Yes (H), C(M)	
Warranty Determination	Yes (M), C(L)	No (L)	No (M)	Yes (H), C(L)	
Repair Cost Determination	Yes (M), C(L)	No (L)	No (M)	Yes (H), C(L)	
Maintenance Cost Determination	Yes (M), C(L)	No (L)	No (M)	Yes (H), C(L)	
Spares Determination	Yes (M), C(L)	No (M)	No (M)	Yes (H), C(L)	
Result Confidence	Low	High	Medium	Medium	
Yes (H): Well suited Yes (M): Can be done, with limitations No (M): In general, not well suited, can be used in certain cases No (L): Not well suited C(L, M, or H): Low, medium, or high cost and effort					

Table B: Suitability of Reliability Methods

Can a model or method predict and/or help improve the reliability of a product? Yes and no. Depends how it will be used, how the results will be interpreted, and what actions will be taken based on the results. Are the actions effective? Only the field failure rate will determine the answer.

Is one method better than the other? Depends on one's philosophy and interests. Proponents of each method will defend their beliefs, often at the exclusion of other methods. A combination of models and processes is probably a better way of assessing and improving the reliability of a product. Every method has advantages and limitations and no method is appropriate for everything.

6. NEEDS FOR RELIABILITY PREDICTIONS

Some of the potential needs for reliability predictions in the electronics industry are:

1. <u>Reliability Improvement</u>. Using fewer or low FR components will result in lower field failures. Lower

temperatures and higher derating will improve reliability.

- 2. <u>Design Tradeoffs</u>. Choices between using a large number of low FR components versus using a lower number of high FR components. Component grade selection (industrial vs. commercial, etc.).
- 3. <u>Lifecycle Cost</u>. Determination of total product cost, on a yearly or lifecycle basis.
- 4. <u>System Availability</u>. Uptime and downtime prediction along with redundancy schemes to improve availability.
- 5. <u>Competitive Analysis, Benchmark</u>. Comparing MTBFs of competing products to get a first approximation of the products' inherent reliability as predicted by models. The higher the complexity of the product, the higher the probability of defects in the field. A common, consistent procedure has to be used to compare "apples with apples." It may not be totally accurate since reliability is design-dependent, but it provides a first order of comparison.
- 6. <u>Warranty</u>. Selection of a suitable warranty period. Overestimating the warranty period will increase total lifecycle costs, while underestimating could render the product less competitive as compared with like products.
- 7. <u>Repair Cost</u>. The cost of repairs during the warranty period. These are part of the product total cost that needs to be factored in during product pricing.
- 8. <u>Maintenance Cost</u>. Assessment of costs related to preventive maintenance of a product. High FR items will need to be replaced more often to minimize system downtime.
- 9. <u>Spares</u>. Determination of number and type of spare parts at the customer or distribution sites. High FR items need to be stocked at higher numbers.

How well a reliability prediction method can address the above outlined needs will determine its rate of success and ultimately its value.

7. VALUE OF MIL-BASED RELIABILITY PREDICTIONS

What is the best method to predict reliability of a product? Obviously, the method that will provide the most accurate results, in the least amount of time, with the least effort, and the least cost.

The value of a reliability prediction, process, method, or any other concept, has to consider benefits and costs. Thus, the simple, well-known equation: Value = Benefit - Cost. If the cost outweighs the benefit there is no real value. In addition, there is an element of time involved due to market pressures sometimes called Time to Market (TTM). Then, value, benefit, and cost are all functions of time: V(t) = B(t) - C(t). The value of a reliability prediction will be different at different points in time, since the benefits and costs vary with time.

The value of MIL-based reliability predictions is variable in time according to V(t) = B(t) - C(t), and depends on the product lifecycle phase. As seen on Figure 3, at the Concept phase, the value is quite high, reaching a maximum at the Development phase because the design is only on paper, with no product to test. The benefit B(t) is high since only blackbox methods can be used based on a bill of material (BOM) and the cost C(t) is low since the prediction is easily derived from established models and databases. The increase in value from Concept to Development is attributed to the design maturity and detail of the product definition.



In addition, when a new product gets introduced, the warranty period, maintenance and repair, and service support costs need to be determined to estimate the total cost of the product during its lifecycle. At this early stage, these quantities can only be estimated by a MIL-based prediction. From that point on, the MIL-based prediction methods decrease in value since Prototypes are built and the reliability of the product can start to be assessed via testing.

The Other prediction methods (via testing, modeling or field failures) have low value at the early stages of the lifecycle, but the value increases rapidly after prototypes are available for testing and the product matures. The value reaches a maximum at the Pilot phase since the products tested are representative of the products going to the field. After Pilot, the value slowly decreases but remains fairly high due to the historical importance of the field failure data. The maximum value of the Other prediction is shown as somewhat lower than the one for the MIL-based prediction because although the benefit of the Other prediction is high, the cost of acquiring this type of data is also high.

8. RELIABILITY PREDICTION VS. FIELD FAILURES

Based on several examples of telecommunication products from different companies, the failure rate behavior of a typical product "A" can be seen on Figure 4.

To protect proprietary information of the companies' data used to construct this graph, the FR numbers shown have been modified from the actual. Also, defects due to "No Defect Found (NDF)," customer errors, and certain abnormalities such as an occasional lot of defective parts, product abuse during shipment, etc., have been removed.

Product "A" reliability prediction on the concept and development phase resulted in a FR=5%. The expectation of this type of product was a FR=1%. Data collected from product "A" field failures for 2 years yielded an average FR=2%. During the first six months, the FR has large fluctuations since few products are in the market, a small number of failures will have a large effect on the FR, and some early failures will escape the factory. After six months, the FR seems to stabilize, until the 9th month where it suddenly sharply increases. This is typical of a large customer

or distributor accumulating failed units and shipping them to the service or repair center all at once. After the first year, the FR seems to be fairly stable at $\sim 2\%$, slightly increasing with time.



Figure 4: Typical FR Behavior of a Telecom Product.

The predicted FR is ~2.5 times higher than the actual. This is expected since MIL-based reliability prediction models are inherently conservative. Telcordia SR-332 indicates that reliability predictions using the generic failure rates in the model are "90% upper confidence level point estimates. This means that there is a 90% chance that the actual device generic failure rate is lower..." And, "there is at least a 90% chance (e.g., perhaps a 95% chance) that the actual failure rate for a unit is lower than the value predicted..." Furthermore, the data used on MIL-based predictions are usually outdated (conservative). By the time data are collected, compiled, verified, published, and used, supplier improvements on component design and manufacturing may have rendered components more reliable.

In addition, using this model is valid for only the middle section of the "bathtub" curve, which best approximates a steady state or constant failure rate. Typically, early failures are screened out in the factory, and wear out for electronic components are much beyond the 2-5 years, which is the lifecycle of a typical telecommunications electronic product. If the product is designed for operating much longer than 5 years, then a different approach should be employed for predictions in the wear out region of the bathtub curve.

Based on product "A" history, a first approximation in the reliability prediction of a subsequent product "B," can use the 2.5 correction factor to bridge the gap between predicted (using a model) and actual reliability. This approximation technique has been used very successfully in the past with accuracy ranging from 80% to 100%. The black box reliability prediction can be used as a baseline, and with the appropriate correction factor, the true reliability can be approximated.

If the typical reliability prediction model follows a blackbox approach, then the predicted reliability will be highly inaccurate (usually conservative rather than optimistic) depending on the model used and how it is used. Most of the time, the reliability prediction is several times more conservative than the actual field behavior. Several methods can be employed to improve accuracy, such as laboratory and field data from suppliers, along with company laboratory and field data (Figure 5).



Figure 5: True Reliability Convergence with Data Infusion.

With every infusion of more deterministic data, the prediction becomes more accurate approaching the true reliability of the product. Typically, a predicted and a demonstrated reliability will be required from a supplier. The demonstrated reliability is usually derived from supplier laboratory (new product) or field (existing product) data. Even the laboratory methods, which use some type of life acceleration technique, would not reveal the true nature or time of failures. This is because debatable acceleration models are used for calculations, failures may be induced due to the testing method, and the products are tested under controlled conditions, while in the field, conditions are not controlled and unpredictable.

Many other factors are involved in field failures. A major factor is software, which has become a different reliability field by itself. Software is usually not adequately addressed in a MIL-based reliability prediction although it may count for a large number of failures in the field, and failures due to software can even be misconstrued as hardware failures. Other "soft" field failures such as cosmetic, labels, etc., are difficult to classify or predict and render field failure monitoring a challenge.

Comparing prediction to field failure data is a challenge due to the time it takes to accumulate, receive, and process meaningful data from the field. Furthermore, field data are biased due to usage, timing, and conditions. Simple processing of this highly convoluted data, without detailed knowledge of the time and cause of failure, can lead to incorrect conclusions.

9. CONCLUSIONS

MIL-based reliability prediction methods are consistent, mathematically simple, but inherently inaccurate, usually erring on the conservative side. This limitation may be overcome with the use of historical data and appropriate correction factors rendering reliability predictions quite accurate on a practical level. The reliability predictions should not be taken at face value, but as a figure of merit or adequate baseline towards comparative studies of design alternatives, evaluation of competitive products, or early forecasting of the total lifecycle cost of a product. The value of these methods is very high at the early stages of the product development where no physical product exists, but the value decreases rapidly as prototypes become available for testing. Employing these methods can increase the baseline reliability of a product, but since product reliability depends mainly on design and enduse conditions, Mil-based reliability predictions are not appropriate in proving or improving the field reliability of a product. Other methods and techniques, which can be analytical, qualitative or quantitative, should be used to prove and improve the field reliability of a product.

REFERENCES

- 1. MIL-HDBK-217
- 2. Telcordia SR-332
- 3. <u>www.reliasoft.com</u>
- 4. <u>www.relexsoftware.com</u>
- 5. Paul A. Tobias, David C. Trindade, Applied Reliability, 2nd edition.

BIOGRAPHY

Manthos Economou Brocade Communications Systems, Inc. 1745 Technology Drive San Jose, CA 95110 <u>meconomo@brocade.com</u>

Manthos Economou is a Senior Staff Engineer – NPI & Reliability at Brocade Communications Systems. He spent twelve years at Hewlett-Packard on Product & Manufacturing Development, and three years at Nortel Networks on Reliability and NPI Management. He received his B.S. & M.S. in Mechanical Engineering from Illinois Institute of Technology, and an M.S. in Engineering Management from Santa Clara University. His current interests include reliability predictions, reliability assessment & budgeting, and product reliability & quality assurance.