RELIABILITY GROWTH CONCEPTS AND TESTING

To improve the reliability of a product, failure data must be collected, the root cause of failures identified, and corrective action(s) implemented. As causes of failures are eliminated by the incorporation of corrective action(s), the reliability of the product will have the potential to improve, or *grow,* to the desired or required level of reliability. This process to eliminate causes of failure is known as reliability growth. The failure data collected for a reliability growth process can originate from a formal reliability growth test, from analysis of failures that occurred during other testing, and/or from analysis of products returned for warranty repairs. The root cause of failures are identified through nondestructive and destructive failure analysis, stress-versus-strength analysis, and/or further testing. The root cause could then be traced back to deficiencies caused by the production process, workmanship, design, test procedures, or operator error. Whatever the cause, the important point to emphasize for reliability improvement is to implement corrective actions.

HISTORY OF RELIABILITY GROWTH

Reliability growth is a reliability improvement technique that has been the foundation for successful products for many decades. In 1946 Aeronautical Radio, Inc. (ARINC), collected and analyzed defective electronic tubes for commercial airlines. In cooperation with the tube manufacturers, the causes of failures were identified and corrective actions were implemented. By the 1950s, ARINC's tube reliability efforts had been applied to military applications to investigate the nature and probable causes of tube failure and the conditions that contribute to it, and, so far as practicable, to make recommendations for application of appropriate corrective measures to eliminate or reduce causes of tube failure (1). J. T. Duane in 1962 developed the learning curve approach to reliability monitoring (2). He recognized a general trend in the improvement of various products under development in terms of the cumulative failure rate. The products included hydromechanical devices, aircraft generators, and an aircraft jet engine. The cumulative number of failures, plotted on log–log paper as a function of cumulative operating hours, produced a nearly straight line for all of the products. The slope of the line showed the rate of growth and indicated the effectiveness of the reliability growth process in identifying and correcting design deficiencies. Duane's work presented the foundation for the monitoring and projection techniques used in military standards and handbooks. In 1975 L. H. Crow (3) observed that the Duane model was equivalent to a nonhomogeneous Poisson process with a Weibull intensity function.

The military standards and handbooks defined the reliability growth methodology as a test, analyze, and fix (TAAF) process (4). Through the late 1960s to the late 1980s the military recognized the importance of conducting TAAF and recognized that the reliability of the drawing board design of a complex product can be improved, and time (and funding) was allocated for that improvement. In 1989 the military emphasized that the TAAF process should not be used to qualify or validate the reliability of a product. The goal is to find failures and eliminate them (5). In June of 1994 the Secretary of Defense issued a memorandum mandating that the military abandon the use of levying military standards and handbooks for defining products in government contracts and use commercial off-the-shelf products as much as possible. Although military standards and handbooks were now obsolete, the planning, monitoring, and projection techniques presented in those documents are still in use today by the commercial manufacturers.

As a requirement for staying ahead of competition, manufacturers of commercial products also conducted forms of reliability growth. Examples are an automobile manufacturer studying warranty returns and implementing improvements in the next model year, and a semiconductor manufacturer conducting accelerated life testing on the next-generation integrated circuit. Program managers soon realized that the reliability growth concept complements the key aspect of total quality management (TQM), that is, to improve continuously (6). To be successful in improving reliability, the growth process must be cost-effective and well-managed with the primary goal of identifying and eliminated failures.

RELIABILITY GROWTH PROCESS

Reliability growth can be applied to hardware or software products or both. Hardware includes items such as a component (semiconductors or integrated circuits), a circuit card assembly, or a complete system (computer, radar, or radio transmitter). Software includes the programs (source code) that operate the systems. When a product requires improvement in the reliability, the level of improvement (say 2, 3, or even 10 times better) is usually determined by calculating the cost of the improvements plus the total life cycle cost (LCC) for the product. For the military, the LCC includes the total cost of the product, which includes acquisition, development, and operating costs. For commercial applications, the LCC usually includes only the manufacturers' perspective of cost, which comprises development and a portion of the operating costs to cover warranty returns (7). The remaining cost, the total cost of ownership (COO), is accounted for by the customer (consumer).

Once the LCC is quantified, the next step is to plot it versus reliability. A graph of LCC versus reliability indicates to the decision makers whether implementing a growth process at that moment would be cost-effective. The decision to proceed could be disastrous if made without full understanding of the technology, complexity, producibility, and suitability of the product. Once it is determined that it will be cost-effective to improve the reliability, there are four key aspects to a successful growth process: planning, data collection, failure analysis, and corrective actions.

Planning

The first step in planning a growth process is to determine the methodology that will be used to collect the failure data. The methodology can be a formal TAAF process (often called a reliability growth test), an accelerated life test, or the use of data from other tests (such as environmental, functional, or safety) and/or from analysis of products returned for warranty repairs. In general, complex items or newly developed items will require a formal TAAF process or an accelerated life test, while low-complexity items or commercial off-the-shelf items may be treated using data from other tests or warranty returns. Procedures for conducting accelerated life testing are discussed in the article "Accelerated life testing in reliability." Factors to consider when determining the methodology are the state of the art of the product, the cost associated with test personnel and resources, the schedule, the number of units under test, the accuracy of the data collected, the test environment, the planned growth, and the failure reporting procedures. The time and funding available to conduct a growth process have the largest influences in selecting the methodology. Performing a trade-off analysis is often beneficial in selecting the methodology. Sample trade-off results for a radio transmitter (RT) are shown in Table 1.

Once the methodology is selected, the next step is to develop ground rules for reviewing failures. Ideally, if there were unlimited resources, time, and funding, all failures should be investigated and corrected. However, due to funding and/or schedule constraints, this ideal case usually does not exist. In addition, during the test, failures will occur that are considered nonrelevant, such as those caused by accidental damage or operator error, failures of the test equipment, or an externally induced overstress that exceeded the design

Table 1. Sample Trade-off Criteria for Radio Transmitter

Methodology	Cost of Conducting Methodology	Number of Items Available for Testing	Test Environment	Length of Test	Advantages (A) and Disadvantages (D)
Formal TAAF	Highest	Limited	Simulation of actual conditions	1000–5000 h	A: controlled test D: calendar time usually double actual test time
Accelerated test	Midrange	Limited	Simulation of actual conditions (may exceed)	250–1000 h	A: expeditious results D: results related to actual conditions
Utilization of other tests	Midrange	Limited	Simulated or actual	Limited, <500 h	A: minimizes duplication of effort D: data limited
Warranty returns	Lowest	All units sold	Actual conditions	1–5 years	A: no test resources required D: exact failure environment may not be known

limits of the product. Time is needed to troubleshoot all failures and develop and implement corrective actions.

During reliability growth testing, the total operating time accrued on the product is only a portion of the total calendar time allotted. Test efficiency is determined by dividing total operating time by calendar time. Experience has shown that most formal TAAF programs have a test efficiency of around 50% (8). Factors affecting the test time are troubleshooting, environmental chamber failures, inadequate quantity of spare test units, and poor supervision.

Data Collection

The accurate collection of data during a growth process plays a significant role in the subsequent failure analysis and corrective action implementation. Formally, a closed-loop corrective action process, termed the failure reporting and corrective action system (FRACAS), is usually established to ensure implementation. The more successful growth programs will collect time-to-failure data, system operating parameters, and information on operating environment(s) at time of failure. Performance data can be collected by recording and monitoring any built-in test parameters and/or by recording data obtained from common test equipment such as spectrum analyzers, digital multimeters, signal analyzers, oscilloscopes, or logic analyzers. Test environment data are collected by using vibration monitor accelerometers, temperature recorders, strain gauges, etc. These data should be available in all methodologies except the review of warranty returns.

During a review of warranty returns the actual use environment may have deviated from the recommended use environment—for example, using a device rated to operate from 0° to 70°C in an application environment outside that range. In addition, the actual performance parameters at time of failure are typically not available. Despite these missing-data problems, eliminating the cause of failures from warranty returns can still provide a cost-effective method of improving the reliability of a product.

Failure Analysis

Once the failure data are collected, the failure analysis process can begin. Of the four key aspects, failure analysis and the subsequent corrective action implementation are two most important. In electrical and electronics engineering, common failure modes include component failures (e.g., disbonds, delamination, die tilt, stress cracks), cracked solder joints, board delaminations, software errors, and loose or broken wires and/or cables. There are five steps needed to determine the root cause:

1. Complete a failure history by documenting the mode failure along with part numbers, revision number, nomenclature, time to failure, environmental conditions, parametric conditions, and description of failure event.
2. Verify the failure by conducting system level tests.
3. Isolate the failure to the lowest repairable assembly, typically a component (integrated circuit, diode, capacitor, etc.).
4. Analyze the failed part using the most cost-effective method.
5. Identify the root cause of the failure.

Using a personal-computer warranty return as an example, we begin by documenting the failure history. The computer user reported that the internal pulse-code modulated modem does not work. The initial testing is to determine whether the problem was caused by hardware or software. System level diagnostics confirmed the modem card had failed. The modem card failure was isolated to a bad digital signal processor chip. Electrical testing confirmed that there was an opening between pins 18 and 32. Scanning electron microscopy revealed that the bond had lifted from the die. Further failure analysis revealed contamination around the bond pads. The cause was isolated to the manufacturing process, where the procedures were modified to prevent the recurrence of the contamination.

In the above example, a structured failure analysis flow was followed. The goal was to perform low-cost nondestructive testing first, and then, if the root cause was still not determined, more elaborate and destructive tests. Other root causes that could have been determined include thermal overstress, electrical overstress, wearout, mechanical damage, or corrosion.

Further information on failure analysis flow and root cause analysis can be found in Refs. 9 and 10.

Typical equipment required for failure analysis includes delidding and cross-sectioning tools, curve tracers, microscopes, power supplies, oscilloscopes, and multimeters. More advanced techniques include equipment to perform scanning electron microscopy, energy-dispersive X-ray spectroscopy, and scanning acoustic microscopy.

At times it may be difficult to duplicate the failure (step 2). In these instances the environment that the product failed in may have to be simulated. The example above discussed a failure of a personal computer, which probably was used in a room temperature (25°C) environment. If the product was a radio transmitter used on an aircraft, the failure might only occur at high altitude, with temperatures exceeding 55°C and relative humidity at 85%. When tested at 25°C the radio transmitter might operate satisfactorily. In this instance, system level testing may have to be performed in an environmental chamber in order to duplicate the failure. In other instances, even with simulated environments, the failure may still not be duplicated. The failure may be intermittent or caused by incorrect use by the operator, or (as often happens in the case of complex systems or those that push the state of the art) there may be built-in test (BIT) inadequacies. The BIT inadequacies may provide incorrect isolation information or provide an indication of failure when one does not exist. These false BIT indications are termed false alarms. As part of the reliability growth process false alarms should also be investigated to determine their root causes. False alarms can be avoided by sampling more often, modeling in greater detail, increasing test tolerances, executing a test repeatedly, or correlating the test indication with other testing (11).

Corrective Actions

We have planned and selected the growth methodology, collected failure data, and determined the root cause of the failure. The final step is to develop and implement corrective actions. The effectiveness of corrective action will be discussed later in this article. By determining the root cause, a corrective action can be developed to eliminate the cause of failure and minimize the chance of that specific failure recurring. The root cause will provide insight into the type of corrective action needed. Common corrective actions from reliability growth tests include hardware design improvements, manufacturing process changes, manufacturing documentation clarifications (to prevent human error), component relocation (board placement), component mounting method changes, component operating parameter derating, software updates and/or corrections, and material selection changes.

There are instances where a corrective action will not be implemented, due to insufficient data, lack of resources (cost or equipment), intermittency of failures (making it impossible to determine the cause), or absence of failure trends. Often there are cost constraints on further analysis that prevent corrective action. In a product that may contain hundreds or thousands of components, it may not be cost-effective to implement a corrective action on the first occurrence of a component failure. In addition, failure analysis may not be conducted on the failed component until a pattern failure exists. A pattern failure is defined as the occurrence of two or more failures of the same part used in the same environment with the same failure mechanism.

However, it is emphasized that, as stated earlier, the only way to achieve reliability growth is to incorporate corrective actions. Program managers (producers) must be aware that although there may not be a short-term cost benefit, there usually is a long-term cost benefit that will help ensure that their customers are satisfied and their business will continue.

MONITORING AND PROJECTION

As corrective actions are implemented and the reliability of the product improves, management often requires a metric to report the progress. Continuous models for repairable products and discrete models for nonrepairable or one-shot products (missiles, rockets) have been developed. For repairable products, the Duane model has been the most common model used.

Duane Model

Duane noticed that the cumulative number of failures, plotted on log–log paper as a function of cumulative operating hours, produced a nearly straight line for the products he tested. The slope of the line showed the rate of growth and indicated the effectiveness of the reliability growth process in identifying and correcting design deficiencies. This phenomenon is mathematically modeled as

$$\lambda_\Sigma = \frac{\sum F}{t} = Kt^{-\alpha} \tag{1}$$

where $\lambda\Sigma$ is the cumulative failure rate, $\sum F$ is the cumulative number of failures, t is the cumulative number of operating hours, K is a constant indicating an initial failure rate, and α is the growth rate. The growth rate α must be between zero and one to model a decreasing failure rate. A growth rate that approaches one represents the maximum growth process achievable. A growth rate of 0.3 to 0.5 is generally accepted as a reasonable value for planning purposes (12).

Upon the completion of a reliability growth process, the current or instantaneous failure rate of product λ_I, can be found from the derivative of the cumulative number of failures, $\sum F$:

$$\lambda_I = \lim_{\Delta \to 0} \frac{\Delta(\sum F)}{\Delta t} = \frac{d(\sum F)}{dt} = (1-\alpha)Kt^{-\alpha} \tag{2}$$

The instantaneous mean time between failures (MTBF) can also be determined graphically by plotting the failures on log–log paper. For each failure, the point estimate cumulative MTBF is determined. The cumulative MTBF is plotted on the y axis and the time to failure on the x axis. A straight line is fitted to the points. The instantaneous MTBF is then determined by drawing a line parallel to and displaced by a factor of $1/(1 - \alpha)$ above this cumulative line.

The Duane model is also useful for predicting or planning the expected or desired reliability growth. An idealized growth curve, as shown in Fig. 1 with sample data, can be developed if an initial MTBF is known and a slope is assumed (12). Crow (13) provides further techniques for determining the initial MTBF. It is important to note that this idealized growth curve serves only as a guideline for assessing progress in terms of the schedule. The desired reliability can only be met if deficiencies are detected, failures analyzed, and corrective actions are implemented. Using Eq. (1), the planned total cumulative test time, t_c, can be derived as

$$t_c = t_i \left(\frac{\Theta_R}{\Theta_i}\right)^{1/\alpha} \tag{3}$$

where t_i is the initial test time, Θ_i is the initial MTBF, Θ_R is the cumulative or required MTBF, and α is the growth rate. Using Eq. (2), the planned instantaneous test time, T_i, can be derived by first converting the initial failure rate, K, to the equivalent instantaneous failure rate. This is derived as $\lambda_I = K(1 - \alpha)$ where λ_I is equal to the instantaneous failure rate, and K is the failure rate to be converted. Then the instantaneous test time (the time for the instantaneous MTBF to reach the desired MTBF) can be determined from

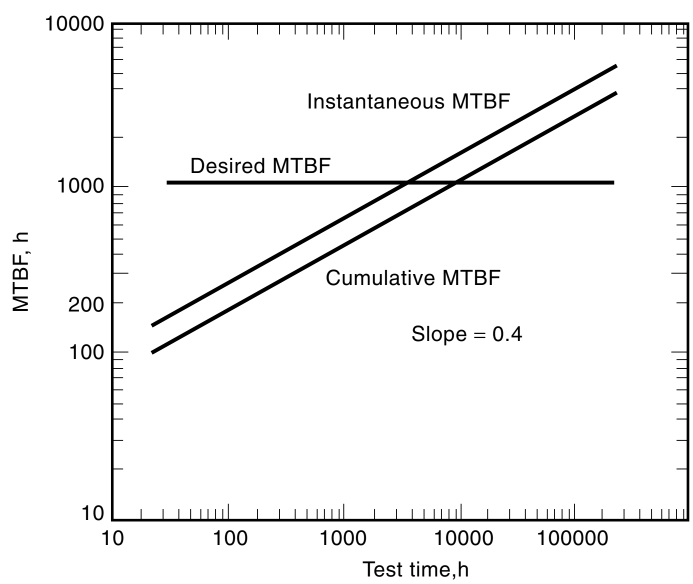

Figure 1. Planned growth curve.

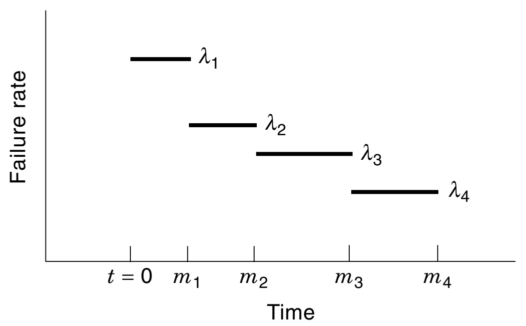

Figure 2. Observation of failure rates between improvements.

$$T = t_{\mathrm{i}} \left(\frac{\lambda_{\mathrm{I}}}{\lambda_{\mathrm{ID}}} \right)^{1/\alpha} \tag{4}$$

where λ_{ID} is equal to the instantaneous failure rate equivalent to the desired MTBF.

US Army Material Systems Analysis Activity Model

In 1975 L. H. Crow (3) observed that the Duane model was equivalent to a nonhomogeneous Poisson process with a Weibull intensity function. This observation was the basis for the development of the AMSAA model. The AMSAA model is explained as follows. Let $m_1 < m_2 < \cdots < m_k$ represent the cumulative test times when design modifications (reliability improvements) are made. The failure rate can be assumed to be constant between the times when reliability improvements are made, as illustrated in Fig. 2. Let λ_i represent the failure rate during the ith time period between modifications (m_i, m_{i-1}). Since constant failure rate is assumed, the number of failures, N_i, during the ith time period has a Poisson distribution with a mean number of failures $\lambda_i(m_i - m_{i-1})$. This is expressed mathematically by

$$\mathrm{Prob}[N_i = n] = \frac{[\lambda_i(m_i - m_{i-1})]^n e^{-\lambda_i(m_i - m_{i-1})}}{n!} \tag{5}$$

where n is an integer.

Let t represent the cumulative test time, and let $N(t)$ represent the total number of product failures by time t. If t is in the first interval (0 to m_1), then $N(t)$ has a Poisson distribution, with mean $\lambda_1 t$. If t is in the second interval, then $N(t)$ is derived by summing the number of failures, N_1, in the first interval with the number of failures in the second interval between m_1 and t. Therefore, in the second interval, $N(t)$ has the mean $\Theta(t) = \lambda_1 m_1 + \lambda_2(t - m_1)$.

When the failure rate is assumed to be constant over a test interval, that is, between reliability improvements resulting in design modifications, then $N(t)$ is said to follow a homogeneous Poisson process, with a mean of the form λt. When the failure rates change with time between two intervals, then $N(t)$ is said to follow a nonhomogeneous Poisson process. For monitoring the reliability growth between design modifications, $N(t)$ follows the nonhomogeneous Poisson process, with the mean value function

$$\Theta(t) = \int_0^t \rho(y)\, dy \tag{6}$$

where the intensity function $\rho(y) = \lambda_{\mathrm{i}}$, and $dm_{i-1} < y < dm_1$. Thus, the probability that exactly n failures occur between the start of the reliability growth process and the completion (total accumulated hours) is

$$\mathrm{Prob}[N(t) = n] = \frac{[\Theta(t)]^n e^{-\Theta(t)}}{n!} \tag{7}$$

where n is an integer. As Δt approaches zero, $\rho(t)\,\Delta t$ approximates the probability of a product failure in the time interval $(t,\ t + \Delta t)$. If the intensity function $\rho(t)$ is equal to λ, a constant for all t, then the failure probability is not changing over time (no trend is established). If $\rho(t)$ is decreasing ($\lambda_1 > \lambda_2 > \lambda_3 > \cdots$), then the failure probability is decreasing, implying reliability growth. If $\rho(t)$ is increasing, the reliability of the product is deteriorating.

The AMSAA model assumes that the intensity function $\rho(t)$ can be approximated by a parametric function defined as $\rho(t) = \lambda\beta t^{\beta-1}$, $t > 0$, $\lambda > 0$, $\beta > 0$, which is recognized as the Weibull failure rate function. When $\beta = 0$, the failure rate or intensity function $\rho(t)$ is constant (exponential case), which is analogous to the homogeneous Poisson process. If $\beta < 1$, then the reliability is improving. If $\beta > 1$, then the reliability is deteriorating. It is noted that the AMSAA model assumes a Poisson process with Weibull hazard rate function. This is not the Weibull distribution; therefore, statistical procedures for the Weibull distribution do not apply. From Eq. (7) the probability that exactly n failures occur between the start of the growth process and the final time t_0 can be determined. The parameter $\Theta(t)$ is the mean value function or, in other words, the expected number of failures expressed as a function of time. For the reliability growth process this function is of the form $\Theta(t) = \lambda t^{\beta}$ with λ and β both > 0.

The cumulative failure rate, $\lambda\Sigma$, can be defined as $\lambda\Sigma = N(t)/t$. If $N(t)/t$ is linear with respect to t on a log–log scale, then this pattern is analogous to the idealized growth pattern recognized by Duane. When no additional reliability improvements (corrective actions) are incorporated after time t_0, future failures will follow an exponential distribution. The instantaneous mean time between failure (MTBF) of the product is obtained as the function $m(t) = (\lambda\beta t^{\beta-1})^{-1}$. The instantaneous MTBF represents the product MTBF that was achieved during the growth process.

The parameters λ and β can be determined graphically from a log–log plot or determined statistically using estimation theory. For statistical estimations the method of maximum likelihood can provide estimates of the parameters λ and β. These statistical estimates can only be used if a nonhomogeneous Poisson process is present. If a significant failure trend, either increasing or decreasing, is not present, then a homogeneous Poisson process exists. One test used to identify such trends is the *central limit theorem test,* or Laplace test (14). If the period of observation ends with a failure (failure truncated), use the test statistic μ_1 generated by

$$\mu_1 = \frac{\sum_{i=1}^{M} X_i - M X_N/2}{X_N (M/12)^{0.5}} \tag{8}$$

where M is the number of failures (N) minus 1, X_N is the time of the last failure, and X_i is the time of the ith failure. If the failure data are time-truncated, use the test statistic μ_2 generated by

$$\mu_2 = \frac{\sum_{i=1}^{N} X_i - Nt_0/2}{t_0(N/12)^{0.5}} \quad (9)$$

where N is the number of failures and t_0 is the total test time. The statistic μ is compared with the standardized normal deviate at the chosen level of significance, Z_α, and if:

- $\mu \leq Z_\alpha$, then significant growth is indicated at the chosen level of significance and the maximum likelihood estimators can be used for estimating parameters λ and β;
- $\mu \geq Z_\alpha$, then significant reliability decay is indicated at the chosen significance level and further corrective action and design changes are needed;
- $-Z_\alpha \leq \mu \geq Z_\alpha$, then the trend is not significant at the chosen significance level, since the data (failure rate) follow a homogeneous Poisson process; additional data should be accumulated.

Critical values of the test statistics can be found in the normal distribution tables. Common two-sided significance level test statistics are 1.960, 1.645, and 1.282 for 5.0, 10.0, and 20.0% levels of significance. If $\mu \leq Z_\alpha$, the estimates of λ and β can be determined by the method of maximum likelihood. For failure-truncated tests, the biased estimate of β is

$$\hat{\beta} = \frac{N}{(N-1)\ln X_N - \sum_{i=1}^{N-1} \ln X_i} \quad (10)$$

The unbiased estimate of β can be determined by multiplying the biased estimate by $(N - 2)/N$:

$$\bar{\beta} = \frac{N-2}{(N-1)\ln X_N - \sum_{i=1}^{N-1} \ln X_i} \quad (11)$$

For failure-truncated tests, the biased estimate of λ is

$$\hat{\lambda} = \frac{N}{X_N^{\hat{\beta}}} \quad (12)$$

The unbiased estimate of λ is

$$\bar{\lambda} = \frac{N}{X_N^{\bar{\beta}}} \quad (13)$$

For time-truncated tests, the biased estimate of β is

$$\hat{\beta} = \frac{N}{N\ln t_0 - \sum_{i=1}^{N} \ln X_i} \quad (14)$$

The unbiased estimate of β can be determined by multiplying the biased estimate by $(N - 1)/2$:

$$\bar{\beta} = \frac{N-1}{N\ln t_0 - \sum_{i=1}^{N} \ln X_i} \quad (15)$$

The biased estimate of λ is

$$\hat{\lambda} = \frac{N}{t_0^{\hat{\beta}}} \quad (16)$$

The unbiased estimate of λ is

$$\bar{\lambda} = \frac{N}{t_0^{\bar{\beta}}} \quad (17)$$

To determine if the collected data fit the AMSAA model, a Cramer–von Mises goodness-of-fit test is used. Table 2 is used to determine the critical value for the test statistic, C_M^2. At the chosen level of significance (α), the indexing parameter is $M = N - 1$, where N is the number of failures that occurred during the growth process. The value calculated from one of the following equations, is then compared with this critical value. If the test is failure-truncated, the calculated value is

$$C_M^2 = \frac{1}{12M} + \sum_{i=1}^{M} \left[\left(\frac{X_i}{X_N} \right)^{\bar{\beta}} - \frac{2i-1}{2M} \right]^2 \quad (18)$$

Table 2. Critical Values of C_M^2 for Parametric Form of the Cramer–von Mises Statistic at Level of Significance α

	Critical Value				
M[a]	α = 0.20	0.15	0.10	0.05	0.01
2	0.138	0.149	0.162	0.175	0.186
3	0.121	0.135	0.154	0.184	0.231
4	0.121	0.136	0.155	0.191	0.279
5	0.121	0.137	0.160	0.199	0.295
6	0.123	0.139	0.162	0.204	0.307
7	0.124	0.140	0.165	0.208	0.316
8	0.124	0.141	0.165	0.210	0.319
9	0.125	0.142	0.167	0.212	0.323
10	0.125	0.142	0.167	0.212	0.324
15	0.126	0.144	0.169	0.215	0.327
20	0.128	0.146	0.172	0.217	0.333
30	0.128	0.146	0.172	0.218	0.333
60	0.128	0.147	0.173	0.221	0.333
100	0.129	0.147	0.173	0.221	0.336

[a]For $M > 100$, use values for $M = 100$.

If the test is time-truncated, then the calculated value is

$$C_M^2 = \frac{1}{12N} + \sum_{i=1}^{N}\left[\left(\frac{X_i}{t_0}\right)^{\bar{\beta}} - \frac{2i-1}{2N}\right]^2 \tag{19}$$

If the calculated value is greater than the tabulated critical value, then the AMSAA model is rejected. If the AMSAA model is not rejected, then the instantaneous MTBF can be determined with $m(t) = (\lambda\beta t^{\beta-1})^{-1}$. Confidence tables (4) developed for the AMSAA model can then be used to determine both the lower and upper confidence bounds around this MTBF.

In some instances, especially for warranty returns, a portion of the data may be missing. Crow developed a technique to estimate reliability growth when data may be missing or some failure times are not known (15). If the intervals of time are known for the data, then the failures can be grouped within each interval. Crow emphasized that this method should be used only in the special case when data were missing. For grouped data, the estimation procedure is somewhat more complicated, since a closed-form equation for β does not exist. Assume that there are k intervals with boundaries $k = 0, x_1, \ldots, x_k$; the maximum likelihood of the shape parameter β is the value that satisfies the equation

$$\sum_{i=1}^{n} n_i \frac{x_i^{\hat{\beta}} \ln x_i - x_{i-1}^{\hat{\beta}} \ln x_{i-1}}{x_i^{\hat{\beta}} - x_{i-1}^{\hat{\beta}}} - \ln x_k = 0 \tag{20}$$

where $x_0 \ln x_0$ is defined to equal zero. Numerical techniques must be employed to solve this equation for β. The scale parameter λ can be estimated by

$$\hat{\lambda} = \frac{\sum_{i=1}^{K} N_i}{x_k^{\hat{\beta}}} \tag{21}$$

Discrete and Software Growth Models

In addition to the Duane and AMSAA reliability models, which were examples of continuous models, discrete models have also been developed. Discrete models differ from continuous models because they measure reliability for one-shot systems, such as a missile or rocket. These products either fail or operate when called into service. Common discrete models include models developed by Lloyd and Lipow (16) and Wolman (17). Other reliability growth models have been developed that model software growth. When software defects (bugs) are removed and corrected during testing and debugging, the number of faults residing in the code is reduced. Musa (18), Jelinski and Moranda (19), and Littlewood and Verrall (20) have developed models that estimate the number of faults in code. For products that include both hardware and software a combination of the models should be used.

EFFECTIVENESS

Ideally, all failures will be corrected and the corrections will eliminate the cause of the failure and will not introduce new failures. Unfortunately, some failures will not be corrected, and corrective actions are not always 100% effective. In addition, there are uncertainties with the projections provided by the growth models.

When failures occur during a reliability growth test, they may or may not be corrected. Crow classified failure modes as either type A or type B. Failure modes with no correction action are termed type A; failure modes that will be corrected are termed type B. Possible causes for not implementing a corrective action (type A failure modes) include unverified failure, intermittency, failure that cannot be duplicated (caused by false BIT), an isolated incident (first-time occurrence—no trend established), funding limitations, or the fact that the design is state of the art. Experience has shown (21) that of the type-B failure modes, an average of 30% will remain in the product, even though they were thought to have been corrected. The proportion to adjust the number of type-B failure modes that will be eliminated (typically 70%) is the growth effectiveness factor, EF. With this factor, the potential growth (9) upon the completion of the growth process can be determined by

$$\text{System}_{\text{GP}} = \frac{1}{\lambda_{\text{A}} + [(1-\text{EF})x\lambda_{\text{B}}]} \tag{22}$$

where $\text{System}_{\text{GP}}$ is the product growth potential, λ_{A} is the observed failure rate of type-A failure modes, and λ_{B} is the observed failure rate of type-B failure modes.

The one variable that has the greatest effect on the growth models is the time of failure. Especially for the AMSAA model, since it is a learning curve approach, the time of the first failure has a significant effect on the calculated MTBF. The uncertainty of growth estimates was discussed in a study (5) conducted by the Department of Defense (DoD). The study conducted Monte Carlo simulations using the AMSAA model to determine the probable uncertainties for the MTBF and growth rates. The study concluded that in an 80% confidence band, if 30 failures were corrected, then the true MTBF value could range from a factor of 0.7 to 1.4 times the estimated value. If only 5 failures were corrected, then the factor range would be 0.4 to 2.6. This does not imply that the AMSAA should not be used. It does imply that more emphasis should be placed on continuously improving and less emphasis should be placed on scoring or calculating an MTBF.

A reliability growth process is a cost-effective method of continuously improving the reliability of a product. Sound engineering judgment should be used to incorporate as many corrective actions as possible and compare the results of the growth process with the predicted (estimated) or calculated (from field data) reliability. Overall, to minimize the uncertainty and maximize the effectiveness of incorporating corrective actions, the more successful growth processes will collect accurate time-to-failure data, system operating performance data, and actual and/or tested operating environmental data at time of failure.

BIBLIOGRAPHY

1. ARINC, *The ARINC Story,* ARINC Incorporated, 1987.
2. J. T. Duane, Learning curve approach to reliability monitoring, *IEEE Trans. Aerospace,* **2**: 563–566, 1964.
3. L. H. Crow, On tracking reliability growth, *1975 Proc. Annu. Reliability and Maintainability Symp.,* Washington, DC, 1975, pp. 438–443.

4. US Army Communications Research and Development Command, *Reliability Growth Management,* MIL-HDBK-189, February 13, 1981.
5. US Air Force, Army, and Navy, *The TAAF Process,* HQ AMC/QA, OASN S&L, HQ USAF/LE-RD, 1989.
6. W. E. Deming, *Out of the Crisis,* Massachusetts Institute of Technology Center for Advanced Engineering Study, 1992.
7. E. R. Carrubba, Integrating life-cycle cost and cost-of-ownership in the commercial sector, *1992 Proc. Annu. Reliability and Maintainability Symp.,* Las Vegas, NV, 1992, pp. 101–107.
8. Reliability Analysis Center, *Reliability Toolkit: Commercial Practices Edition,* 1995, p. 426.
9. J. R. Devaney, G. L. Hill, and R. G. Seippel, *Failure Analysis, Mechanisms, Techniques, & Photo Atlas,* Monrovia, CA: Failure Recognition and Training Services, Inc., 1983.
10. M. Pecht (ed.), *Product Reliability, Maintainability, and Supportability Handbook,* Boca Raton, FL: CRC Press, 1995.
11. W. R. Simpson and J. W. Sheppard, *System Test and Diagnosis,* Boston, MA: Kluwer, 1994.
12. D. T. O'Connor, *Practical Reliability Engineering,* 3rd ed., Chichester: Wiley, 1991, pp. 292–300.
13. L. H. Crow, On the initial system reliability, *1986 Proc. Annu. Reliability and Maintainability Symp.,* Las Vegas, NV, 1986, pp. 115–119.
14. D. R. Cox and P. A. W. Lewis, *The Statistical Analysis of Series of Events,* New York: Wiley, 1966.
15. L. H. Crow, Reliability growth estimation with missing data—II, *Proc. Annu. Reliability and Maintainability Symp.,* Los Angeles, CA, 1988, pp. 248–253.
16. D. K. Lloyd and M. Lipow, *Reliability: Management Methods and Mathematics,* Englewood Cliffs, NJ: Prentice-Hall, 1962.
17. W. Wolman, Problems in system reliability analysis, in M. Zelen (ed.), *Statistical Theory in Reliability,* Madison, WI: University of Wisconsin Press, 1963.
18. J. D. Musa, A. Iannino, and K. Okumoto, *Software Reliability Measurement, Prediction and Application,* New York: McGraw-Hill, 1987.
19. Z. Jelinski and P. B. Moranda, Software reliability research, in W. Freiberger (ed.), *Statistical Computer Performance Evaluation,* New York: Academic Press, 1972.
20. B. Littlewood and J. L. Verrall, A Bayesian reliability growth model for computer software, *Record IEEE Symp. Computer Software Reliability,* New York, 1973.
21. L. H. Crow, Reliability values in the development process, *1985 Proc. Annual Reliability and Maintainability Symp.,* Philadelphia, 1985, pp. 383–388.

WALTER TOMCZYKOWSKI
ARINC