PER: A Process Control Methodology for Development Testing

Larry H. Crow, Ph.D., Crow Reliability Resources

Key Words: Reliability management strategy metrics, Process Control, Continuous Evaluation, Control Charts

SUMMARY & CONCLUSIONS

It is not unusual for a system to fail to meet its reliability requirement during development reliability growth testing due to problems that management could have corrected if the issues had been brought forward in a timely manner. In order for management to act it must have simple, but convincing, information on key parameters during testing that define success. If the criteria for these parameters are not met at any point during testing then the system may not be on track to meet requirements. This indicates an alert to management that needs to be investigated and appropriate changes made. Within the Department of Defense the concept of Continuous Evaluation is defined as that process which provides the continuous flow of information regarding system status to include planning, testing, data compilation, analysis, evaluation, conclusions, and reporting. The objective of this paper is to present a simple continuous evaluation control chart methodology that effectively assesses the data during development testing and evaluates if the management strategy and design are sufficient to support meeting the target reliability. The methodology utilizes a 'dial up' feature where the target requirement is an input and the methodology determines key parameters and control chart limits that are necessary to meet the requirement. If the parameters fall outside the limits then a caution is raised which allows management to investigate the situation and take early corrective action if necessary.

1. INTRODUCTION

A system may fail to meet its reliability requirement for a number of reasons. They are all important but there are two that occur frequently and should be addressed as early as possible in a development testing program. One is that the inherent capability of the design is insufficient to reach the requirement no matter how much corrective action takes place or how much testing is conducted. In terms discussed later in this paper this means that the growth potential in the design is inadequate. If this is the case it usually means that at least part or all of the system will need to be redesigned for reliability. A redesign for reliability means that the same function is achieved but in a different way so that entire failure modes are eliminated or failure rates greatly reduced.

A second problem occurs when the system's engineering design is adequate but not enough of the system's failure intensity is being addressed by management for corrective action. In this case the reliability cannot grow to the target reliability or the growth potential. In order to reach the target reliability the management strategy will need to change. This situation should be flagged early so that valuable test resources and schedule are not greatly impacted.

In this paper we discuss a continuous evaluation approach for management to track two key reliability metrics addressing the problems noted above. This tracking methodology is in terms of two control charts. If these metrics fall outside certain control limits based on our target reliability objective then a caution is raised indicating that further management attention may be warranted.

Notation

λ	Scale parameter for Crow (AMSAA) model
β	Shape parameter for Crow (AMSAA) model
t	Test time
T	Total test time
$r(\cdot)$	Crow (AMSAA) model failure intensity
λ_I	Initial system failure intensity
λ_{GP}	Growth potential failure intensity
P	Percent failure intensity due to Type A modes
K	Percent failure intensity due to Type BD modes

2. BACKGROUND

To lay the groundwork for the continuous evaluation control chart methodology we will first review two reliability growth models. The first model is the Crow (AMSAA) model applied to test-fix-test data and the second model is the Extended Model applied to test-find-test data. For test-fix-test corrective actions are incorporated during the test. For test-find-test all corrective actions are delayed and incorporated at the end of the test. The reader is referred to Refs. 2,3,4,6, for addition information and examples on the models discussed below.

2.1. CROW (AMSAAA) Basic Model for Test- Fix- Test

The Duane postulate, Ref. 1, for reliability growth during test-fix-test development testing stated that the instantaneous system failure rate at cumulative test time t is $r(t) = \lambda\beta t^{\beta-1}$, where $0 < \lambda$ and $0 < \beta$ are parameters.

Crow (Ref. 2) modeled the Duane postulate stochastically as a non-homogeneous Poisson process

(NHPP) with intensity function

$$r(t) = \lambda \beta t^{\beta - 1}, \qquad (1)$$

thus allowing for statistical procedures based on this process for reliability growth analyses. This model is applicable to test-fix-test data. The parameter λ is referred to as the scale parameter and β is the shape parameter. For $\beta = 1$, there is no reliability growth. For $\beta < 1$, there is positive reliability growth. That is, the system reliability is improving due to corrective actions. For $\beta > 1$, there is negative reliability growth. Under the Crow (AMSAA) basic model, the achieved or demonstrated failure intensity at time T, the end of the test, is given by r(T).

Suppose a development testing program begins at time 0 and is conducted until time T and stopped. Let N be the total number of failures recorded and let $0 < X_1 < X_2 < \ldots X_N \leq T$ denote the N successive failure times on a cumulative time scale. We assume that the Crow (AMSAA) model assumptions apply to this set of data. Under the model the maximum likelihood estimates (MLEs) for λ and β are:

$$\hat{\lambda} = \frac{N}{T^{\hat{\beta}}}, \qquad \hat{\beta} = \frac{N}{\sum_{i=1}^{N} \ln\left(\frac{T}{X_i}\right)} \, . \qquad (2)$$

It is important to note that the Crow (AMSAA) test-fix- test model does not assume that all failures in the data set received a corrective action. Based on the management strategy some failures may receive a corrective action and some may not. This topic is discussed next.

2.2 Crow Extended Model for Test- Find -Test

Suppose a system is tested for time T. During the testing, problem failure modes are identified, but all corrected actions are delayed and incorporated at the end of the test phase. This is test-find-test. These delayed corrective actions are usually incorporated as a group and the result is generally a distinct jump in the system reliability. The Extended Model (see Refs. 3, 4) estimates this jump in reliability due to the delayed fixes. This is called a "projection."

The projection model places all failures into two groups, A and BD. Type A failure modes are all modes such that if seen during test no corrective action will be taken. This accounts for all modes for which management determines that it is not cost-effective to increase the reliability by a design change. Type BD failure modes are all modes such that if seen during test a delayed corrective action will be taken. This Type A and Type BD designation plus the effectiveness of the fixes define the reliability growth management strategy. The basic projection model assumes that the group of Type A failure modes has a constant failure intensity λ_A, the i-th individual Type BD failure mode follows the exponential distribution with failure rate λ_i, and the group of Type BD failure modes has a constant failure intensity λ_{BD}.

Example 2. Test-Find-Test

For the data in Table 1, the system is tested for T=400 hours with a total of N=42 failures. Each failure is designated as either a Type A failure mode (no corrective action) or Type BD failure mode (corrective action). There are $N_A = 10$ Type A failures and $N_{BD} = 32$ Type BD failures during the test.

In Table 1 there are M = 16 unique Type BD failure modes, which means there are 16 distinct corrective actions incorporated into the system at the end of test. The total number of failures for the j-th observed distinct Type BD mode is denoted by N_j and the total number of Type BD failures seen during the test is $N_B = \sum_{j=1}^{M} N_j$.

j	X	M	j	X	M
1	1	B	2	2	B
2	2	B	2	2	B
3	4	B	2	2	A
4	5	B	2	2	B
5	5	B	2	2	B
6	6	A	2	3	B
7	7	B	2	3	B
8	9	B	2	3	A
9	1	B	3	3	A
1	1	A	3	3	B
1	1	B	3	3	B
1	1	B	3	3	B
1	1	B	3	3	B
1	1	B	3	3	B
1	1	B	3	3	A
1	1	B	3	3	B
1	1	B	3	3	B
1	2	A	3	3	B
1	2	A	4	3	B
2	2	B	4	3	A
2	2	A	4	3	B

Table 1: Test-Find-Test Data and Modes

An effectiveness factor (EF), d_j, Ref. 3, is the fraction decrease in λ_j after a corrective action has been made for the j-th Type BD mode. The failure rate of the j-th Type BD failure mode after a corrective action is $(1-d_j)\lambda_j$. In practice the Extended Model EFs are assigned based on engineering assessments, test results, and other factors. Studies indicate that an average EF d of .70 is typical for a reliability growth program. Individual EFs may vary. See Ref. 5.

3. CONTINUOUS EVALUATION MANAGEMENT OBJECTIVE

In the test-find-test example we test the system for 400 hours. Under test-find-test we calculate the current achieved MTBF as the cumulative $MTBF = \frac{400}{42} = 9.5$. (See Ref. 4). From this 9.5 MTBF the Extended Model with effectiveness factors given in Ref. 4 projects a MTBF of 15.1 hours if all of the 16 unique BD failure modes received a corrective action. If we keep testing new Type BD unique failure modes will be seen and the projection at a later time would usually be larger than 15.1. In this projection analysis the likelihood of reaching a reliability requirement is not directly addressed continuously during the 400 hour test. What would be useful is a methodology that addresses the requirement and continuously gives feedback during testing on critical parameters that affect attaining the requirement.

The objective of the methodology presented in this paper is to provide an early indication of potential problems in regard to meeting the stated reliability requirement or target. This indication is a caution and flags a situation where there may be a reasonable risk associated with the design or management strategy relative to attaining the reliability target. This methodology is designed to help give management additional lead time to make design or management strategy changes instead of finding out late in testing that there are issues that need management attention. This methodology is the focus of this paper. To develop these methods we need additional concepts and definitions.

4. GROWTH POTENTIAL

The growth potential is the maximum reliability the system can attained with the system design and the management strategy. The growth potential is reached when all Type BD failure modes have been seen and corrective actions implemented. This happens when the rate in which we are discovering new Type BD failure modes is equal to zero The estimated growth potential failure intensity is

$$\hat{\lambda}_{GP} = \frac{N_A}{T} + \sum_{j=1}^{M} (1 - d_j) \frac{N_j}{T} \quad (3)$$

The estimated growth potential MTBF is

$$\hat{M}_{GP} = \frac{1}{\hat{\lambda}_{GP}} . \quad (4)$$

In the test-find-test example, see Ref. 4,Metric 29, the growth potential MTBF is estimated to be 22.4.

5. CONTINUOUS EVALUATION KEY EQUATION 1

When we are planning a reliability growth program we do not know the actual BD failure modes for the growth potential calculation or the associated effectiveness factors. What we do know is that the typical average effectiveness factor is generally in the area of .7. For planning purposes we can choose an expected average depending on the program, previous history, funding, etc, but the value would generally not be expected to be significantly higher than .7 without considerable justification.

How can we use the growth potential to help manage a reliability growth testing program? The growth potential MTBF must be above the target reliability we wish to attain. The higher the growth potential is above the target reliability the quicker during the testing we will reach the target reliability. Therefore, two conditions must be satisfied. First, the basic inherent design must be sufficient to support a growth potential above the target. This means that if we correct all failure modes with reasonable effectiveness we could be at the target reliability at some point in the future. This will be not the case, for example, if the design is such that the initial test reliability is very low. In addition, even if the inherent design capability is adequate we must implement an actual reliability growth management strategy to correct problems throughout testing so that the growth potential is viable and realistic.

The methodology in this paper sets a value above the target for the growth potential. We then monitor the operating system MTBF and management strategy against this growth potential. Control limits are established that raises a caution if MTBF and management strategy conditions are not met for this growth potential.

The level for the growth potential above the target reliability can be chosen so that if we estimate we are below this level it may indicate a concern. For planning purposes we will assume an average effective factor d for the program, for example d = .7 and we choose a value for the growth potential based on the target reliability we wish to attain. This may be a percent, such as 10 % or 20%, above the target reliability, for example. The higher this percentage the quicker the target reliability can be reached. This growth potential then determines certain design and management strategy parameters that we can measure during testing and control. If these parameters fall outside certain ranges this may indicate that the growth potential is too low and the target cannot be attained.

For planning purposes we will use an average effectiveness factor d and the growth potential equation

$$\lambda_{gp} = \lambda_A + (1-d)\lambda_{BD} . \quad (5)$$

At the start of development testing the system will have an initial MTBF M_I and corresponding failure intensity λ_I. The initial failure intensity is given by

$$\lambda_I = \lambda_A + \lambda_{BD} \quad (6)$$

Now let $K = \frac{\lambda_{BD}}{\lambda_A + \lambda_{BD}}$ be the fraction of the initial failure intensity addressed by a corrective action. Also, let P = 1-K be the fraction of the initial failure intensity not being addressed by a corrective action. That is, $P = \frac{\lambda_A}{\lambda_A + \lambda_{BD}}$.In general K should be large and P should be low.

We can now determine how low P should be based on our target MTBF, our target Growth Potential and our assumed average Effectiveness Factor d. This can be done by using our Key Equation 1 which relates target Growth Potential, the Effectiveness Factor d, and the management strategy fraction K. This key equation is

$$\lambda_I = \frac{\lambda_{GP}}{(1 - dK)} . \quad (7)$$

6. CONTROL CHART LIMITS

In this section we will illustrate the application of Key Equation 1 in providing a basic process control continuous evaluation methodology to be used throughout a development testing phase. This will be given in steps using the test-find-test data and can serve as a template for general applications. We first address the minimum initial MTBF for K =1. This sets an absolute lower limit on the initial MTBF.

Step 1. Specify a target MTBF to be attained (e.g. 25 hrs)

Step 2. Specify a minimum growth potential margin above the target (e.g. 10 %).

Step 3. Calculate minimum growth potential MTBF (e.g. 27.5 hrs)

Step 4. Specify an average effectiveness factor for planning (e.g. d = .7)

Step 5. Set K = 1 in Key Equation 1.

Step 6. Solve for minimum initial MTBF ($M_I = 8.25$)

The application of this result is the following: If the cumulative MTBF during the test-find-test test phase is not greater than 8.25 it is unlikely that the inherent design can support a growth potential of 27.5 or greater. If the cumulative MTBF falls below 8.25 we run a risk of not being able to reach a MTBF of 25 hrs. during our test phase even if we correct every failure mode seen during the test (no Type A modes or K = 1). If the operational cumulative MTBF is below 8.25 during testing then a redesign of the system or subsystems should generally be considered as an option.

In the above analysis of minimum cumulative MTBF during test we assumed no Type A failure modes (K = 1). Although it is a target best practice during development to not have any Type A modes, it may happen that at any specific time during the testing this goal may not be possible because of such factors as incomplete failure diagnostics or ongoing failure analysis to determine root cause. Therefore, a maximum margin for Type A can be accounted for in defining a practical management strategy. An evaluation of Department of Defense reliability growth data for major systems shows that the average K for successful programs is about .95 or greater. This means that, on average, the fraction of the failure intensity not being addressed by corrective actions, P = 1-K is .05 or less. This percent will vary by program and requirements. However, a reasonable rule of thumb for Type A modes appears to be 5% percent or less for new technology systems. The 5% rule is used in this paper to illustrate the methodology. That is, we do not want P to be greater than .05 during the testing, i.e. ($P \le .05$. , $K \ge .95$.

Therefore, to set our limits for the continuous evaluation control chart we set K = .95 in Step 5 above. We then have

Step 5. Set K = .95 in Key Equation 1.

This gives

Step 6. Solve for minimum initial MTBF ($M_I = 9.2$).

For our stated growth potential this calculation sets the practical limits for the minimum MTBF and the parameter P. In this example, we want the cumulative MTBF to be greater than 9.2 and we want at least 95% of the system's failure intensity to be addressed by corrective actions (or no more than 5 % of the system's failure intensity not being addressed by corrective action). We can monitor these limits with control charts and act if these metrics fall outside the stated bounds. Because P = 1- K we can equivalently choose to either monitor $K \ge .95$ for Type BD modes or $P \le .05$ for Type A modes. In this example we choose to monitor P to illustrate the methods. With this methodology there are two control charts. Control Chart One is for the MTBF and Control Chart Two is for P.

7. CONTROL CHART ONE

Using Key Equation 1 and Steps 1-6 for the stated K (e.g. K = .95) we established our minimum MTBF of 9.2 for Control Chart One. For time points 100, 200, 300, 400, during the testing we plot in Figure 1 time divided by the cumulative number of failures against our control limit of 9.2. We want the cumulative MTBF to stay above the limit. This is the case in this application. This is acceptable as long as the percent Type A failure intensity remains below 5 %. This parameter is addressed next in Control Chart Two.

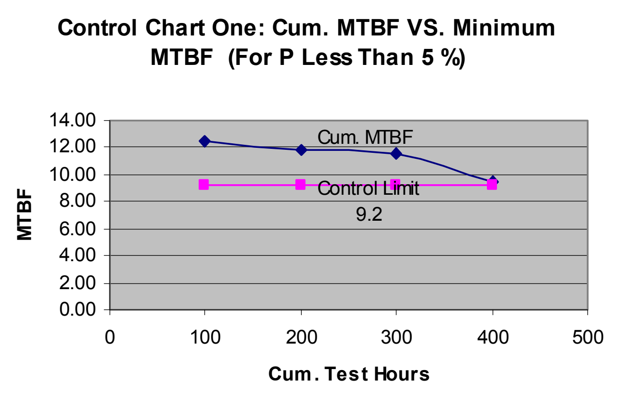

Figure 1: Control Chart One For Table 1 Data

8 CONTINUOUS EVALUATION KEY EQUATION 2

In this section we address a control chart for P. The process control metric P is the fraction of the system failure intensity that is classified as Type A. That is, that part of the failure intensity not addressed by a corrective action in accordance with out management strategy. As noted above we want this fraction to be small during reliability growth. development program.

During the testing we assume that the system failures follow the Crow (AMSAA) model discussed above. That is, we assume the system failure intensity r(t) during the testing equals $r(t) = \lambda\beta t^{\beta-1}$. In the special case of test-find-test we generally expect a constant failure intensity or $\beta = 1$. In this case the estimate of beta during testing should be consistent with beta equal to one, and this can be tested using confidence bounds. The fraction of the failure intensity r(t) classified as Type A is P.

Based on a coherent management strategy, each time we have a system failure during the test we have a constant probability P that the failure is classified as a Type A mode and a constant probability K = 1-P that the failure is classified as a Type BD failure mode. That is, the failure intensity for Type A failures is $r_A(t) = P\lambda\beta t^{\beta-1}$ and the failure intensity for Type BD failures is $r_{BD}(t) = K\lambda\beta t^{\beta-1}$.

Conditioned on the total number of system failures N(T) = n, the random variables $\ln\left(\frac{T}{X_i}\right)^{\beta}$, 1=1,…,n, unordered, are distributed as exponential random variables with mean 1, and $Y = \sum_{i=1}^{n} \ln\left(\frac{T}{X_i}\right)^{\beta}$ is a Gamma random variable with mean n. Now, a coherent management strategy will randomly place the n random variables into an A bin with probability P or a BD bin with probability K = 1-P. The $N_A(T)$ failure times X_{A_i} go into the A bin and the $N_{BD}(T)$ failure time X_{BDj} go into the BD bin. The mean of $\sum_{i=1}^{N_A(T)} \ln\left(\frac{T}{X_{A_i}}\right)^{\beta}$ is nP and the mean of $\sum_{i=1}^{N_{BD}(T)} \ln\left(\frac{T}{X_{BDj}}\right)^{\beta}$ is nK. Note that $\ln(x)^{\beta} = \beta\ln(x)$.

The estimate of P is given below by Key Equation 2

$$P_{est} = \frac{\sum_{i=1}^{N_A(T)} \ln(\frac{T}{X_{A_i}})}{\sum_{i=1}^{n} \ln(\frac{T}{X_i})} \qquad (8)$$

We can show that the mean of P_{est} is P for a coherent management strategy. That is, P_{est} is an unbiased estimate of the fraction of the system failure intensity P that is not addressed by corrective action. This is an unbiased estimate of P that utilizes both the number of Type A failures and also when they occurred during test. The estimate P_{est} is also discussed in Ref. 6

9. CONTROL CHART TWO

For a control chart on P we calculate the estimate P_{est} at time points during the testing and visually compare it to the maximum control limit, e.g. 5 %. If the estimate remains above this maximum limit then a caution is raised. In this case the situation can be investigated early to determine if changes in the program are warranted. Because K = 1-P we could choose to put a control chart either on K using $K_{est} = P_{est}$ with equal results.

We apply this methodology to our data set at times T = 100, 200, 300, 400. At these points we calculate Key Equation 2 (eq. (8)). See Figure 2.

We see in Figure 2 that the percent failure intensity curve is above the maximum control limit. This is a very strong indication that the percent of the failure intensity allocated to Type A failures is too large to reach the target reliability of 25. This situation should be evaluated and consideration given to changing previous Type A failure modes to Type BD. For example, we could convert mode A1, which currently will not receive a corrective action, to mode BD17, which will now receive a corrective action. See Table 2.

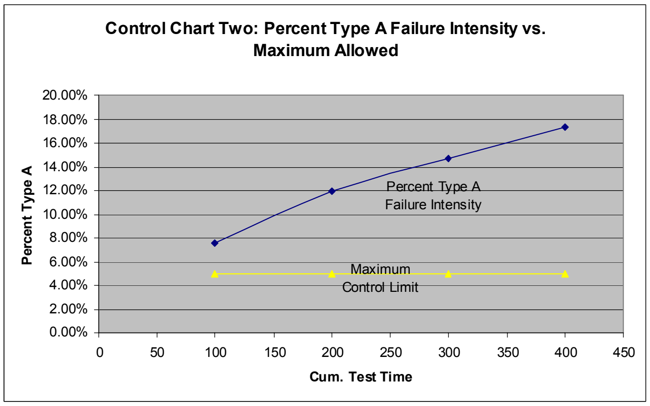

Figure 2: Control Chart Two For Table 1 Data

j	X	M	j	X	M
1	1	B	2	2	B
2	2	B	2	2	B
3	4	B	2	2	A
4	5	B	2	2	B
5	5	B	2	2	B
6	6	B	2	3	B
7	7	B	2	3	B
8	9	B	2	3	B
9	1	B	3	3	A
1	1	B	3	3	B
1	1	B	3	3	B
1	1	B	3	3	B
1	1	B	3	3	B
1	1	B	3	3	B
1	1	B	3	3	B
1	1	B	3	3	B
1	1	B	3	3	B
1	2	B	3	3	B
1	2	A	4	3	B
2	2	B	4	3	B
2	2	B	4	3	B

Table 2: Test-Find-Test Data and Modes

For this management strategy, the corresponding Control Chart Two is given below in Figure 3.

Both Control Chart One and Control Chart Two are required to jointly manage the MTBF and Type A failure intensity fraction P relative to the stated growth potential. Figure 1, together with Figure 3, indicate that the system design and management strategy are currently in control with regard to these parameters and no management action is needed. The Extended Model (Ref. 4) can be used to further monitor progress toward the requirement of 25 hrs.

The Control Chart application in this paper utilizes only point estimates and therefore reasonable consideration should be given to statistical variability in the data. However, the objective should be to keep the MTBF control line above the limit and the P control line below the limit. Also, if a system has a low MTBF then a strong management policy may be to allow very few Types A modes.

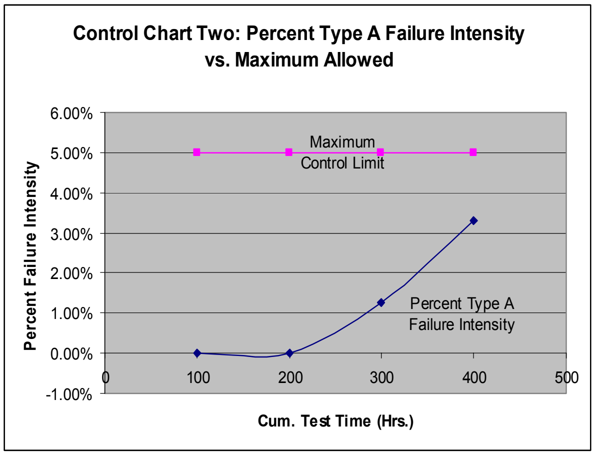

Figure 3: Control Chart Two For Table 2 Data

REFERENCES

1. J. T. Duane, Learning Curve Approach to Reliability Monitoring, IEEE Transactions on Aerospace. Vol. 2, 1964, PP.563-566.
2. L. H. Crow, Reliability Analysis for Complex, Repairable Systems, in Reliability and Biometry, ed. by F. Proschan and R.J. Serfing, pp.379-410, 1974, Philadelphia, SIAM.
3. L. H. Crow, Reliability Growth Projection from Delayed Fixes, Proceedings 1983 Annual Reliability and Maintainability Symposium, pp.84-89. and Michael Cohen, eds. 2002. National Academy Press.
4. L. H. Crow, An Extended Reliability Growth Model for Managing and Assessing Corrective Actions. Proceedings of the 2004 Annual RAMS Symposium, January 2004, Los Angeles, CA.
5. L. H. Crow, Methods and Studies to Improve the Effectiveness of Reliability Tasks. Proceedings of the 2005 Annual RAMS Symposium, January 2005, Alexandra, VA.
6. L. H. Crow, Useful Metrics for Managing Failure Mode Corrective Action. Proceedings of the 2006 Annual RAMS Symposium, January 2006, Newport Beach, CA.

BIOGRAPHY

Larry H. Crow, Ph. D.
109 Clifts Cove Blvd.
Madison, AL 35758 U. S. A.

e-mail Crowrel@knology.net

Dr. Larry H. Crow is president of CRR. Previously Dr. Crow was VP, Reliability & Sustainment Programs, at ALION Science and Technology, Huntsville, AL. From 1985 to 2000 Dr. Crow was Director, Reliability, at General Dynamics ATS- formally Bell Labs ATS. From 1971-1985, Dr. Crow was chief of the Reliability Methodology Office at the US Army Materiel Systems Analysis Activity (AMSAA). He developed the Crow (AMSAA) reliability growth model, which has been incorporated into US DoD handbooks, and national & international standards. He chaired the committee to develop Mil-Hdbk-189, *Reliability Growth Management* and is the principal author of that document. Dr. Crow is a Fellow of the American Statistical Association, and the Institute of Environmental Sciences and Technology. He is a Florida State University Alumni Association Distinguished Alumnus and the recipient of the FSU "Grad Made Good" Award for the Year 2000, the highest honor given to a graduate by Florida State University.