A Methodology for Managing Reliability Growth during Operational Mission Profile Testing

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

Key Word: Operational Testing, Reliability Management, Reliability Growth, Mission Profiles

SUMMARY & CONCLUSIONS

It is common practice for systems to be subjective to operational testing during their development program. The objective of this testing is to evaluate the performance, including reliability, of the system under conditions that represent actual use conditions. Because of expense, resources, schedule, and other considerations, these operational tests rarely represent exactly the actual use conditions. Rather, stated mission profile conditions are specific for the operational testing. These mission profiles conditions are typically general statements that guide the testing on an average basis during the testing. Because of practical constraints the elements that make up the mission profile conditions are typically tested under varying schedules with the intent that on average the mission profile conditions are met. It is also common practice that reliability corrective actions are incorporated into the system during this type of testing. That is, the test is often an operational mission profile reliability growth test. Under these conditions, we usually have a lack of structure for managing the elements that make up the mission profiles, which makes it very difficult to have an agreed-on methodology for estimating the system's reliability. This is especially true if reliability growth is occurring. Many systems fail operational testing because key assessments parameters can not be made in a straightforward clear manner so that management can take timely and appropriate action. This paper addresses this issue and presents a methodology currently being applied on major Department of Defense programs for operational reliability growth testing.

1. INTRODUCTION

Operational testing for many systems is generally an attempt to subject the system to conditions close to the actual environment expected under customer use. Sometimes the stated intent is for a demonstration test where corrective actions are not the prime objective. However, it is not unusual for a system to have a lower than desired reliability and the test becomes a reliability growth test with corrective actions. Therefore, important valid key parameters are needed to properly assess this situation and make cost-effective and timely decisions. This is often difficult in practice because of the conduct of the mission profiles during the test.

To illustrate the practical situation addressed in this paper, we consider an example where there are three mission profile tasks. In this example, suppose a system is required to conduct a particular function a discrete number of times for each hour of system operation (Task 1), be required to move a fixed number of miles under condition E1 for each hour of system operation (Task 2), and be required to move a fixed number of miles under condition E2 for each hour of system operation (Task 3). During the practical conduct of operational testing these three guidelines are generally met individually as averages. For example, the actual as-tested profile for Task 1 may not be uniform relative to the stated mission guidelines during the testing. If a task is operated below the stated guidelines for a period of time this could mask a major reliability problem if the analysis used that data directly. This often leads to issues in how to analyze the data and evaluate reliability growth. This becomes a significant program risk because an important aspect of effective reliability risk management is not to wait until the end of the test to have an assessment of the reliability performance. As we move through the testing, the actual percent of time under each of the mission profile tasks will rarely, if ever, balance continuously to the stated average. The problem is how to evaluate the reliability growth under these conditions. A common analysis method is to piece the reliability assessment together by evaluating each element of the profile separately and weighting the respective contributions to the system reliability. In our example, the reliability of Task (1) must also be converted to hours and the total reliability converted to MTBF, the typical parameter of interest. It is not a well defined methodology and does not account for reliability growth in an easy to understand manner. It is therefore not unusual for two separate organizations, for example the customer and the developer, to analyze the same data and have different MTBF numbers. Consequently, there is a need for a rigorous methodology for reliability management during operational reliability growth testing that does not reply on piecewise analysis and avoids the issues noted above.

This paper presents a methodology currently being used by the Department of Defense to manage system reliability during operational mission profile testing. This methodology will work equally well for any commercial product, for example, trucks, printers, etc. The methodology approach draws information from particular plots of the operational test data and inserts key information into a growth model. This improved methodology does not piece the analysis together. but gives a direct MTBF mission profile estimate of the system's reliability that is directly compared to the MTBF requirement. No conversion for a discrete task to a continuous scale such as hours is needed. A key advantage is that the methodology is well defined and all organizations involved will arrive at the same reliability assessment with the same data. These features of the methodology allows for a structured approach for reliability management decisions and data analysis during mission profile operational testing. This paper will address the background that is the basis of the methodology, discuss the procedures and related models, and illustrate real-world applications by numerical example.

Notation

λ	Scale parameter for Crow Extended model
β	Shape parameter for Crow Extended model
α	Growth rate equal to 1- β
t	Test time
T	Total test time
MTBF	Mean time between failures
DMTBF	Demonstrated mean time between failures
Xj	The j-th successive failure time
N	Total number of failures

2. BACKGROUND

To lay the groundwork for this operational test evaluation methodology, we will first review some reliability growth model concepts used in this paper.

2.1 Crow Extended Model Background

Suppose a system is tested for time T. During the testing, problem failure modes are identified. Some corrected actions may be incorporated during the test and other corrective actions may be delayed and incorporated at the end of the test phase. This is test-fix-find-test. We will often incorporate these delayed corrective actions as a group at a time T and the result is generally a distinct jump in the system reliability. The Crow Extended Model (see Refs. 2, 3) estimates this jump in reliability due to the delayed fixes. This is called a “projection.” The “as tested” reliability immediately before the jump is the demonstrated or assessed reliability. This is also estimated by the Extended model utilizing the Crow (AMSAA) model.

The Crow Extended model places all failures into three groups: A, BC, 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 BC are those failure modes such if seen will be corrected during the test. These are usually easy to correct problems which are corrected at the time of failure. Type BD failure modes are all modes such that if seen during test, a delayed corrective action will be taken. Type BD failure mode corrective actions are generally more extensive than Type BC corrective actions and have a greater impact on the reliability improvement. This Type A, Type BC, and Type BD designation plus the effectiveness of the fixes define the reliability growth management strategy. See data in Table 1.

Again, under general growth testing, we would pick a time T, for example T = 4000 in Table 1, stop the test, and incorporate the delayed corrective actions. The time T is both a point in which we incorporate corrective actions and also an assessment point to estimate the demonstrated and projected reliability (typically an MTBF).

Table 1: Failure Times and Failure Mode Designations, T= 4000 Hrs.

J	X_J	Mode	J	X_J	Mode
1	7	BC1	19	1555	BD10
2	10.5	BC2	20	1918	BD11
3	15	BC3	21	1926	BC8
4	17	BD1	22	1947	BC9
5	79	A	23	2450	BD12
6	102	BD2	24	2467	BD11
7	113	BD3	25	2490	BD12
8	812	BD4	26	2601	BD1
9	837	BC4	27	2635	BD8
10	866	BD5	28	2747	BD6
11	977	A	29	3040	BD9
12	992	BC5	30	3154	BD4
13	996	BD6	31	3171	A
14	1002	BD7	32	3249	BD10
15	1509	BD8	33	3502	BD3
16	1516	BC6	34	3552	BC11
17	1522	BD2	35	3867	BD12
18	1529	BD1			

To make these concepts concrete, we define (1) a stopping point as a time in which we stop the test and incorporate delayed corrective actions and (2) an assessment point is a test time at which we want to make valid reliability assessments (demonstrated or projected) based on data to date at that point. Under common applications of the Extended model, a stopping point is also an assessment point.

When the failure times are known, the Extended model demonstrated or assessed reliability for the system configuration on test at the assessment point is derived from the Crow (AMSAA) basic model. See Refs. 1,3.

21 Assessment Problems During Mission Profile Testing

If the operational mission profile tasks are not continuously being exercise exactly in accordance with their averages throughout the entire test, then we cannot use the actual successive times of failure to apply the Extended model or the Crow (AMSAA) model. This is because if a mission profile task is being exercised well above or below its expected usage over a period of time then the actual failures times are not represented of the reliability. For example, we may see clustering of failures and then long periods of no failures, depending on the profiles actually being exercised. This is indicated by the data in Table 1. In this case it would not be unusual for the model goodness-of-fit test to give a failed result if all the actual failure times were used. For example, if a task has a high failure rate but is not being executed over the first 100 hours of test, then an Extended or Crow (AMSAA) model assessment over this period would not reflect the actual impact of this task on the system reliability. The objective of this paper is to present a practical methodology that will allow for a direct application of the Extended Model and the Crow (AMSAA) model under these conditions. In this paper we will illustrate the methodology utilizing only the Crow (AMSAA) basic model and estimating the demonstrated reliability and other related reliability parameters. The Extended model projection application would use addition information such as the failure mode classifications in Table 1 and effectiveness factor assignments (see Refs. 2, 3).

In our methodology we do not use the actual failure times but rather use the grouped data method utilizing the number of failures over intervals. This is discussed next. This discussion is in the context of the Crow (AMSAA) model and establishes the basis for the methodology's grouped data analysis approach. With additional Extended model grouped data input other parameter can be estimated. The grouped data version of the Extended model is in the Reliasoft RGA 6 software and documentation and for general applications gives a demonstrated reliability estimate, a projection, and other parameters such as growth potential.

2.2 Crow (AMSAA) Grouped Data Model Background

The Crow (AMSAA) model grouped data method addresses the situation where the actual failure times may not be known or we do not desired to use them. In this case the total test period is partitioned into K intervals and the number of failures in each interval is known. It is not required that the intervals be of the same length.

Let the length of the q^{th} interval be L_q , $q = 1,\ldots,K$. Also, let $T_1 = L_1$, $T_2 = L_1 + L_2$, …, etc, be the accumulated time through the q^{th} interval. Let N_q be the total number of failures in the q^{th} interval and let $T_K = T$. See Table 2.

Table 2: Grouped Data for Test-Fix-Test

Interval	No. of Failures	Length	Accum. Time
1	N_1	L_1	T_1
2	N_2	L_2	T_2
...			
q	N_q	L_q	TS_q
K	N_K	L_K	T_K

The Crow (AMSAA) model failure intensity is estimated by

$$\hat{r}(T) = \hat{\lambda}\hat{\beta}T^{\hat{\beta}-1} \qquad 1)$$

where the values $\hat{\lambda}$ and $\hat{\beta}$ satisfy

$$\sum_{q=1}^{K} N_q \left[\frac{\left\{ \left[T_q\right]^{\hat{\beta}} Ln\left[T_q\right] - \left[T_{q-1}\right]^{\hat{\beta}} Ln\left[T_{q-1}\right] \right\}}{\left[T_q\right]^{\hat{\beta}} - \left[T_{q-1}\right]^{\hat{\beta}}} \right] = 0 \qquad (2)$$

$$\hat{\lambda} = \frac{N}{T^{\hat{\beta}}}. \qquad (3)$$

where $N = N_1 + N_2 \ldots + N_q$ is the total number of failures.

The achieved or demonstrated MTBF is estimated by

$$\hat{M}(T) = \left[\hat{r}(T)\right]^{-1}. \qquad (4)$$

3. CORRECTIVE ACTION STRATEGY

For the typical reliability growth situation under 'normal" applications, a BC mode is corrected during the test and a BD mode is corrected at the end of the test, for example 4000 hours. For mission profile testing we generalized these definitions. For mission profile testing discussed in this paper we define a BC mode as a mode corrected at the time of failure and a BD mode as a mode that is delayed and corrected at a time past the first occurrence of that mode. For most practical applications, the two definitions are often the same.

5. TESTING METHODOLOGY

In order to have valid Crow (AMSAA) model assessments, we require that the operational mission profile be conducted in a structured manner. This testing methodology involves convergence points during the testing. A convergence point is a time during the test when all the operational mission profile Tasks all meet their expected averages. In this paper we will require at least three convergence points. During the testing BC modes are corrected at the time of failure and BD modes are delayed and corrected later. The incorporation of a BD mode may occur any time during the testing and we may correct one BD mode or several at same time. These stopping points to fix the BD modes may occur throughout the testing. We wish to calculate a valid estimate of the demonstrated reliability at time T using the Crow (AMSAA) model. With this methodology, the assessment point T must also be a convergence point. The test times between the convergence points do not have to be the same length.

5.1 Example of Convergence Points

To make the concepts concrete, suppose we have a military system with Task 1 firing a gun, Task 2 moving under environment E1 and Task 3 moving under environment E2. For every hour of operation, the operational profile states that the system be in environment E1 70% of the time and in environment E2 30% of the time. In additional, for each hour of operation we must fire the gun 10 times. Now, in general, it is difficult to manage an operation test so that these operational profiles are met exactly and continuously throughout the test. Our methodology requires that these conditions be met on the average at the convergence points. In practice, this can almost always be done with proper program and test management.

As noted earlier, the assessment point of interest, time T, must also be a convergence point. The methodology for controlling the convergence at these points involves monitoring a graph for each of the tasks. In our example we would have three graphs. Each graph would have a line with the expected average as a function of hours, and the corresponding actual. When an actual test profile for a task intersects the average line, then we have a convergence point for that task. A mission profile convergence point is a point when all of the tasks converge at the same time. In our example, this means a mission profile convergence point is when all three mission profile tasks meet their averages at the same time. Again, with this methodology we require at least three mission profile convergence points, one of which is the stopping point T. In our example, the total test time is 4000 hours. We choose converge points at 1000, 2500, 3200

Table 3: Firing Task Profile Testing For First 1000 Hrs

System Hrs	Expected Firings	Actual Firings	System Hrs	Expected Firings	Actual Firings
50	5	0	550	55	80
100	10	0	600	60	80
150	15	0	650	65	80
200	20	0	700	70	80
250	25	10	750	75	80
300	30	15	800	80	90
350	35	40	850	85	95
400	40	60	900	90	100
450	45	60	890	95	100
500	50	60	1000	100	100

Table 4: E1 Task Profile Testing For First 1000 Hrs

System Hrs	Expected Hrs	Actual Hrs	System Hrs	Expected Hrs	Actual Hrs
50	35	50	550	385	350
100	70	100	600	420	400
150	105	150	650	455	450
200	140	200	700	490	500
250	175	250	750	525	550
300	210	300	800	560	550
350	245	300	850	595	550
400	280	300	900	630	600
450	315	300	950	665	650
500	350	300	1000	700	700

Table 5: E2 Task Profile Testing For First 1000 Hrs

System Hrs	Expected Hrs	Actual Hrs	System Hrs	Expected Hrs	Actual Hrs
50	15	0	550	165	200
100	30	0	600	180	200
150	45	0	650	195	200
200	60	0	700	210	200
250	75	0	750	225	200
300	90	0	800	240	250
350	105	50	850	255	300
400	120	100	900	270	300
450	135	150	890	285	300
500	150	200	1000	300	300

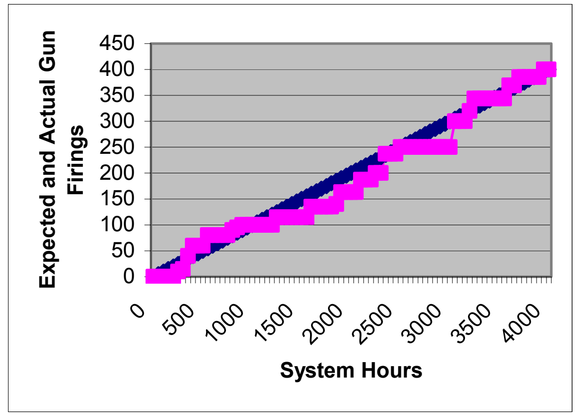

Figure 1: Firing Task Profile Testing For First 4000 Hrs

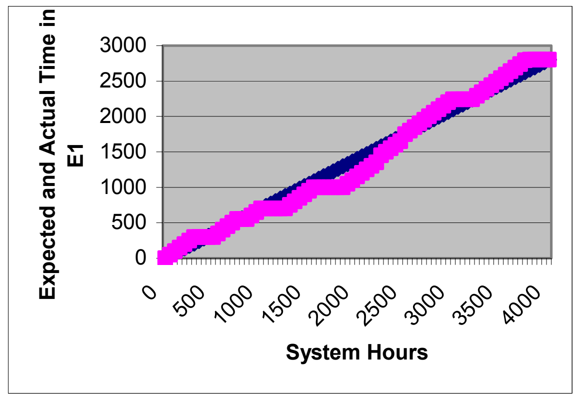

Figure 2: E1 Task Profile Testing For First 4000 Hrs

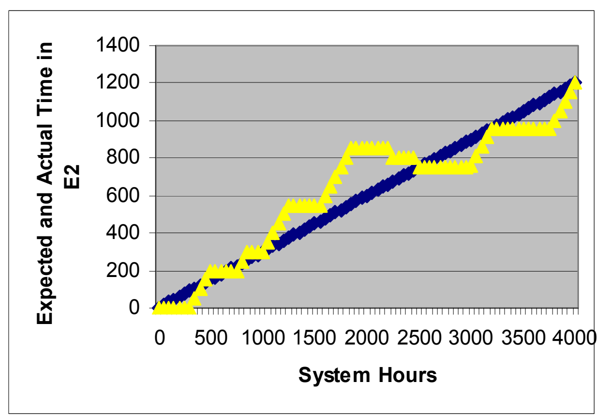

Figure 3: E2 Task Profile Testing For First 4000 Hrs

and 4000 hours. We manage the testing profiles so that at these times the actual operational test profiles equal the expected for the three profiles. For practical applications, a convergence point is when the actual profiles is exactly or is "close to" the expected profile. The profiles for the first 1000 hours of system operation are given in Tables 3 to 5 for the three tasks, firings, E1, and E2. This shows the profiles up to the first convergence point at 1000 system test hours.

For the 4000 hours of system testing we manage the testing in order to have the convergence points at times 1000, 2500, 3200 and 4000 hours. We can monitor the actual and expected task profiles. See Figures 1, 2, 3.

6. CROW (AMSAA) MODEL APPLICATION

The objective of having the convergence points is so we will be able to apply the Crow (AMSAA) model in such a way that the demonstrated reliability estimate and other key reliability growth parameters, such as growth rate, can be estimated in a valid fashion. To do this we apply the grouped data version of the Crow (AMSAA) model. For reliability growth assessments using grouped data, we only use the grouped information between interval time points in the testing. In our application, these time points are the convergence points 1000, 2500, 3200, 4000. Actual failure times are not used.

During the testing Type A, Type BC, and Type BD failure modes are seen and corrective actions are incorporated for the Type BC and BD problems. At any time during the testing, there may be BD modes which have been seen but the delayed corrective action has not yet been incorporated. At the assessment point, in our example T = 4000, we use the Crow (AMSAA) model to estimate the current reliability and the growth rate. For this basic application we do not need to input Type A, BC, and BD information as in the full Extended model. See Table 6.

Table 6: Grouped Data Input For First 4000 Hrs

Interval	No. of Failures	Length	Accum. Time
1	13	1000	1000
2	12	1500	2500
3	6	700	3200
4	4	800	4000

The parameters of the Crow (AMSAA) model are then estimated using equations (2) and (3). These parameter estimates are given in Figure 4 below in the format of Reliasoft RGA 6. Note that in this application the model goodness-of-fit test passed (at 0.1 significance level). This would not be the case if the actual failure time in Table 1 were used in the analysis. For the mission profile testing up to 4000 hours, positive reliability growth is indicated by the beta being less than one at 0.7083. The demonstrated mean time between failures (DMTBF) for the configuration on test at 4000 hours is estimated to be 161 hours. The growth rate is one minus beta or 0.2917. It is very important to note that the reliability estimate of 161 is valid for a direct comparison against the mission profile reliability requirement or goal without any adjustment or tailoring. This is the key property of this methodology and is the main result of this paper.

The mission profile reliability growth curve over 4000 hours is given in Figure 5.

This is a simple and direct methodology for evaluating a system's reliability and related growth parameters during operational mission profile testing. This methodology is practical and can be applied for any number of mission profile tasks. With this information evaluated and updated throughout the testing, management can better control risks affecting the attainment of the operational reliability objectives. This test and analysis methodology should be planned before the test starts and monitored with the graphs to assured proper implementation.

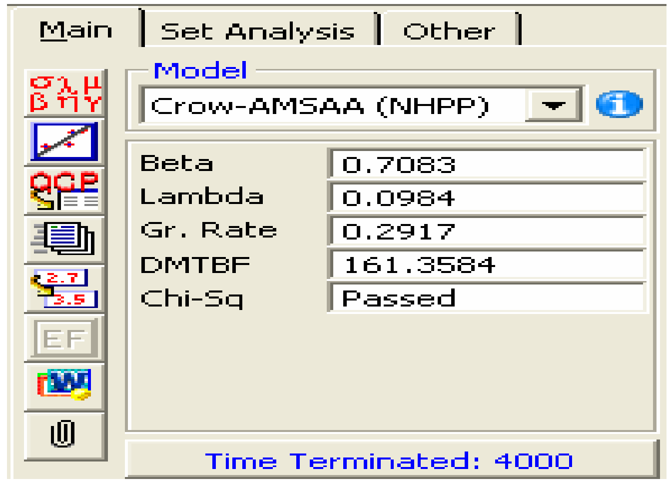

Figure 4: Profile Testing Analysis For First 4000 Hrs

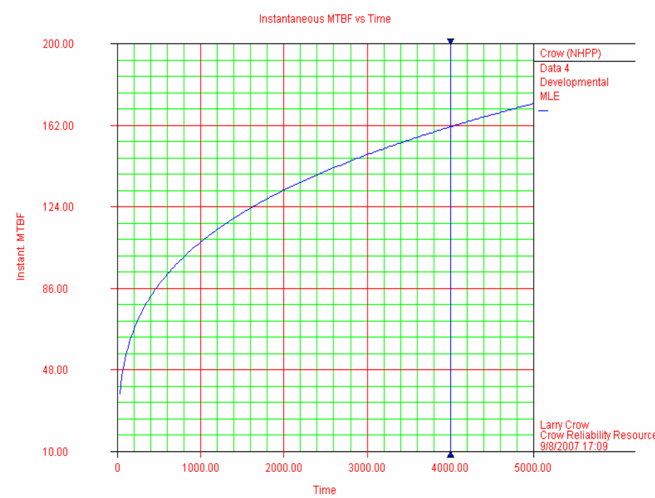

Figure 5: Profile Testing Reliability Growth Curve For First 4000 Hrs

ACKNOWLEDGMENTS

I would like to thank the referee for the excellent review and constructed comments and suggestions. I would also like to thank James Ruma, Vice President, General Dynamics Land System (GDLS), Sharad Kumar, Senior Director, GDLS, and Dmitry Tananko, Reliability Manager, GDLS, for first defining the problem addressed in this paper, and the acceptance and successful application of this methodology.

REFERENCES

1. 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.
2. L. H. , Reliability Growth Projection from Delayed Fixes, *Proceedings 1983 Annual Reliability and Maintainability Symposium*, pp.84-89.

3. An Extended Reliability Growth Model for Managing and Assessing Corrective Actions. *Proceedings of the 2004 Annual RAM Symposium*, January 2004, Los Angeles, CA.

BIOGRAPHY

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

Internet (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.