MIL-HDBK-189

13 February 1981

MILITARY HANDBOOK

RELIABILITY GROWTH MANAGEMENT

NO DELIVERABLE DATA REQUIRED BY THIS DOCUMENT

FSC RELI

DEPARTMENT OF DEFENSE
Washington, DC 20301

RELIABILITY GROWTH MANAGEMENT
MIL-HDBK-189

1. This Military Handbook is approved for use by all Departments and Agencies of the Department of Defense.

2. Beneficial comments (recommendations, additions, deletions) and any pertinent data which may be of use in improving this document should be addressed to: Headquarters, U S Army Communications Research and Development Command, ATTN: DRDCO-PT, Fort Monmouth, NJ 07703.

FOREWORD

1. The government's materiel acquisition process for new military systems requiring development is invariably complex and difficult for many reasons. Generally, these systems require new technologies and represent a challenge to the state of the art. Moreover, the requirements for reliability, maintainability and other performance parameters are usually highly demanding. Consequently, striving to meet these requirements represents a significant portion of the entire acquisition process and, as a result, the setting of priorities and the allocation and reallocation of resources such as funds, manpower and time are often formidable management tasks.

2. Reliability growth management procedures have been developed for addressing the above problem. These techniques will enable the manager to plan, evaluate and control the reliability of a system during its development stage. The reliability growth concepts and methodologies presented in this handbook have evolved over the last few years by actual applications to Army, Navy and Air Force systems. Through these applications reliability growth management technology has been developed to the point where considerable payoffs in the effective management of the attainment of system reliability can now be achieved.

3. This handbook is written for use by both the manager and the analyst. Generally, the further into the handbook one reads, the more technical and detailed the material becomes. The fundamental concepts are covered early in the handbook and the details regarding implementing these concepts are discussed primarily in the latter sections. This format, together with an objective for as much completeness as possible within each section, have resulted in some concepts being repeated or discussed in more than one place in the handbook. This should help facilitate the use of this handbook for studying certain topics without extensively referring to previous material.

CONTENTS

CONTENTS (Continued)

CONTENTS (Continued)

1. SCOPE

1.1 Purpose. This handbook provides procuring activities and development contractors with an understanding of the concepts and principles of reliability growth, advantages of managing reliability growth, and guidelines and procedures to be used in managing reliability growth. It should be noted that this handbook is not intended to serve as a reliability growth plan to be applied to a program without any tailoring. This handbook, when used in conjunction with knowledge of the system and its development program, will allow the development of a reliability growth management plan that will aid in developing a final system that meets its requirements and lowers the life cycle cost of the fielded systems.

1.2 Application. This handbook is intended for use on systems/equipments during their development phase by both contractor and government personnel.

2. REFERENCED DOCUMENTS

2.1 Issues of documents. The following documents of the issue in effect on date of invitation for bids or request for proposal, form a part of this handbook to the extent specified herein.

STANDARDS

MILITARY

MIL-STD-721 - Definitions of Effectiveness Terms for Reliability, Maintainability, Human Factors, and Safety.

MIL-STD-756 - Reliability Prediction.

MIL-STD-785 - Reliability Program for Systems and Equipment Development and Production.

MIL-STD-781 - Reliability Design Qualification and Production Acceptance Tests: Exponential Distribution.

MIL-STD-499 - Engineering Management.

(Copies of specifications, standards, drawings, and publications required by contractors in connection with specific procurement functions should be obtained from the procuring activity or as directed by the contracting officer.)

3. DEFINITIONS

3.1 Reliability growth. The positive improvement in a reliability parameter over a period of time due to changes in product design or the manufacturing process.

3.2 Reliability growth management. The systematic planning for reliability achievement as a function of time and other resources, and controlling the ongoing rate of achievement by reallocation of resources based on comparisons between planned and assessed reliability values.

3.3 Terms. The definitions of terms not called out herein shall be in accordance with MIL-STD-721.

4. GENERAL STATEMENTS

4.1 Benefits of Reliability Growth Management. The initial prototypes for a complex system with major technological advances will invariably have significant reliability and performance deficiencies that could not be foreseen in the early design stage. The prototypes, therefore, are subjected to a development testing program to surface problems so that improvements in system design can be made. The ensuing system reliability and performance characteristics will depend on the number and effectiveness of these fixes. The ultimate goal of the development test program is to meet the system reliability and performance requirements.

Experience has shown that programs which rely simply on a final demonstration by itself to determine compliance with the reliability requirements do not, in many cases, achieve the reliability objectives with the allocated resources. Emphasis on reliability performance prior to the final demonstration could substantially increase the chance of meeting these objectives. This can be accomplished by the utilization of reliability growth management. This involves setting interim reliability goals to be met during the development testing program and the necessary allocation and reallocation of resources to attain these goals. A comprehensive approach to reliability growth management throughout the development program consists of planning, evaluating and controlling the growth process.

Reliability growth planning addresses program schedules, amount of testing , resources available and the realism of the test program in achieving the requirements. The planning is qualified and reflected in the construction of a reliability growth program plan curve. This curve establishes interim reliability goals throughout the program. To achieve these goals it is important that the program manager be aware of reliability problems during the conduct of the program so that he can effect whatever changes are necessary, e.g., increased reliability emphasis. It is, therefore, essential that periodic assessments of reliability be made during the test program (e.g., at the end of a test phase) and compared to the planned reliability growth values. These assessments provide visibility of achievements and focus on deficiencies in time to affect the system design. By making appropriate decisions in regard to the timely incorporation of effective fixes into the system commensurate with attaining the milestones and requirements, management can control the growth process.

This handbook provides methodology and concepts to assist in reliability growth planning and a structured approach for reliability growth assessments. The planning aspects in this handbook address the planned growth curve and related milestones. The assessment techniques are based on demonstrated and projected values which are designed to realistically evaluate reliability in the presence of a changing configuration.

The planned growth curve and milestones are only targets. They do not imply that reliability will automatically grow to these values. On the contrary, these values will be attained only with the incorporation of an adequate number of effective design fixes into the system. This requires dedicated management attention to reliability growth. The methods in this handbook are for the purpose of assisting management in making timely and appropriate decisions to ensure sufficient support of the reliability engineering design effort throughout the development testing program.

4.2 Management's Role.

The various techniques associated with reliability growth management do not, in themselves, manage. They simply make reliability a more visible and manageable characteristic. Every level of management can take advantage of this visibility by requesting reliability growth plans and progress reports for review. Without this implementation, reliability growth cannot truly be managed.

High level management of reliability growth is necessary in order to have available all the options for difficult program decisions. For example, high level decisions in the following areas may be necessary in order to ensure that reliability goals are achieved:

- Revise the program schedule
- Increase testing
- Fund additional development effort
- Add or reallocate program resources
- Stop the program until interim reliability goals have been demonstrated

Although some of these options may result in severe program delay or significant increase in costs, they may have to be exercised when major reliability difficulties occur.

4.3 Basic Reliability Activities.

Reliability growth management is part of the system engineering process (MIL-STD 499). It does not take the place of the other basic reliability program activities (MIL-STD 785) such as predictions (MIL-STD 756), apportionment, failure mode and effect analysis, and stress analysis. Instead, reliability growth management provides a means of viewing all the reliability program activities in an integrated manner.

4.4 Reliability Growth Process.

4.4.1 Basic Process. Reliability growth is the result of an iterative design process. As the design matures, it is investigated to identify actual or potential sources of failures. Further design effort is then spent on these problem areas. The design effort can be applied to either product design or manufacturing process design. The iterative process can be simply visualized as a feedback loop as in Figure 4.1. This illustrates that there are three essential elements involved in achieving reliability growth:

(a) Detection of failure sources,

(b) Feedback of problems identified, and

(c) Redesign effort based on problems identified.

Furthermore, if failure sources are detected by testing, a fourth element is necessary:

(d) Fabrication of hardware.

And, following redesign, detection of failure sources serves as:

(e) Verification of redesign effect.

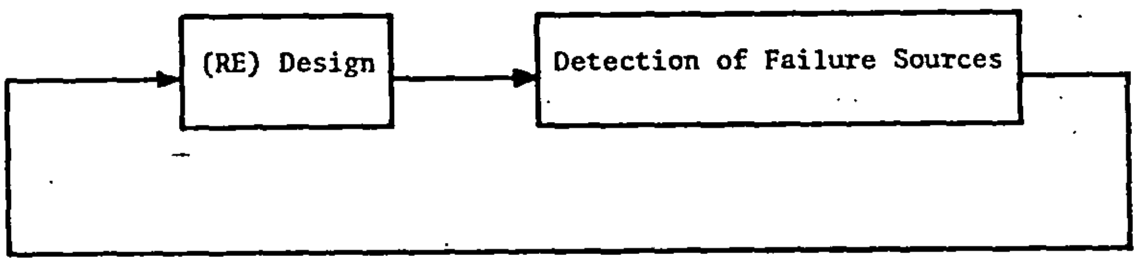


FIGURE 4.1 Reliability growth feedback model.

4.4.2 Growth Rate. The rate at which reliability grows is dependent on how rapidly activities in this loop can be accomplished, how real the identified problems are, and how well the redesign effort solves the identified problems without introducing new problems. Any of these activities may act as a bottleneck. The cause and degree of the bottleneck may vary from one development program to the next, and even within a single program may vary from one stage of development to the next.

Figures 4.1, 4.2 and 4.4 illustrate the growth process and associated management processes in a skeleton form. This type of illustration is used so that the universal features of these processes may be addressed. The representation of an actual program or program phase may be considerably more detailed. This detailing may include specific inputs to, and outputs from, the growth process, additional activity blocks, and more explicit decision logic blocks.

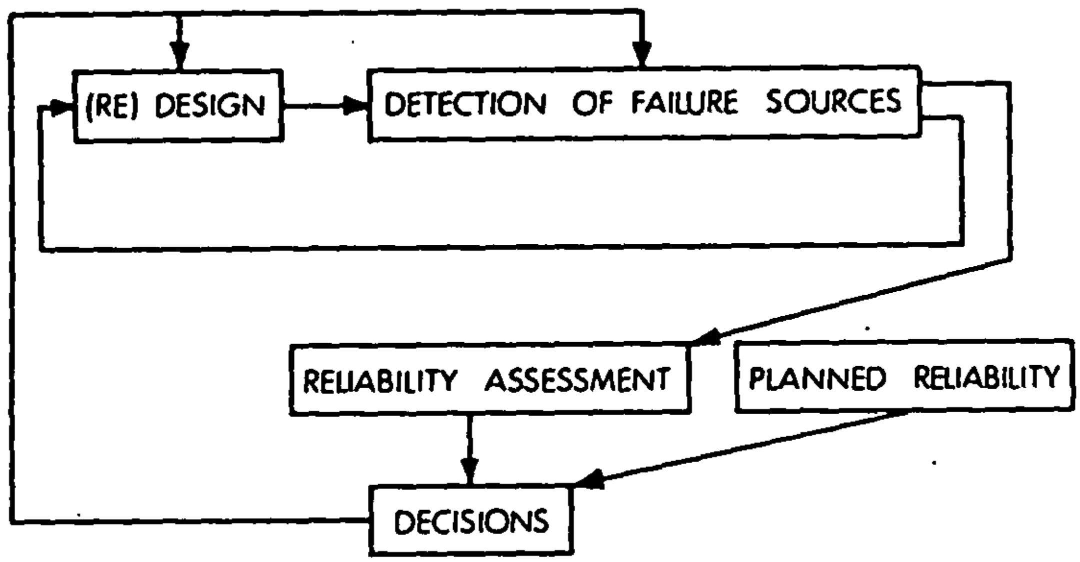


Figure 4.2 Reliability Growth Management Model (Assessment).

4.5 Reliability Growth Management Control Processes.

4.5.1 Basic Methods. There are two basic ways that the manager evaluates the reliability growth process. The first method is to utilize assessments (quantitative evaluations of the current reliability status) that are based on information from the detection of failure sources. The second method is to monitor the various activities in the process to assure himself that the activities are being accomplisehd in a timely manner and that the level of effort and quality of work are in compliance with the program plan. Each of these methods complement the other in controlling the growth process.

4.5.2 Comparison of Methods. The assessment approach is results oriented; however, the monitoring approach, which is activities oriented, is used to supplement the assessments and may have to be relied on entirely early in a program. This is often necessary because of the lack of sufficient objective information in the early program stages.

4.5.3 Assessment. Figure 4.2 illustrates how assessments may be used in controlling the growth process. Reliability growth management differs from conventional reliability program management in two major ways. First, there is a more objectively developed growth standard against which assessments are compared. Second, the assessment methods used can provide more accurate evaluations of the reliability of the present equipment configuration. A comparison between the assessment and the planned value will suggest whether the program is progressing as planned, better than planned, or not as well as planned. If the progress is falling short, new strategies should be developed. These strategies may involve the reassignment of resources to work on identified problem areas or may result in adjustment of the timeframe or a re-examination of the validity of the requirement. Figure 4.3 illustrates an example of both the planned reliability growth and assessments.

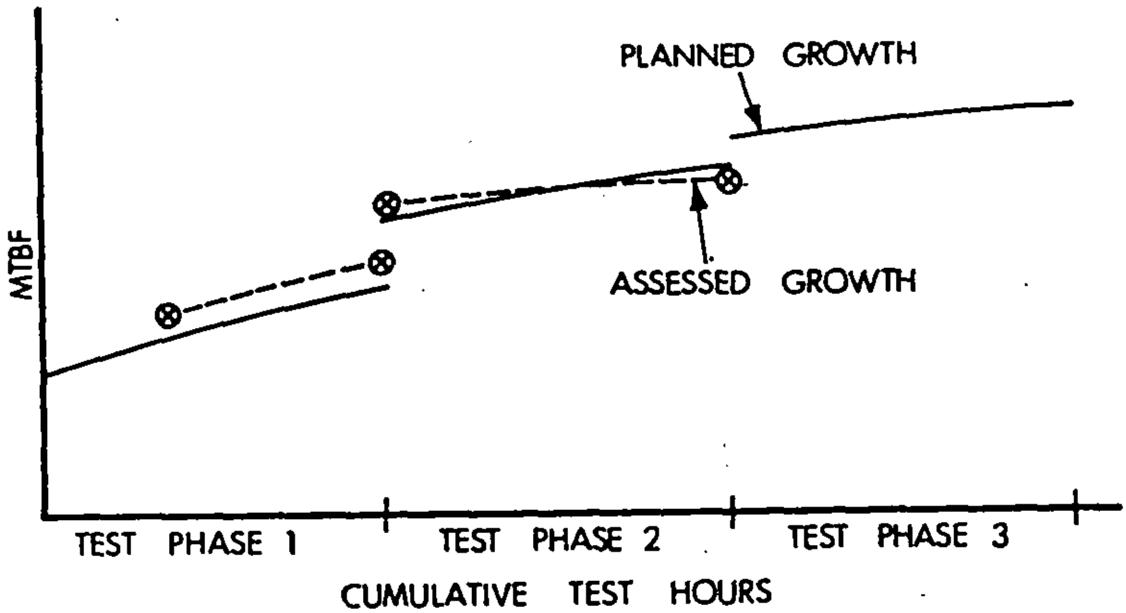


Figure 4.3 Planned Growth and Assessments.

4.5.4 Monitoring. Figure 4.4 illustrates control of the growth process by monitoring the growth activities. Since there is no simple way to evaluate the performance of the activities involved, management based on monitoring is less definitive than management based on assessments. Nevertheless, this activity is a valuable complement to reliability assessments for a comprehensive approach to reliability growth management. But standards for level of effort and quality of work accomplishment must, of necessity, rely heavily on the technical judgment of the evaluator. Monitoring is intended to assure that the activities have been performed within schedule and meet appropriate standards of engineering practice. It is not intended to second-guess the designer, e.g., redo his stress calculations. One of the better examples of a monitoring activity is the design review. The design review is a planned monitoring of a product design to assure that it will meet the performance requirements during operational use. Such reviews of the design effort serve to determine the progress being made in achieving the design objectives. Perhaps the most significant aspect of the design review is its emphasis on technical judgment, in addition to quantitative assessments of progress.

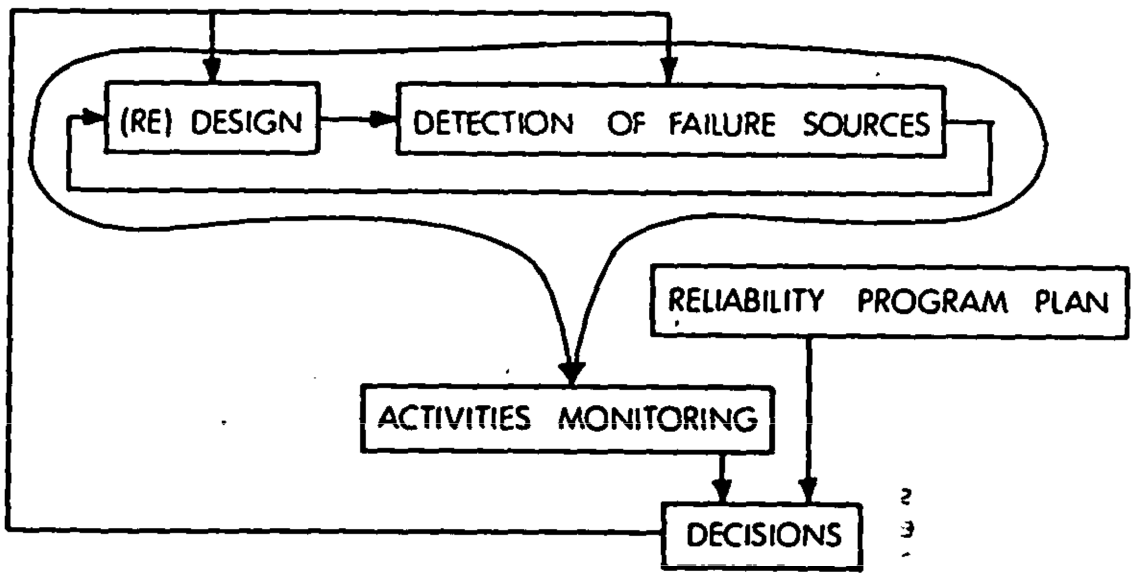


Figure 4.4 Reliability Growth Management Model (Monitoring).

4.6 Test Phase and Total Program Management Concepts.

Two approaches, or levels of consideration, for planning and controlling reliability growth are commonly used. One approach treats reliability growth on a global basis over the entire development program. The other approach treats reliability growth on a phase-by-phase or local basis. Both approaches are fundamental to the management of reliability growth. Figures 4.5 and 4.6 illustrate the appropriate approach for several growth management activities concerned with the analysis of previous programs, constructing the planned growth curve, and determining demonstrated and projected reliability values.

4.6.1 Analysis of Previous Programs. Analysis of previous similar programs are used to develop guidelines for predicting the growth during future programs. Of particular interest are the patterns of growth observed and the effect of program characteristics on initial values and growth rates. The analysis may be performed on either the overall program or individual program phases, or both.

4.6.2 The Planned Growth Curve. The planned growth curve is an important part of the reliability growth management methodology and is considered essential to any reliability program. This curve is constructed early in the development program generally before hard reliability data are obtained and is typically a joint effort between the program manager and contractor. Its primary purpose is to provide management with guidelines as to what reliability can be expected at any stage of the program and to provide a basis for evaluating the actual progress of the reliability program based upon generated reliability data. The planned growth curve is constructed on a phase-by-phase basis.

4.6.3 Demonstrated Reliability. A demonstrated reliability value is based on actual test data and is an estimate of the current level of reliability. The assessment is made on the system configuration currently undergoing test, not on an anticipated configuration. The demonstrated value is determined on a phase-by-phase basis.

4.6.4 Projected Reliability. A reliability projection is an estimate of reliability that can be expected at some future point in the development program. The projection is based on the achievement to date and future program characteristics. Projection is a particularly valuable analysis when a program is experiencing difficulties, because it enables investigation of program alternatives. Projections may be made by extrapolating a growth curve or by engineering assessments.

5. DETAILED STATEMENTS

5.1 Reliability Growth Management Concepts.

In planning a development program, methods are needed for quantifying the reliability growth and resources so that a proposed

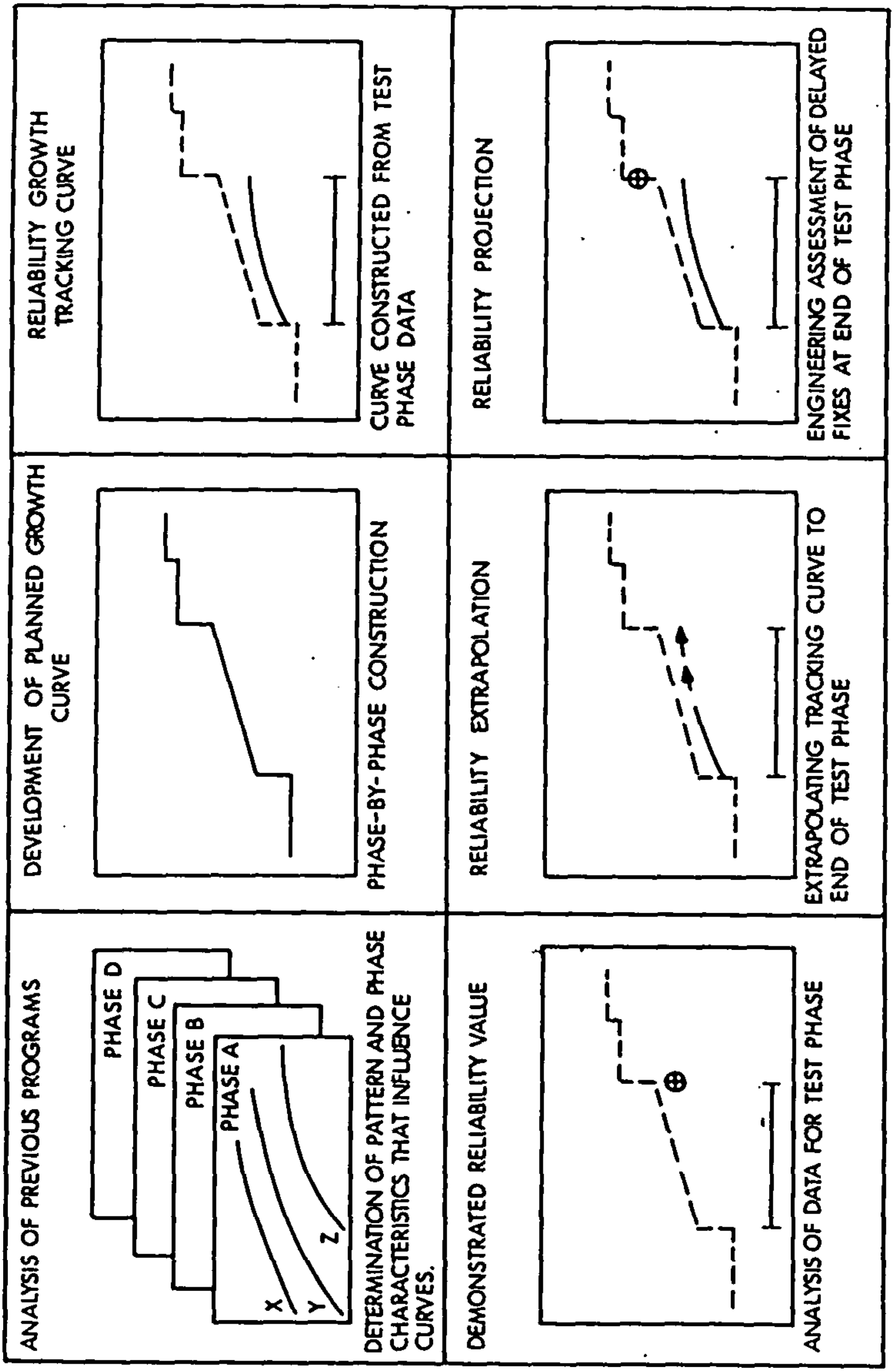


Figure 4.5 Activities that Address Reliability Growth on a Local, Phase-By-Phase Basis.

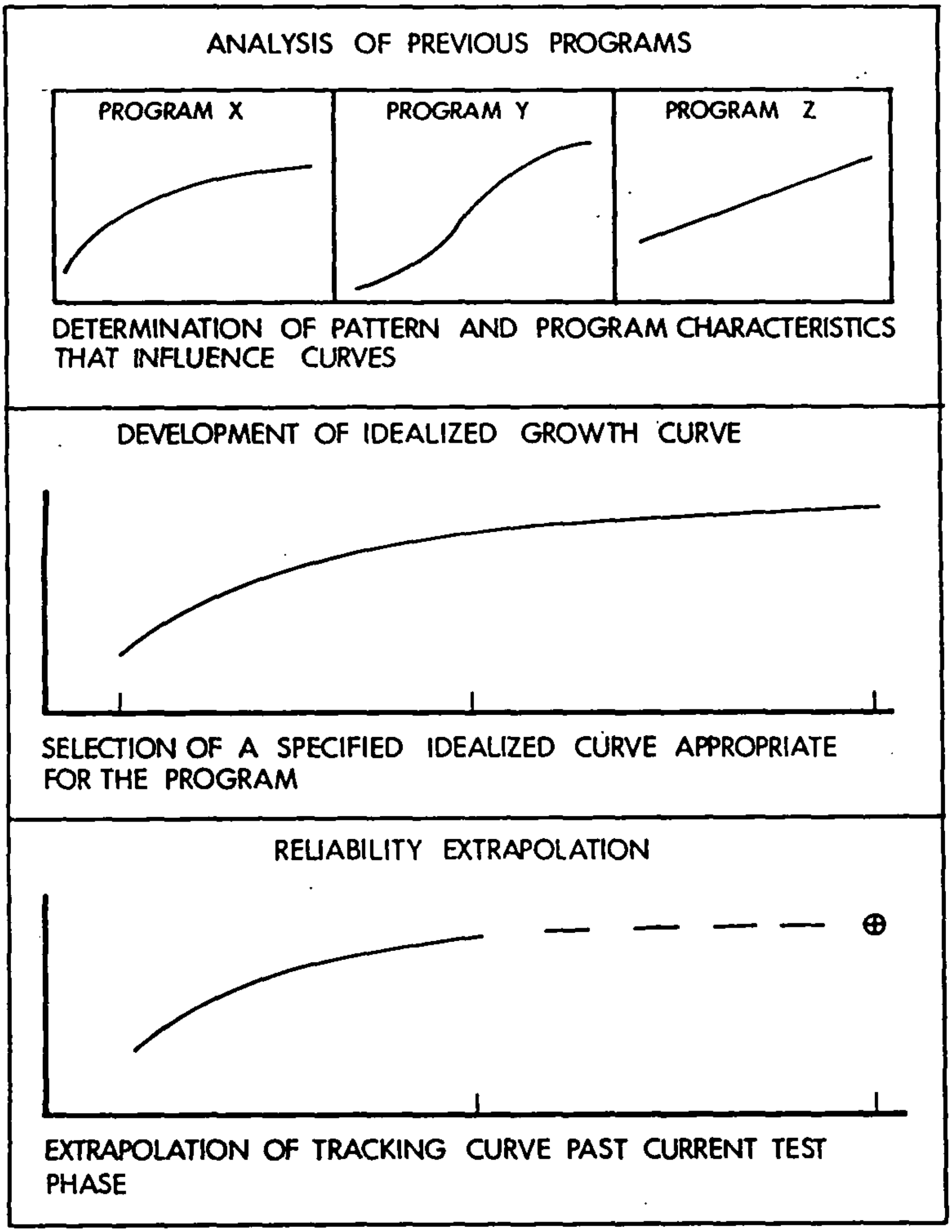


Figure 4.6 Activities That Address Reliability Growth on a Global Basis for the Entire Program.

program can be evaluated to determine if the reliability requirements and interim milestones are realistically achievable. Also, throughout development testing, management needs to evaluate progressively the reliability status of the system. By making assessments of the system reliability performance based on test data, the program manager has a tool for evaluating the adequacy of the development effort and for assessing the likelihood of attaining future reliability requirements and goals.

This section will present a formal reliability growth management format which is designed to:

a. Help the program manager plan the development program so as to best utilize available resources and set milestones.

b. Help the program manager obtain realistic estimates of current and future system reliability throughout the program.

c. Provide the program manager with a standard procedure for objectively evaluating the reliability status of the program.

Three types of reliability growth curves will be considered for the management and control format. These are idealized, planned, and tracking. This section discusses these three growth curve concepts and their respective use for planning and controlling reliability. The application of these concepts is generally not routine and would, as for most procedures, need to be tailored for the particular situation under consideration.

5.1.1 Development Program. In a general sense, the development program for a complex system is usually constructed in the manner illustrated in Figure 5.1.

5.1.2 Major Test Phases. The test portion of the program can also be divided into major phases or segments of test time. A major test phase is a distinct period of time during development when the system is subjected to development testing and subsequent fixes made. A major test phase will generally lie totally within a development stage but a development stage may have more than one major test phase.

5.1.2.1 Basic Types of Test Programs. During a major test phase, one of three basic types of test and fix programs is conducted. The primary distinction among these programs is when fixes are incorporated into the system.

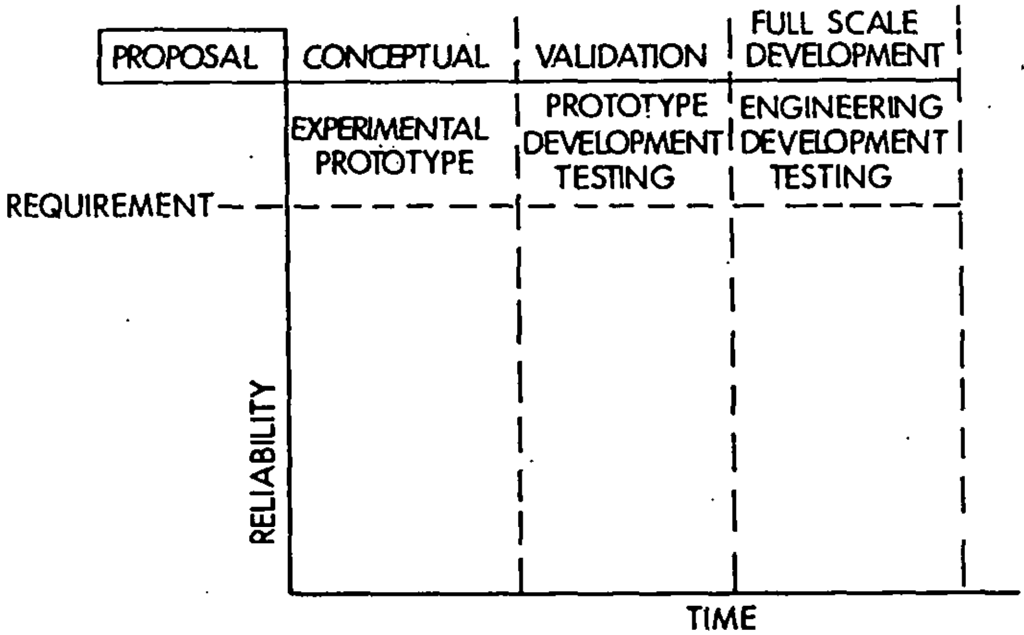


Figure 5.1 Typical Stages of a Development Program.

5.1.2.1.1 Test-fix-test. During a test-fix-test program the system is tested and problem failure modes determined. When a fix is found for a problem failure mode it is incorporated into the system which is retested to verify the fix and surface new problem areas. Since only a few fixes will generally be incorporated into the system at any one time, the reliability growth of the system during the test phase will typically be approximated by a smooth curve. See Figure 5.2.

5.1.2.1.2 Test-find-test. During a test-find-test program the system is also tested to determine problem failure modes. However, unlike the test-fix-test program fixes are not incorporated into the system during the test. Rather, the fixes are all introduced into the system at the end of the test phase and before the next testing period. Since a large number of fixes will generally be incorporated into the system at the same time, there is usually a significant jump in system reliability at the end of the test phase. The fixes incorporated into the system between test phases are called delayed fixes. See Figure 5.3.

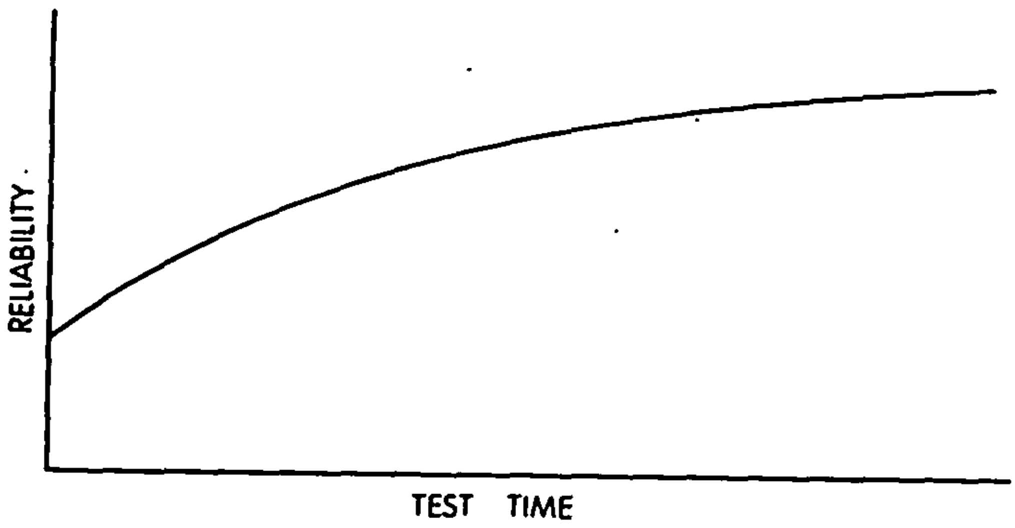


Figure 5.2 Reliability Growth During a Test-Fix-Test Program.

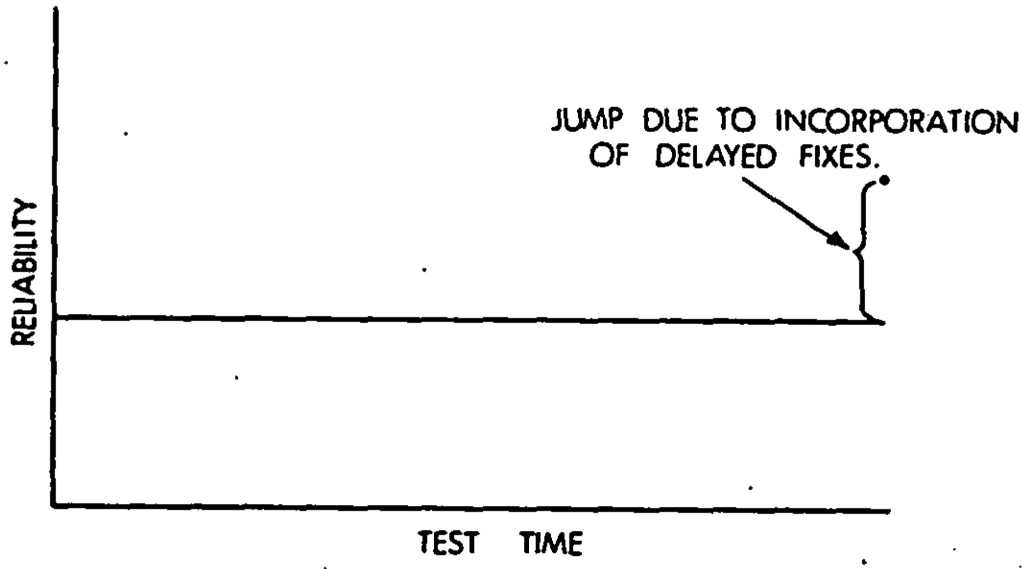


Figure 5.3 Reliability Growth During a Test-Find-Test Program.

5.1.2.1.3 Test-fix-test With Delayed Fixes. A type of test program which is commonly used in development testing is a combination of the two types discussed above. In this case, some fixes are incorporated into the system during the test while other fixes are delayed until the end of the test phase. Consequently, the system reliability will generally grow as a smooth process during the test phase and then jump due to the introduction of the delayed fixes. See Figure 5.4.

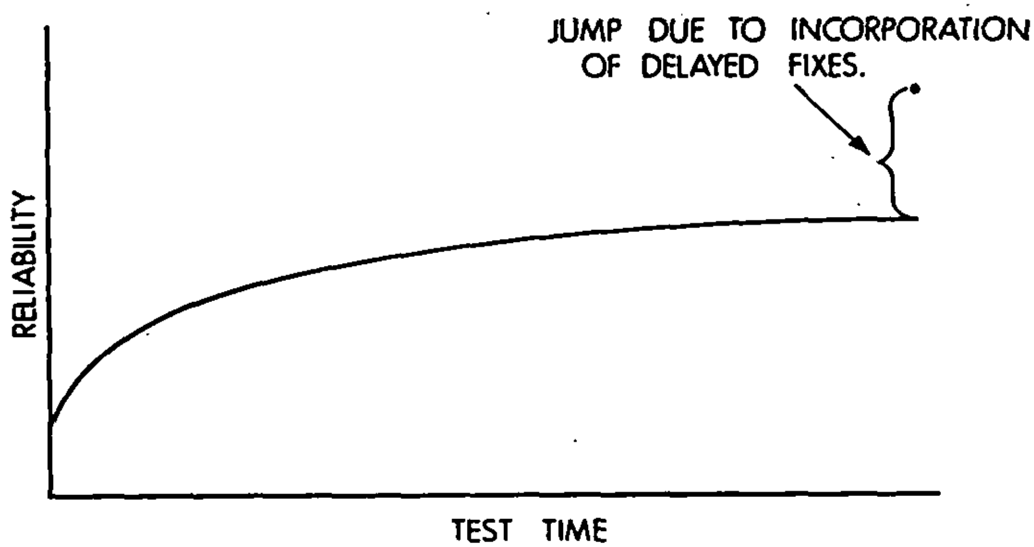


Figure 5.4 Reliability Growth During a Test-Fix-Test with Delayed Fixes Program.

5.1.3 Reliability Growth During Development Testing. The development testing program will usually consist of several major test phases, and within each test phase the testing may be conducted according to any one of the three test programs discussed above.

As an example, suppose that testing were conducted during the validation and full-scale development stages of the program. Each stage would be considered at least one major test phase, implying a minimum of two major test phases for the program. In this case, there would be $3^2 = 9$ general ways the reliability may grow during the development testing. A development stage may consist of more than one distinct test phase. For example, suppose that testing is stopped part-way through the full-scale development stage and delayed fixes incorporated into the system. The testing may, in this case, be considered as two major test phases during this stage. If the program had three major test phases then there would be $3^3 = 27$ general ways reliability may grow.

For purposes of discussion, assume that the development testing program consists of two major test phases. There would, of course, be $3^2 = 9$ possible test programs for reliability growth. See Figure 5.5.

Figure 5.5.1 illustrates testing conducted according to a test-fix-test program for each of the test phases. Figure 5.5.2 illustrates testing conducted according to a test-fix-test program for both test phases with delayed fixes incorporated at the end of the second phase, resulting in a jump in reliability. Note that for the plan illustrated in Figure 5.5.2, the impact of the delayed fixes at the end of the test cannot be evaluated from actual test results. Hence, whether or not the jump is sufficient to achieve the requirement will not be known until after production. On the other hand, the reliability can be evaluated during the test itself and compared against the goal to be achieved by the introduction of fixes during the test. Compare this with the test program illustrated in Figure 5.5.1. In this case, the reliability can be progressively evaluated throughout the test phase and the final estimate at the end of test would be compared with the reliability goal. If a determination is made during the test that the goal will not be met with the present effort, then corrective action can be taken before the end of the test phase.

The reliability growth is often depicted as a function of test time for evaluation purposes. For management and presentation purposes it may be desirable to portray reliability growth as a function of calendar time. This representation, of course, is a direct function of the program schedule. Figure 5.6 shows the reliability growth of a system as a function of test time (flight number) and calendar time.

<u>5.1.4 Idealized Reliability Growth Curve</u>. The reliability growth for a system is often depicted in a smooth fashion after some point in the development program, as shown in Figure 5.7. In general, however, a smooth process does not convey the way reliability will actually grow during development, as noted in Section 5.1.3. This smooth representation of reliability growth is basically idealistic. If we divide the development testing program into its major test phases and join by a smooth curve reliability values for the system during the test phases, then the resulting curve represents the general, overall pattern for reliability growth. The horizontal line prior to the smooth curve in Figure 5.7 represents the baseline or initial reliability of the system. This baseline for reliability together with the smooth curve is called the idealized reliability growth curve. The idealized curve is very useful in quantifying the overall development effort and serves as a significant tool in the planning of reliability growth.

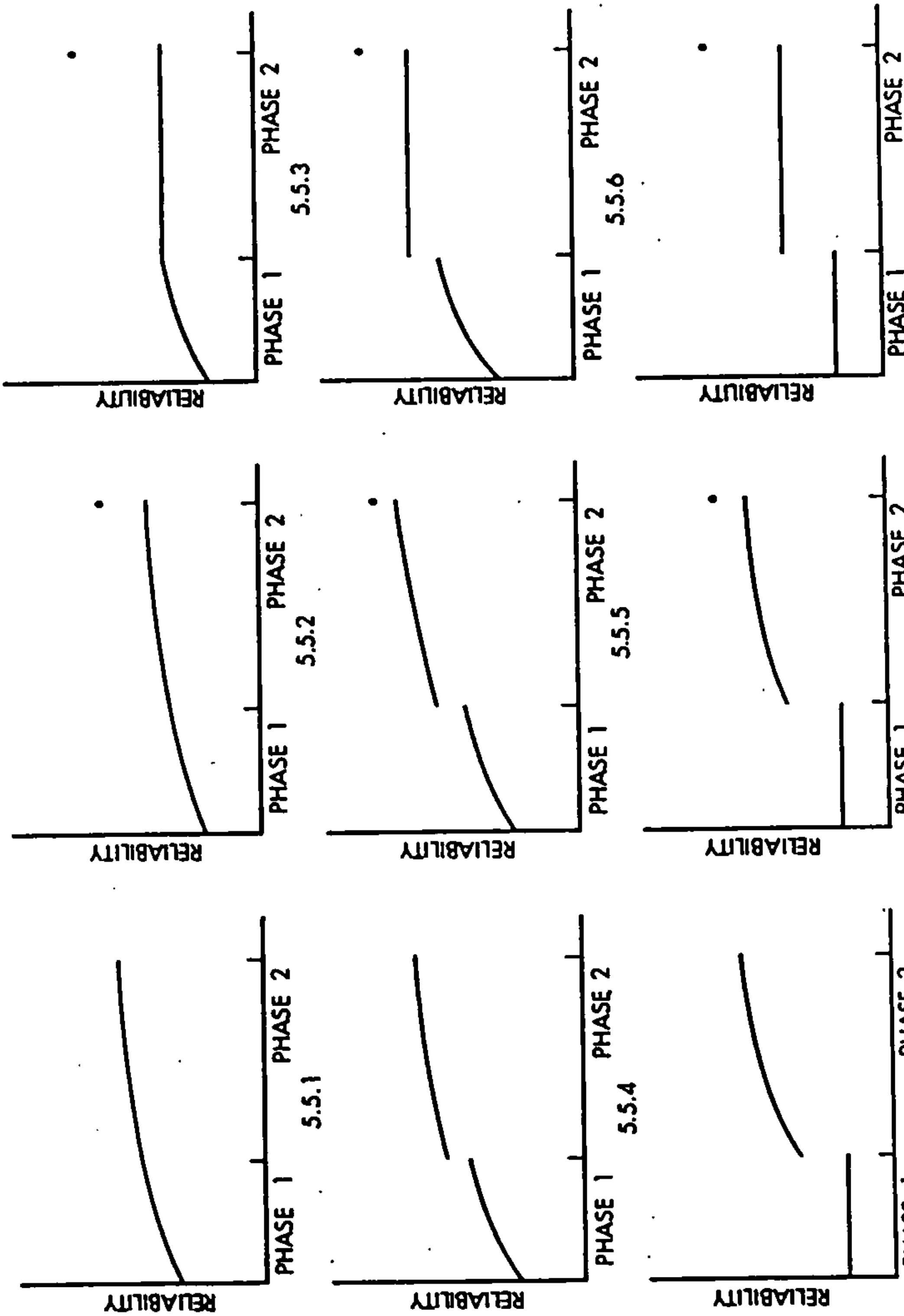


Figure 5.5 Nine Basic Test Programs for Two Test Phases.

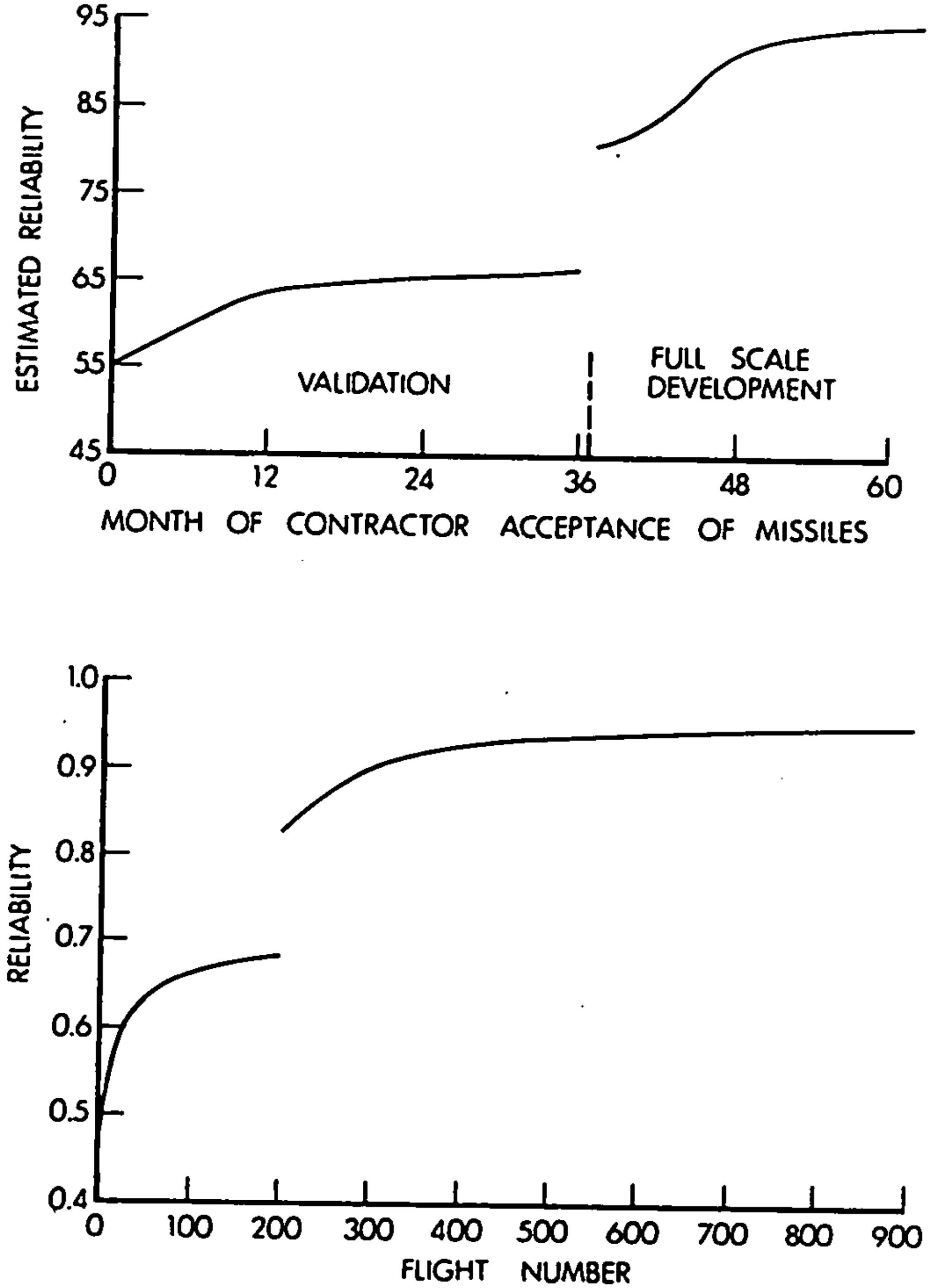


Figure 5.6 Reliability Growth Curve for a Missile as a Function of Calendar Time and Flight Number.

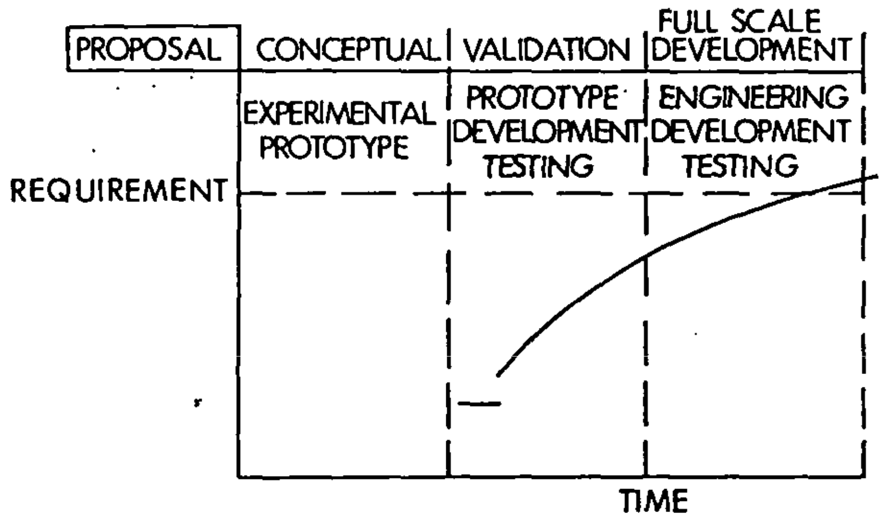


Figure 5.7 Idealized Reliability Growth Curve.

5.1.5 Planned Reliability Growth Curve. The planned reliability growth curve is constructed early in the development program, generally before hard reliability data are obtained, and is typically a joint effort between the program manager and contractor. Its purpose is to give a realistic and detailed indication of how the system reliability is planned to grow during development with the allocated resources such as time, money, personnel, the available prototypes for testing, the test schedule, major emphasis of the various tests, management control, type of hardware, etc.

There should be one planned growth curve for the entire system and one planned growth curve for each major subsystem which has a reliability requirement. These planned growth curves should be acceptable to both the contractor and program manager and should be compatible with the idealized growth curves which reflect the overall planned growth patterns.

For each major test phase the planned curve should indicate the type of development testing program for reliability that will be conducted, namely, test-fix-test, test-find-test, or test-fix-test with delayed fixes. The initial reliability planned for the beginning of each major test phase should be given along with the reliability growth to be achieved due to the incorporation of fixes during the test and due to the introduction of delayed fixes. The planned growth curve should meet or exceed the reliability requirement at the end of the program where the requirement is to be met.

Figure 5.8 is an example of a planned growth curve and the corresponding idealized curve. (See also Figure 5.9.) A point on the planned curve at any given time in the program is the level of reliability to be achieved at that time by the contractor.

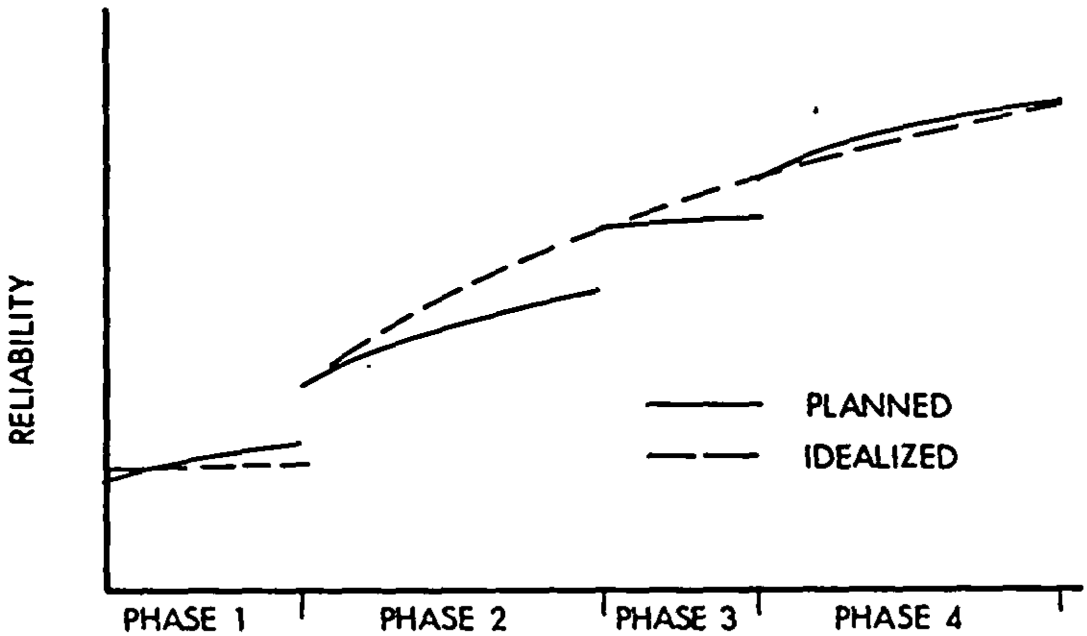


Figure 5.8 Example of Planned Growth Curve and Corresponding Idealized Curve.

To be an effective management tool it is not sufficient for a planned growth curve to be simply an increasing curve which goes through the reliability requirement at the end of the program. The planned growth curve should reflect a level of reliability to be attained at each milestone which is sufficient to ensure that the program manager has viable alternatives available at major decision points to meet future reliability goals. An evaluation of the alternatives available at each major decision point will indicate the risk associated with a proposed planned growth curve.

5.1.6 Tracking Reliability Growth. The planned growth curve and interim goals provide a yardstick which can be used to gauge the progress of the reliability effort. To obtain an objective evaluation, the program manager needs a demonstrated numerical measure of the system reliability during the development testing program based on the test data. In addition to the demonstrated estimate, it is often desirable to make reliability projections beyond the time associated with the most recent test data.

Since typical development test programs are conducted on a phase-by-phase basis, the reliability evaluations are normally conducted on that same phase-by-phase basis. At the end of each phase, therefore, the incidents occurring during the test must be classified in accordance with the failure definition before a reliability assessment can take place. (See Section 5.4.1.3.4.)

5.1.6.1 Demonstrated Reliability Value. A demonstrated reliability value is a reliability estimate based on test data for the system configuration under test at the end of the test phase. If, for example, design changes are proposed but have not been introduced into the system by the end of the test then the impact of these fixes on the system's reliability would not be considered in the determination of the demonstrated value. This estimate is based on actual system performance of the hardware tested and not of some future configuration. A demonstrated reliability value should be determined at the end of each test phase.

The demonstrated reliability is usually determined by one of two methods. When appropriate, the preferred method is reliability growth analysis. However, should the data not lend itself to this type of analysis, then the second method, an engineering analysis, should be used.

5.1.6.1.1 Reliability Growth Analysis. During a test phase the configuration of the system may be changing with the introduction of fixes for problem failure modes. Consequently, in the presence of reliability growth the data from the earlier part of the test phase would not be representative of the current configuration. On the other hand, the most recent test data, which would best represent the current configuration, may be limited so that an estimate based upon the recent data would not, in itself, be sufficient for a valid determination of reliability. Because of this situation, reliability growth models are often employed. These are mathematical models of the reliability growth process which are useful for combining test data to obtain a demonstrated estimate in the presence of a changing configuration.

5.1.6.1.2 Engineering Analysis. The nature of the data may be such that normal reliability growth procedures cannot be employed and an engineering analysis must be performed. This technique involves adjusting estimates determined from the test data to reflect the impact that fixes have had on the reliability of the system. However, adjustments can only be made when the fixes have been proven to be effective based on verified testing.

5.1.6.2 Projected Reliability. While the demonstrated reliability estimate provides an estimate of the current system reliability based upon the test data, it is usually informative and, indeed, often necessary to project the reliability beyond the present test time. Reliability projections are needed so that timely decisions can be made based upon what is expected at some future point, such as the beginning of the next test phase, the end of the development phase, etc.

Projections may be based on test data, engineering judgment, and other pertinent information. A projection can account for proposed fixes to be incorporated after the end of the test phase and for late fixes that were incorporated near the end of the test phase but may not be reflected fully in the demonstrated reliability value because of limited test exposure. A projection, by its very nature, will generally be less precise than the demonstrated value, but it serves the basic purpose of quantifying the present reliability effort relative to the achievement of future milestones.

5.1.7 Reliability Growth During Operational Testing. Operational testing will usually be a test-find-test type program conducted in the latter stages of each major test phase. Results may reflect incorporation of fixes from earlier development testing (DT) test-fix-test or test-find-test programs and be useful in determining the effectiveness of the fixes. The operational nature of the testing may result in reliability estimates which are inconsistent with DT results obtained under different conditions. This may be true even if the same failure definition is applied to results. Abrupt changes or inconsistencies in the reliability growth pattern emerging from operational testing should be carefully assessed. Reliability growth predictions should consider the potential impact of operational conditions on the reliability estimates. Exposure to the operational conditions by means of operational testing (OT) early in the development program can be useful in determining this potential impact in latter stages of development.

5.1.8 Management Guidelines. Although there is no absolute guarantee that a reliability goal will be met, planning and controlling the reliability growth process adds assurance that realistic objectives will be met within the program constraints, and reduces the risk of accepting a system with significant reliability deficiencies. The general concepts associated with planning and controlling the reliability growth during development testing have been discussed relative to the roles of the idealized, planned and tracking reliability growth curves. The utilization of these concepts in themselves, do not, of course, cause reliability growth. These are concepts and methods for realistically setting objectives and assessing what has been achieved against interim goals and the requirement.

In planning the reliability growth, the major role of the idealized curve is to quantify the overall development effort so that the growth pattern can be evaluated relative to the basic objectives and resources of the particular program under consideration. Section 5.2.6 discusses the general construction of idealized growth curves. A typical idealized growth curve profile, as discussed in Section 5.2.6.1, is illustrated in Figure 5.9.

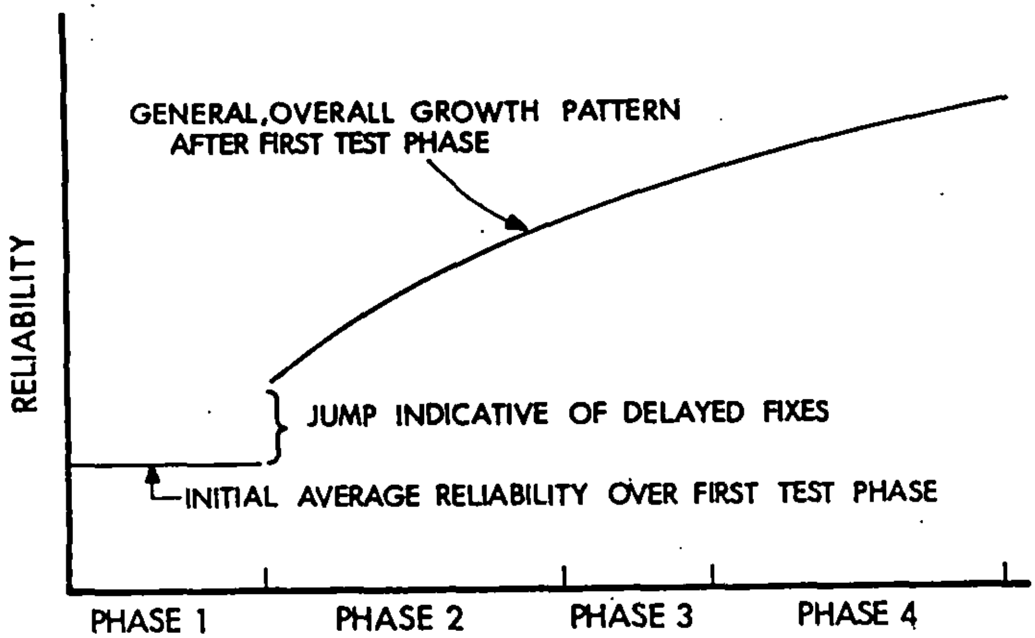


Figure 5.9 Example of Idealized Growth Curve.

The planned growth curve lays out a more detailed plan of how the reliability growth will actually be achieved. The proper construction of the planned growth curve forces a thorough consideration of the allocated resources, test schedules and many other important factors which are characteristic of the program. See Figure 5.10.

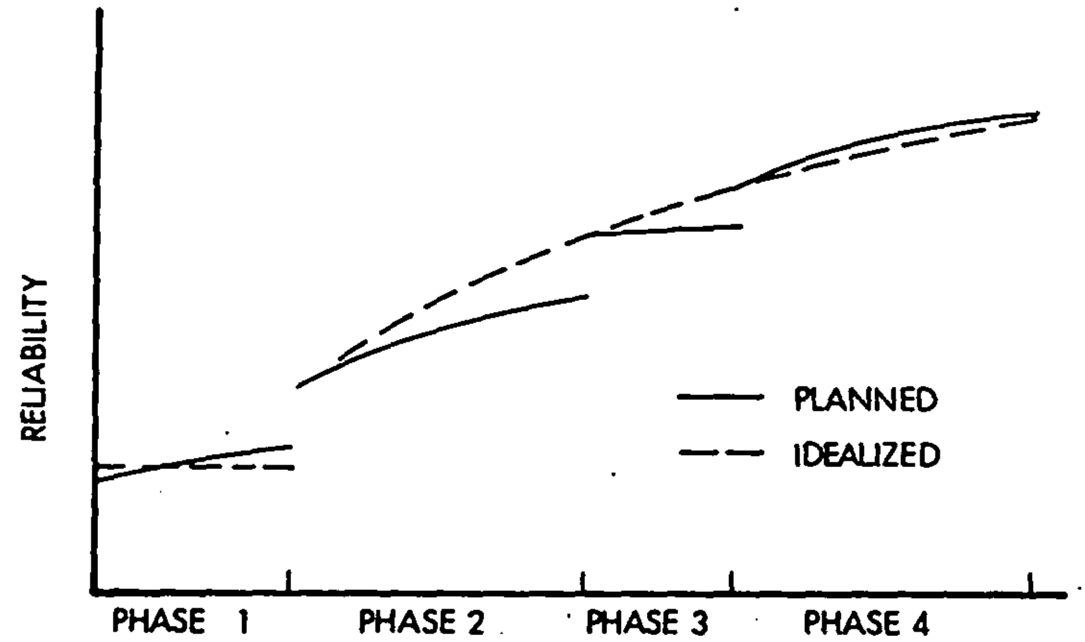


Figure 5.10 Example of Planned Growth Curve.

The development testing program should be planned so that the program manager will have viable alternatives available to him at each major decision point. That is, the program manager should have sufficient remaining resources to take meaningful corrective action, when necessary, in order to achieve the reliability objectives. For example, if the interim goals are set too low, then they may be met during development testing with little or no reliability growth. However, toward the end of the program or test phase, there is likely to be a situation where the reliability must increase significantly in a very short period of time in order to meet the objectives. If fixes are incorporated into the system at a late date, there may not be sufficient remaining test time to evaluate, from actual system performance, the impact of these changes. Hence, the program manager may have no real assurance that the system reliability is acceptable prior to making a major program decision, and the remaining resources may be insufficient to correct the situation if the objectives were not met.

The development testing program should be structured so that the reliability growth can be effectively tracked. Reliability growth tracking is conducted on a phase-by-phase basis. The type of testing planned for the test phases is a major factor in the amount of information that will be provided to the program manager for evaluating the progress of the program and assessing the likelihood of achieving the interim goals and final requirement.

During a test-fix-test program fixes are incorporated into the system and the system is further tested. This additional testing provides information on how effective the fixes are that were previously introduced. Hence, data from this testing plan can be used to evaluate the progress of reliability during the test itself. By measuring this progress, the program manager can assess the likelihood of attaining the goal set for the end of the testing phase. By testing and verifying, the program manager has a means of surfacing major problem areas before the end of the test phase so that time and other resources remain to take corrective action if necessary. With sufficient test time allocated so that the data provide meaningful information, the program manager has a method to significantly control the reliability growth effort.

During a test-find-test program fixes are not incorporated into the system until the end of the testing phase. Therefore, the impact of these fixes cannot be ascertained until the next test phase. If the reliability level is not satisfactory as a result of these fixes, the program manager cannot take corrective action until the next test phase when this deficiency is surfaced. If corrective action is, in fact, necessary there will, of course, be less time and other resources available than if the problem was recognized earlier. Thus, in this regard, there is a higher risk associated with the test-find-test program than with test-fix-test. Moreover, during the last test phase there would be no verification of fixes until the system was tested after production. At this point, there would generally be no remaining resources to recover the program if in fact the final reliability requirements were not met.

For the test-fix-test with delayed fixes program, part of the reliability increase is achieved as a consequence of the fixes incorporated into the system during the test itself with the remainder achieved as a result of the introduction of delayed fixes at the end of the test period. Therefore, during the test the data may be used to measure how well the program is progressing in relation to the interim goal to be achieved at the end of testing. A measure of the impact of the delayed fixes on the system's reliability will not be available until the next test phase. By not having this information until the next test phase, there is more risk associated with this test plan than with the test-fix-test plan. However, there is less risk than with the test-find-test plan because information can usually be obtained from the data regarding the impact of the fixes introduced during the test.

There are three primary reliability values on the planned curve of interest when tracking the reliability growth during a major test phase. These are: (a) the reliability value (A) planned for the beginning of the test phase, (b) the reliability (B) to be achieved as the result of incorporating fixes into the system during the test, and (c) the reliability value (C) to be achieved as the result of introducing delayed fixes into the system at the end of the test phase. See Figure 5.11.

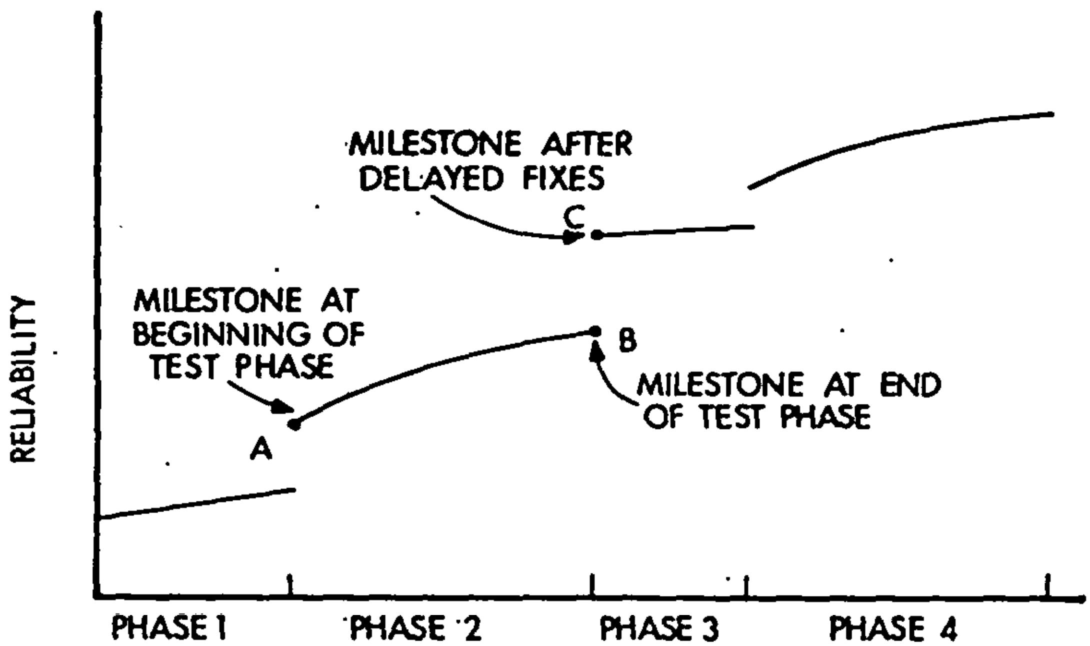


Figure 5.1.1 Reliability Milestones Associated with a Major Test Phase.

In the early stage of the test phase, estimates of the system reliability would be compared to A to determine if the initial reliability was satisfactory. As the testing continues new data are generated which can be used for obtaining additional estimates of the reliability. These can be compared to the planned growth curve. If a growth rate is established, e.g., from a growth model, then the reliability growth curve can generally be extrapolated. If the extrapolation is to the end of the test phase, then this estimate would be compared to B. It is important to note that the estimate would not be compared to C, since the extrapolation is based on the calculated growth rate determined from the fixes incorporated into the system during the test phase. See Figure 5.12 . If the extrapolation indicates that it is unlikely that goal B will be met with the present effort, then the program manager can take appropriate corrective action before the end of the test period.

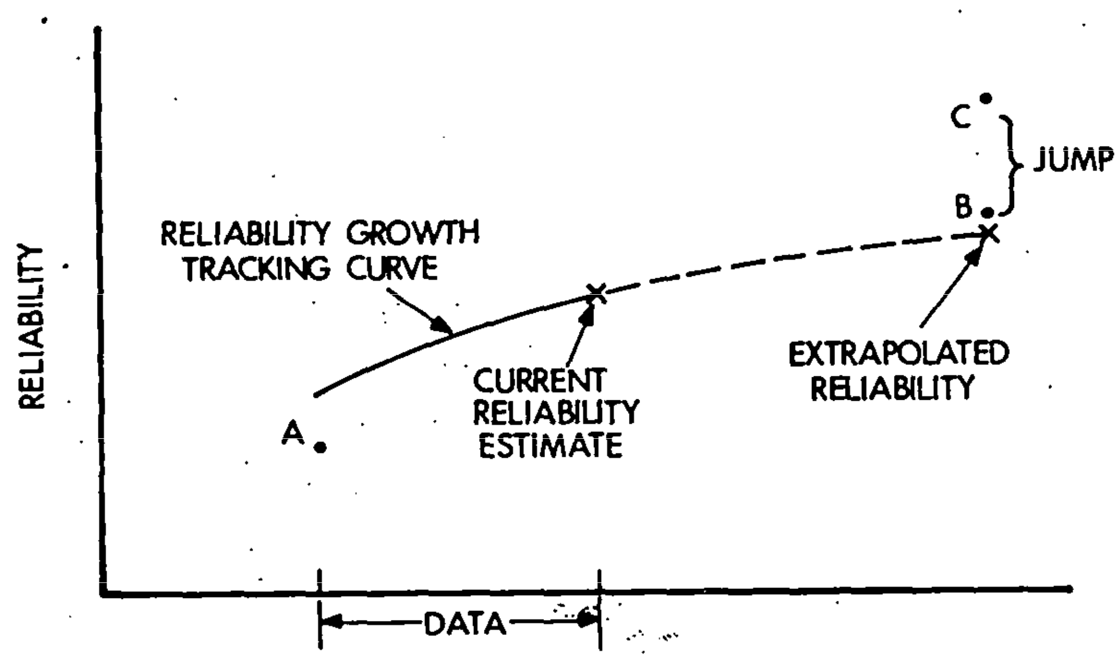


Figure 5.12 Reliability Growth Curve for Major Test Phase.

At the end of a test phase incidents would be classified in accordance with the failure definition, and demonstrated and projected reliability values determined. The demonstrated value is a reliability estimate for the configuration of the system on test at the end of the test phase. The demonstrated value is based on data generated during the test phase. The projected value is an estimate of the reliability expected going into the next phase. This estimate is based on an engineering assessment of the delayed fixes to be introduced into the system at the end of the test phase. See Figure 5.13. The demonstrated and projected values at the end of a test phase are compared to milestones B and C, respectively.

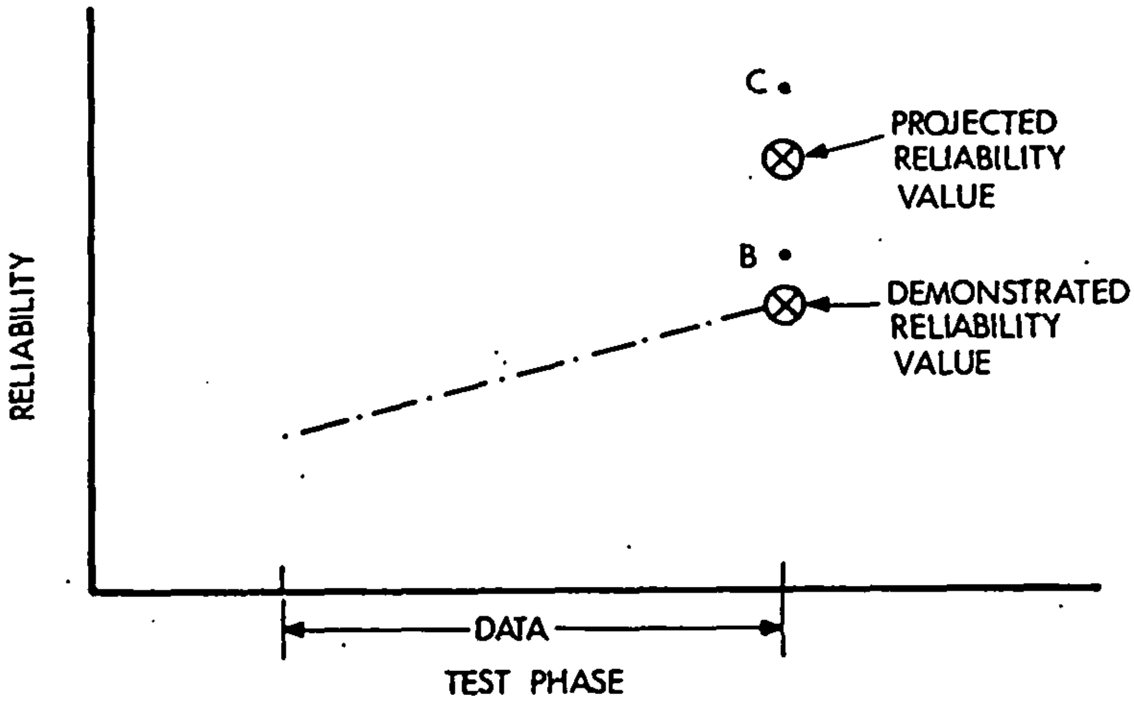


Figure 5.13 Demostrated and Projected Reliability Values at End of Test Phase.

5.2 Planned Growth Curves. Development of the planned growth curve is an application of the "lessons learned" from previous program experiences to predict the growth that can be expected in a future program. The importance of this curve must be understood. When hard reliability data have begun to be generated, the results will be compared with the predicted values given by the planned curve to determine if the reliability growth is progressing satisfactorily. For information on the management uses of the planned growth curve, see Section 5.1.

5.2.1 General Development of the Planned Growth Curve. The detailed planned growth curve provides, as precisely as possible, the phase-by-phase development of reliability improvement that is expected. Each test phase must be carefully considered to determine the type of testing that will be conducted and the impact of the fixes that can be anticipated. The role of the idealized growth curve is to substantiate that the planned growth follows a learning curve which, based on previous experience, is reasonable and can be expected to be achieved. The following paragraphs describe how planned growth curves may be developed for specific programs. Every program can, however, be expected to require some modification of the suggested procedures.

5.2.2 General Concepts for Construction. In general, there are two basic approaches for constructing planned growth curves. The first method is to determine the idealized growth pattern that is expected or desirable, and to use this as a guide for the detailed planned curve. The second method is just the reverse. In this case a proposed planned curve is first developed which satisfies the requirement and interim milestones. The idealized curve is then constructed and evaluated to determine if this learning curve is reasonable when compared to historical experience. If not acceptable a new detailed curve would need to be developed.

5.2.3 Understanding the Development Program. Development of planned growth curves requires a fairly complete understanding of the proposed development program, particularly the reliability program and all other program activities and constraints that will affect reliability. In the case of mechanical equipment, an understanding of the hardware is useful in evaluating the delays that should be associated with design changes. For complex test programs a logic diagram should be used to show the relationships between those phases in which failure modes will be found and those phases which will have the resultant "fixes" in the hardware. The expected policy for incorporating fixes must be understood. For systems with high reliability, the expected number of failures during the test program should be determined to give an indication of the number of fixes that can be anticipated. For initial estimating purposes this may be based on the starting MTBF.

5.2.4 Portraying the Program in Total Test Units. Although the planned growth curve is usually portrayed in final form as a function of calendar time for management use, the analytical development of the curve is done as a function of test units. Test units may be hours, miles, rounds, or similar units; and in some cases, the use of multiple units (e.g., both miles and rounds) may be appropriate. Figure 5.14 shows an example of a development program portrayed in calendar time, and Figure 5.15 shows the same program portrayed in cumulative miles.

5.2.5 Determining the Starting Point. The initial reliability for a system under development will typically not be known at the time when the planned curve is developed. A starting point for the planned

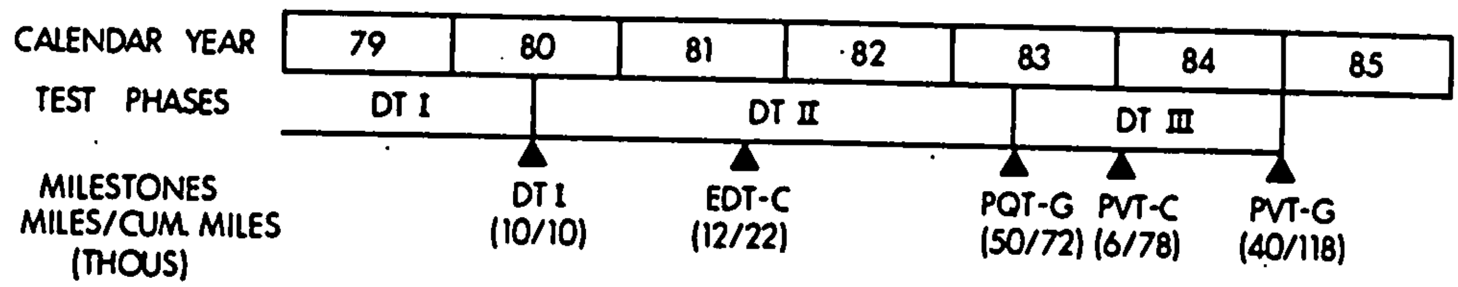


Figure 5.14 Development Program Portrayed in Calendar Time.

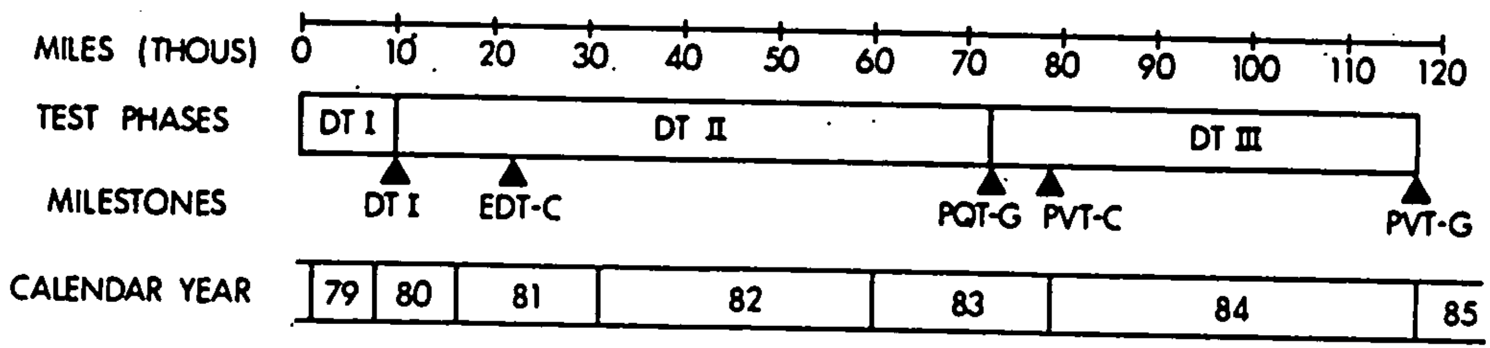


Figure 5.15 Development Program Portrayed in Test Units.

growth curve may, however, be determined from (1) using information from previous programs on similar systems, (2) specifying a minimum level of reliability that management requires to be demonstrated early in order to have assurance that the reliability goals will be met, and (3) conducting an engineering assessment of the design together with any previous test data that may exist, e.g., bench test, prototype test.

The practice of arbitrarily choosing a starting point, such as 10% of the requirement, is not recommended. Every effort to obtain information even remotely relevant to a realistic starting point should have been exhausted before an arbitrary figure can be used.

5.2.5.1 Example of Determining a Starting Point. A planned growth curve is to be developed for a ground vehicle development program. One of the first steps in this process is to determine a starting point for this curve.

To establish a starting point, the reliability growth experience of a predecessor system is analyzed. It is found that an initial MMBF (mean miles between failures) of 183 miles was demonstrated during early engineering development. The predicted MMBF was 580 miles. So, at this point in development, the achievement was 183/580 = .32 of predicted. The system under development has about the same degree of design maturity as did its predecessor; but since the reliability program emphasis is somewhat greater, it is expected that perhaps .35 of the prediction, rather than .32, will be achieved. With a prediction of 410 miles for the current system, .35 (410) = 143 would be expected as a starting point. To further rationalize this estimate, some pre-development testing of the proposed system resulted in 5 failures in 493 miles. No significant design changes were incorporated during test, so the MMBF may be estimated as 493/5 = 99 miles. Some design change is planned prior to engineering development testing. Using engineering analysis methods similar to those described in Appendix A, it is estimated that 2 of these failures will be affected by design change. It is also estimated that the design changes will be 70% effective. The MMBF expected on entering engineering design testing is then 493/(5 - .7(2)) = 137 miles. This value gives additional support to the estimate of 143 miles.

5.2.6 Development of the Idealized Growth Curve. During development, management should expect that certain levels of reliability be attained at various points in the program in order to have assurance that reliability growth is progressing at a sufficient rate to meet the requirement. The idealized curve portrays an overall characteristic pattern which is used to determine and evaluate intermediate levels of reliability and construct the program planned growth curve. Growth profiles on previously developed, similar type systems provide significant insight into the reliability growth process and are valuable in the construction of idealized growth curves. Reliability growth information on previous programs should be used whenever possible to develop the idealized curve directly or as input into a model for development of the idealized curve.

5.2.6.1 Idealized Growth Model Based on Learning Curve Concept. If documented reliability histories for similar type systems are not available to provide a basis for the idealized curve of the system under consideration, then a general method based on the learning curve concept is an alternative. Appendix B provides a survey of various growth models. If the learning curve pattern for reliability growth assumes that the cumulative failure rate versus cumulative test time is linear on log-log scale, then the following method is appropriate for construction of the idealized growth curve. This method is based on the test phase structure of a development program for reliability growth, as discussed in Section 5.1.2. This approach gives a realistic method for placing the initial MTBF at the proper point in time and portrays a growth pattern which has a meaningful interpretation in terms of test phase reliability growth.

5.2.6.1.1 Summary of Method. The idealized growth curve M(t) discussed in this section has the form shown in Figure 5.16 and portrays a general profile for reliability growth throughout system testing. The idealized curve has the baseline value M_I over the initial test phase which ends at time t_1. The value M_I is the average MTBF over the first test phase. From time t_1 to the end of testing at time T, the idealized curve M(t) increases steadily according to a learning curve pattern till it reaches the final reliability requirement M_F. The slope of this curve on the log-log plot in Figure 5.16 is the growth parameter α. The parametric equation for M(t) on this portion of the curve is

$$M(t) = M_I \left(\frac{t}{t_1}\right)^{\alpha} (1-\alpha)^{-1}.$$

5.2.6.1.2 Basis of Model. This model assumes that the cumulative failure rate versus cumulative test time is linear on log-log scale when plotted at the ends of test phases or reporting periods. See Figure 5.17. It is not assumed that the cumulative failure rates follow the same pattern within test phases. In fact, if delayed fixes are incorporated into the system at the end of a test phase, or the reliability is held constant during a test phase, then this linear pattern within test phases would not hold.

To illustrate this approach let $t_1, t_2, \ldots, t_k$ denote the cumulative test times which correspond to the ends of test phases. It is assumed that $N(t_i)/t_i$ versus t_i, $i = 1, 2, \ldots, K$, are linear on log-log scale, where $N(t_i)$ is the cumulative number of failures by time t_i. That is, $\log N(t_i)/t_i$ is linear with respect to $\log t_i$. This implies that $\log N(t_i)/t_i$ can be expressed as $\log N(t_i)/t_i = \sigma - \alpha \cdot \log t_i$, where σ and α are, respectively, intercept and slope parameters. Let λ_I denote the initial average failure rate for the first test phase, i.e., $\lambda_I = N(t_1)/t_1$. Since $\log \lambda_I = \sigma - \alpha \log t_1$, it follows that $\sigma = \log \lambda_I + \alpha \log t_1$.

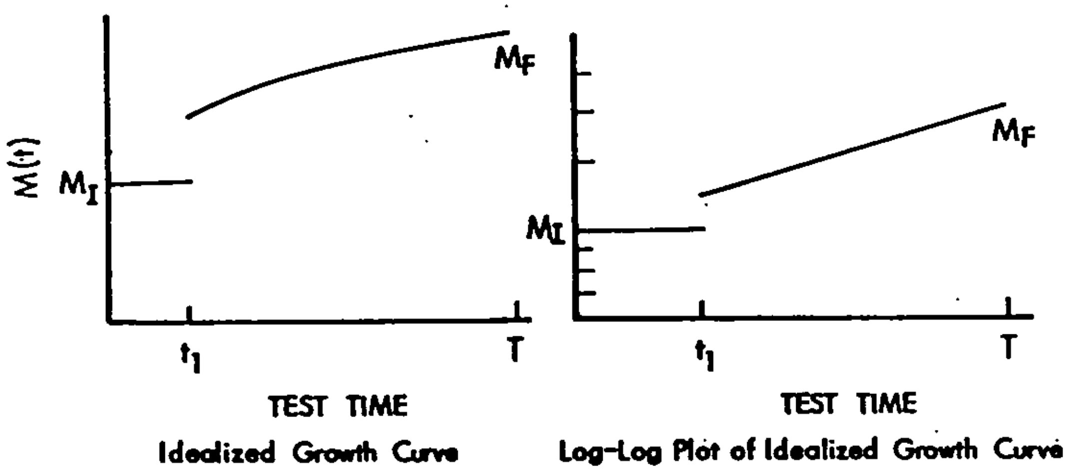


Figure 5.16 Idealized Growth Model.

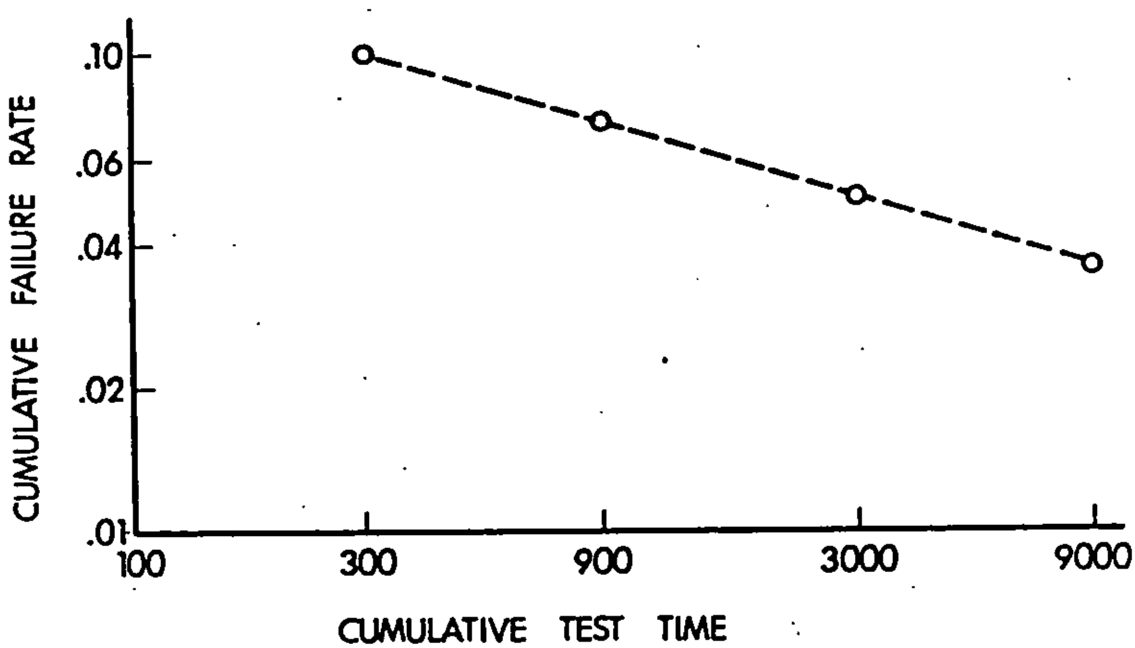


Figure 5.17 Example of Log-Log Plot at Ends of Test Phases.

Therefore, $\log N(t_i)/t_i = \log \lambda_I - \alpha \log\left(\frac{t_i}{t_1}\right)$. Consequently, the cumulative failure rate can be expressed as

$$N(t_i)/t_i = \lambda_I\left(\frac{t_i}{t_1}\right)^{-\alpha}.$$

This gives

$$N(t_i) = \lambda_I t_i\left(\frac{t_i}{t_1}\right)^{-\alpha}, \text{ or equivalently, } N(t_i) = \lambda_I t_1\left(\frac{t_i}{t_1}\right)^{1-\alpha}.$$

The average failure rate over the test interval t_{i-1} to t_i (the i-th test phase) is the total number of failures during this period divided by the length of the interval t_i-t_{i-1}. Therefore, the linearity of the cumulative failure rates at ends of test phases implies that the average failure rate λ_i for the i-th test phase is

$$\lambda_i = \frac{N(t_i) - N(t_{i-1})}{t_i-t_{i-1}}$$

where $N(t_i) = \lambda_I t_1 \left(\frac{t_i}{t_1}\right)^{1-\alpha}$. See Figure 5.18.

In terms of failure rate, this result for the average failure rates over the test phases is all that can be concluded from the linearity on log-log scale of the cumulative failure rates at ends of test phase. The reliability growth of the system in terms of MTBF is reflected by the increase in the average MTBF's $m_i = 1/\lambda_i$ over the test program. See Figure 5.19.

Now, the curve defined by $r(t) = \frac{d}{dt} N(t) = \lambda_I(1-\alpha)\left(\frac{t}{t_1}\right)^{-\alpha}$

crosses the average failure rates λ_i for each test phase. See Figure 5.20.

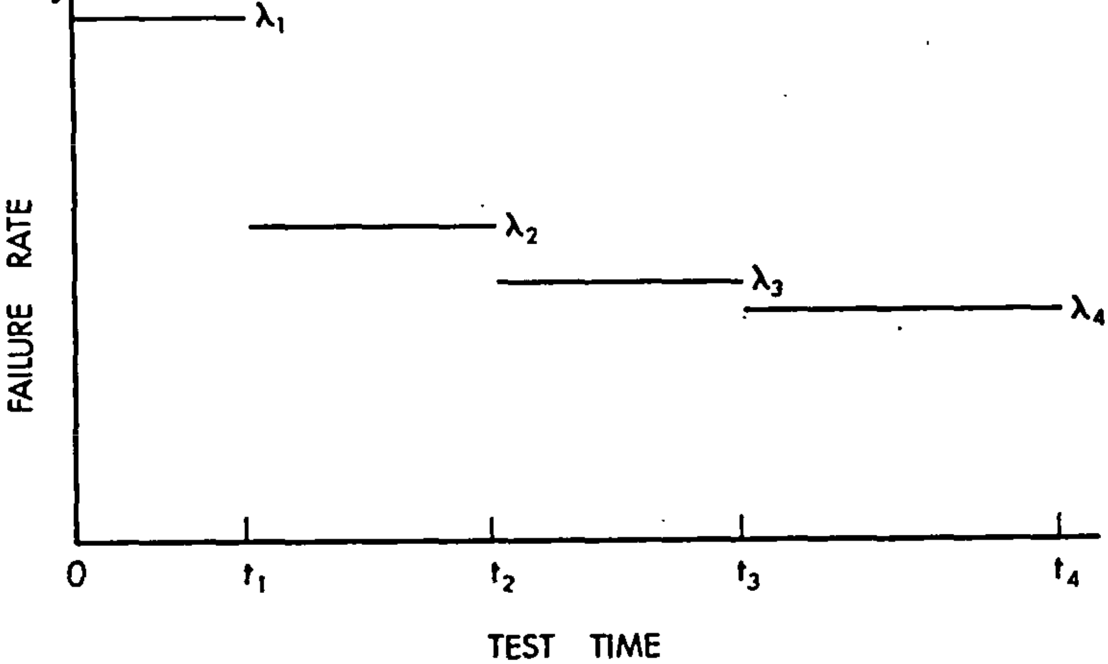


Figure 5.18 Average Failure Rates over Test Phases.

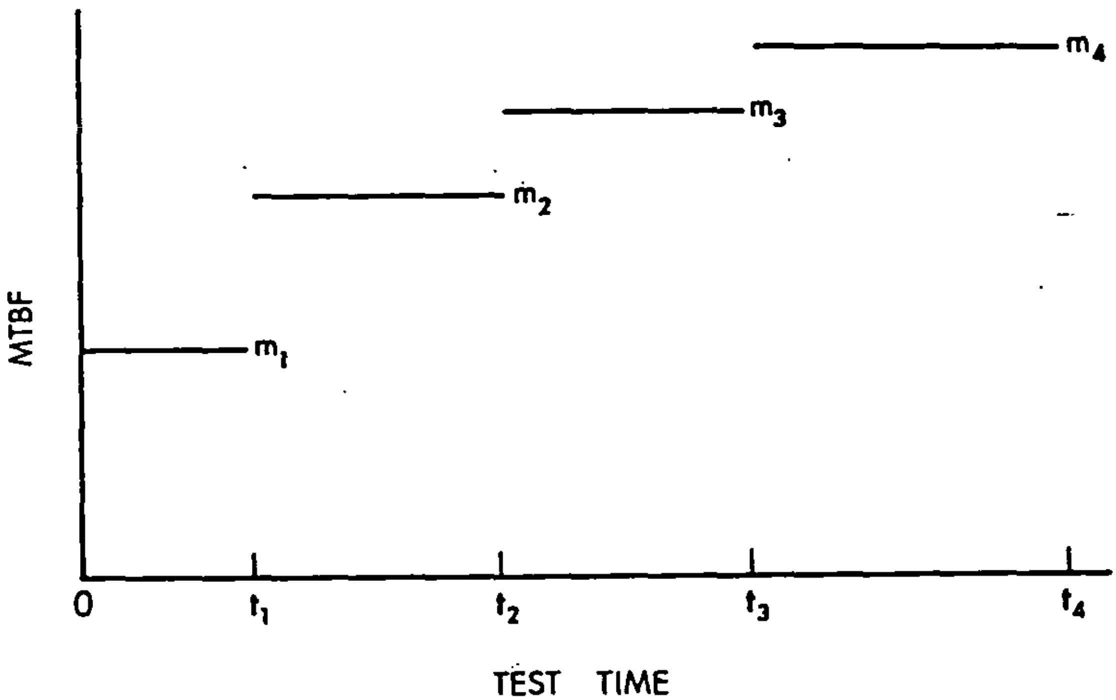


Figure 5.19 Average MTBF's over Test Phases.

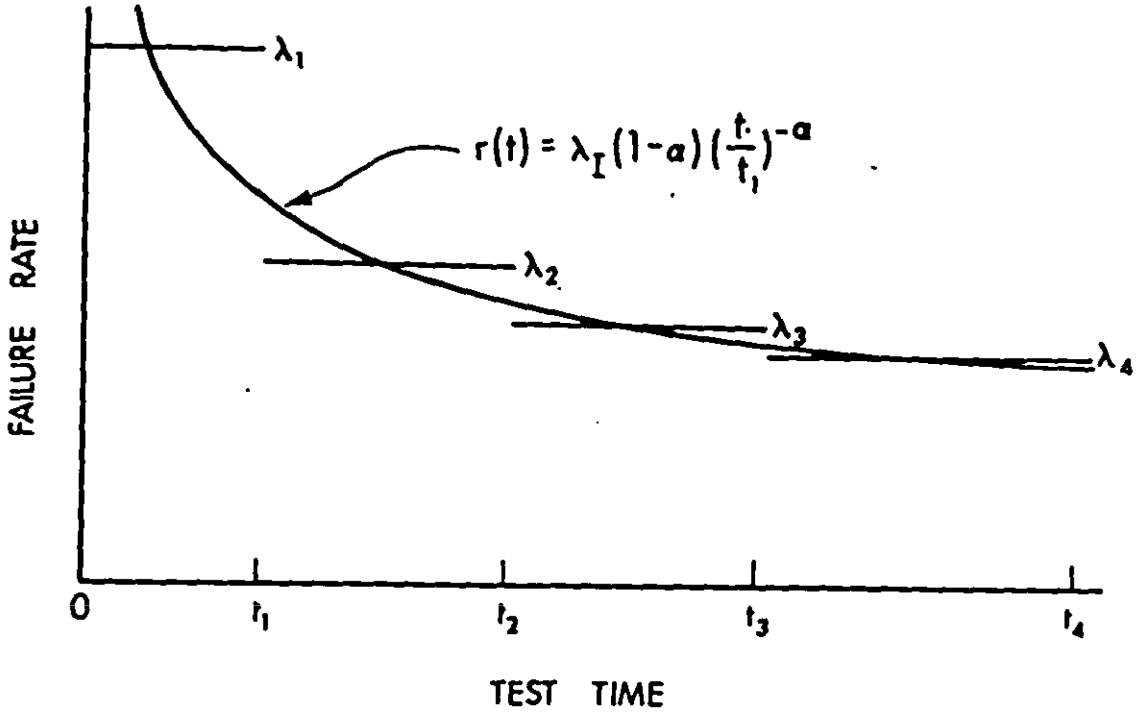


Figure 5.20 Average Failure Rates and r(t) Curve.

For any test phase the area under the curve $r(t)$ is equal to the area under the average failure rate. Therefore, for any test phase the average failure rate can be determined from $r(t)$. The reciprocal of the curve $r(t)$, $\overline{m}(t) = (r(t))^{-1} = M_I\left(\frac{t}{t_1}\right)^{\alpha}(1-\alpha)^{-1}$ also crosses the average MTBF m_i for each test phase. See Figure 5.21.

The actual underlying pattern for reliability growth is represented by the increase in the test phase average MTBF's. The growth in the individual test phase does not follow the smooth line $\overline{m}(t)$. In particular, note that the curve $\overline{m}(t)$ gives a value of 0 at test time 0, which is, of course, not a realistic value for the actual system MTBF at the beginning of development testing. However, the curve $\overline{m}(t)$ can generally be viewed as reflecting a meaningful trend for the average MTBF's after the first test phase. See Figure 5.22.

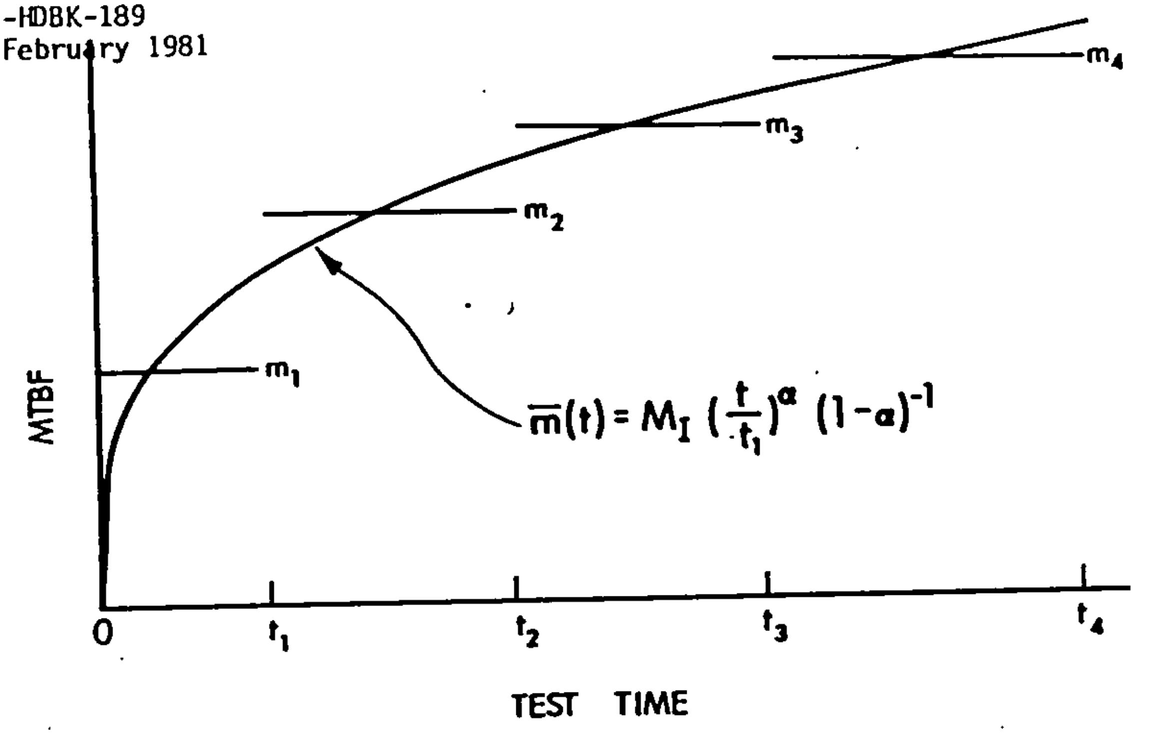


Figure 5.21 Average MTBF's and $\overline{m}(t)$ Curve.

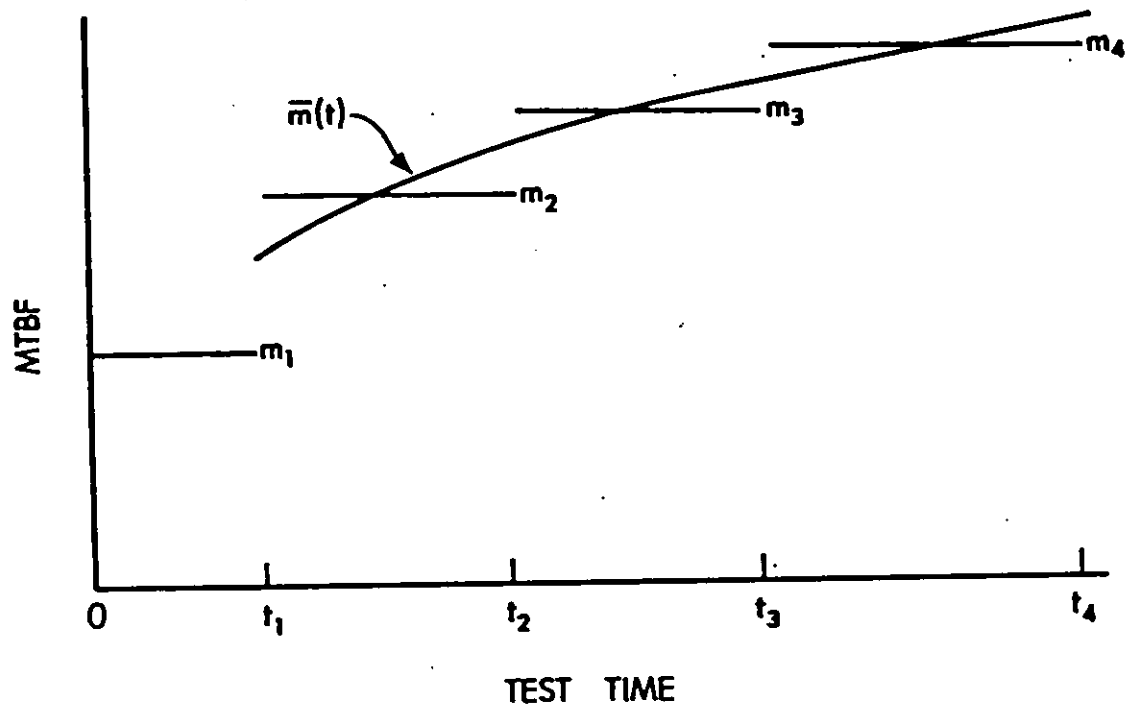


Figure 5.22 Average MTBF's and Modified $\overline{m}(t)$ Curve.

It is further noted that the baseline for reliability growth in terms of average MTBF's is the initial average MTBF $M_I = 1/\lambda_I$. Therefore, a practical and meaningful idealized growth curve is one that equals M_I over the first test phase and equals the curve $\overline{m}(t)$ over the remaining test time. This curve is denoted by M(t). See Figure 5.23 and Figure 5.24.

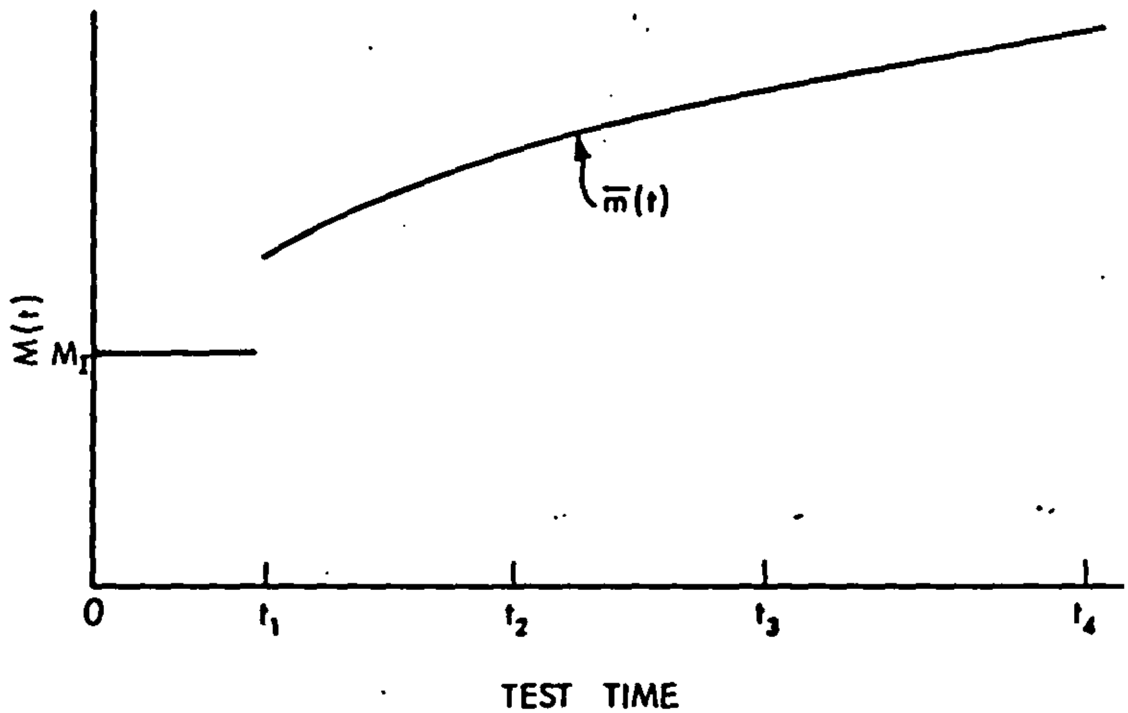


Figure 5.23 Idealized Growth Curve.

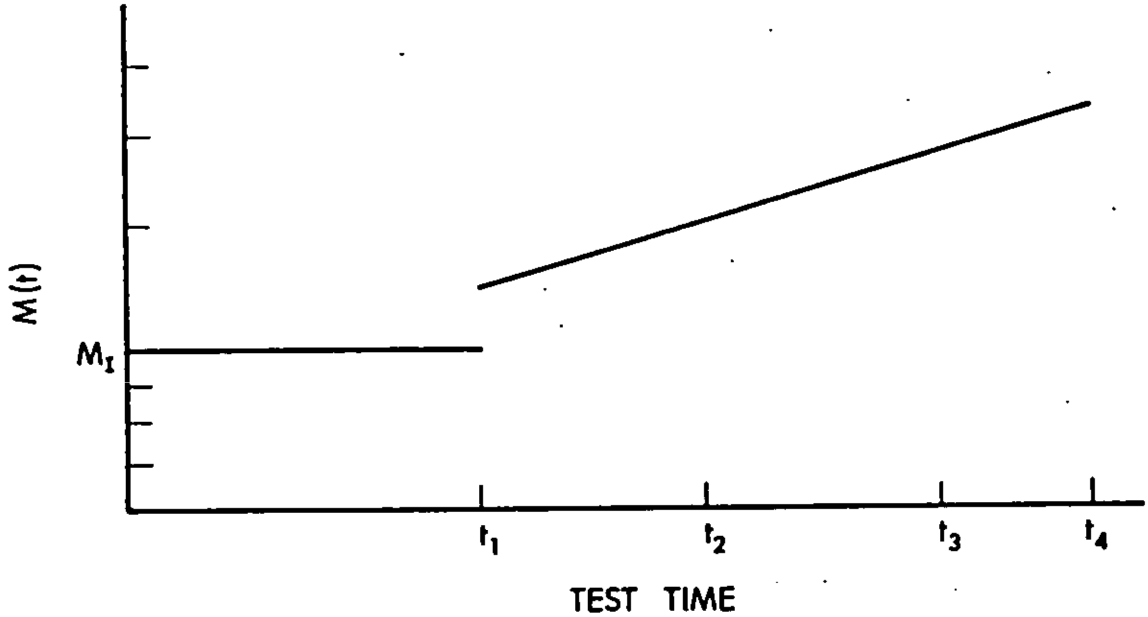


Figure 5.24 Log-Log Plot of Idealized Growth Curve M(t).

The idealized growth curve shows that the initial average MTBF over the first test phase is M_I, and that reliability growth from this average begins at t_1. This jump is indicative of delayed fixes incorporated into the system at the end of the first test phase. The idealized curve M(t) is a guide for the average MTBF over each test phase. Further, given that

$$M(t) = M_I\left(\frac{t}{t_1}\right)^{\alpha}(1-\alpha)^{-1} \quad \text{for } t>t_1,$$

then the average failure rate and the average MTBF for the i-th test phase can be determined by

$$\lambda_i = \frac{N(t_i)-N(t_{i-1})}{t_i-t_{i-1}}, \text{ and } m_i = 1/\lambda_i, \text{ where } N(t_i) = \lambda_I t_1\left(\frac{t_i}{t_1}\right)^{1-\alpha}.$$

See Figure 5.25.

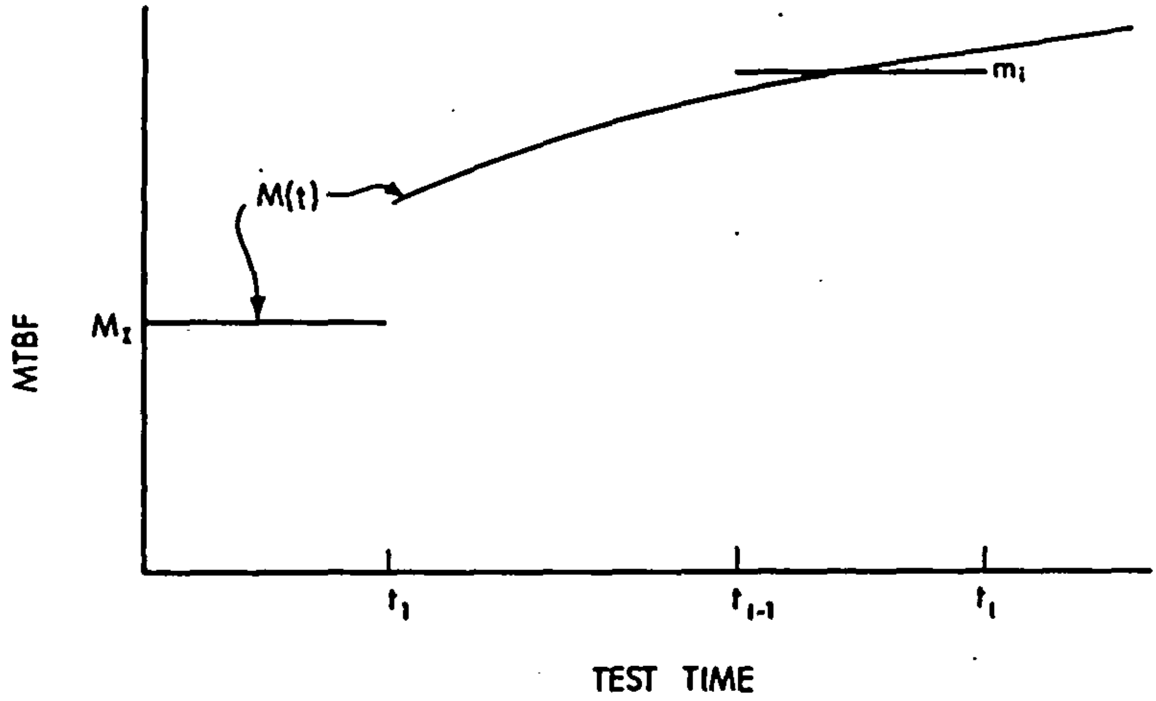


Figure 5.25 Average MTBF over i-th Test Phase.

In the application of the idealized growth curve model, the final MTBF value M_F to be attained at time T is set equal to M(T), i.e., $M_I\left(\frac{T}{t_1}\right)^{\alpha}(1-\alpha)^{-1} = M_F$. Also, the parameters M_I and t_1 of this model have the physical interpretations that M_I is the initial average MTBF for the system and t_1 is the length of the first test phase in the program. The parameter α is a growth parameter.

5.2.6.2 Procedures for Using Idealized Growth Curve Model. This section contains problems, solutions, and numerical examples which illustrate the application of the idealized growth model discussed in Section 5.2.6. The following notation and formulas are given for completeness.

Notation:

a. T - the cumulative test time over the test program.

b. $t_1, t_2, \ldots, t_k$ - the cumulative test times corresponding to the ends of test phases ($t_k = T$).

c. $N(t_i)$ - the cumulative number of failures by time t_i.

d. $H_i = N(t_i)-N(t_{i-1})$ - the number of failures during the i-th test phase.

e. $\lambda_i = H_i/(t_i-t_{i-1})$ - the average failure rate over the i-th test phase.

f. M_F - the final MTBF at time T.

g. $M_i = 1/\lambda_i$ - the average MTBF over the i-th test phase.

h. $\lambda_I = \lambda_1$ - the subscript I denotes initial average failure rate.

i. $M_I = 1/\lambda_I$ - the initial average MTBF.

j. α = growth parameter.

Model:

a. The idealized growth model M(t) is given by

$$M(t) = \begin{cases} M_I \text{ for } 0<t<t_1 \\ M_I\left(\frac{t}{t_1}\right)^{\alpha} (1-\alpha)^{-1} & \text{for } t>t_1. \end{cases}$$

where t_1 is the end of the first test phase.

b. Under this model

$$M_F = M_I\left(\frac{T}{t_1}\right)^{\alpha} (1-\alpha)^{-1}$$

and

$$N(t_i) = \lambda_I \, t_1 \left(\frac{t_i}{t_1}\right)^{1-\alpha}.$$

5.2.6.2.1 Case 1. How to Determine the Idealized Growth Curve.

Objective: Determine the idealized growth curve.

Given Conditions: T - the cumulative test time over the program.

t_1 - the test time for the first test phase.

M_I - the average MTBF over the first test phase.

M_F - the final MTBF at time T.

Solution for Case 1: Set $M_F = M_I\left(\frac{T}{t_1}\right)^{\alpha}(1-\alpha)^{-1}$ and find α such that $\frac{M_F}{M_I} = \left(\frac{T}{t_1}\right)^{\alpha}(1-\alpha)$. That is, find α such that

$$\log\left(\frac{M_F}{M_I}\right) = \alpha \log\left(\frac{T}{t_1}\right) - \log(1-\alpha).$$

Then the idealized curve is given by

$$M(t) = \begin{cases} M_I & 0<t<t_1 \\ M_I\left(\frac{t}{t_1}\right)^{\alpha}(1-\alpha)^{-1} & \text{for } t>t_1. \end{cases}$$

See Figure 5.26.

The following expression for α is a good second order approximation that is sufficient whenever α is less than 0.5:

$$\alpha = -\log\left(\frac{T}{t_1}\right) - 1 + \left[(1 + \log\left(\frac{T}{t_1}\right))^2 + 2\log\left(\frac{M_F}{M_I}\right)\right]^{1/2}.$$

The logarithms in this expression are natural logarithms.

Example of Case 1: Suppose that the initial MTBF for the system is estimated to be 45 hours and a final MTBF of 110 hours is desired after 10,000 hours of testing. For this program the first test phase is 1,000 hours. This is the point where delayed fixes will first be introduced into the system. Further, some reliability growth is planned during the first test phase so that an average MTBF of $M_I = 50$ hours is anticipated during the first phase. Determine the idealized growth curve. The parameter $\alpha = .23$ is found as the solution to

$$\log\frac{110}{50} = \alpha \log\frac{10000}{1000} - \log(1-\alpha).$$

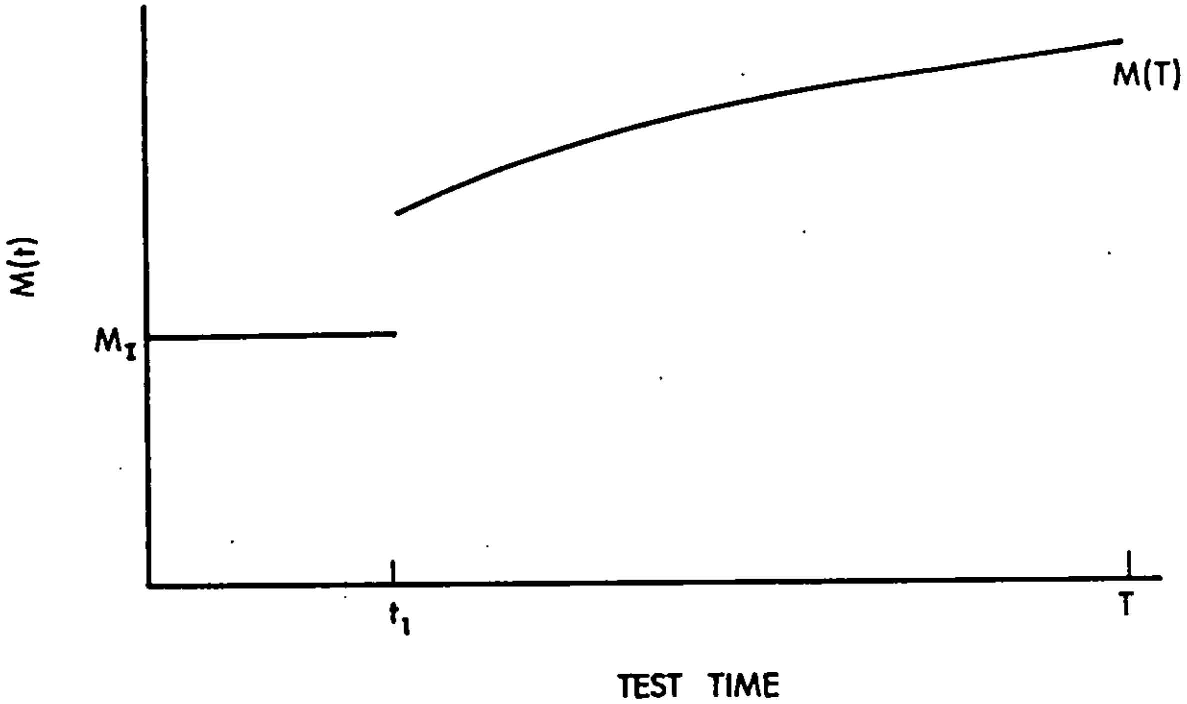


Figure 5.26 Idealized Growth Curve.

Therefore, if $\alpha = .23$ is acceptable the idealized growth curve is given by

$$M(t) = \begin{cases} 50 & 0<t<1000 \\ \frac{50}{.77}\left(\frac{t}{1000}\right)^{.23} & t>1000 \end{cases}$$

See Figure 5.27.

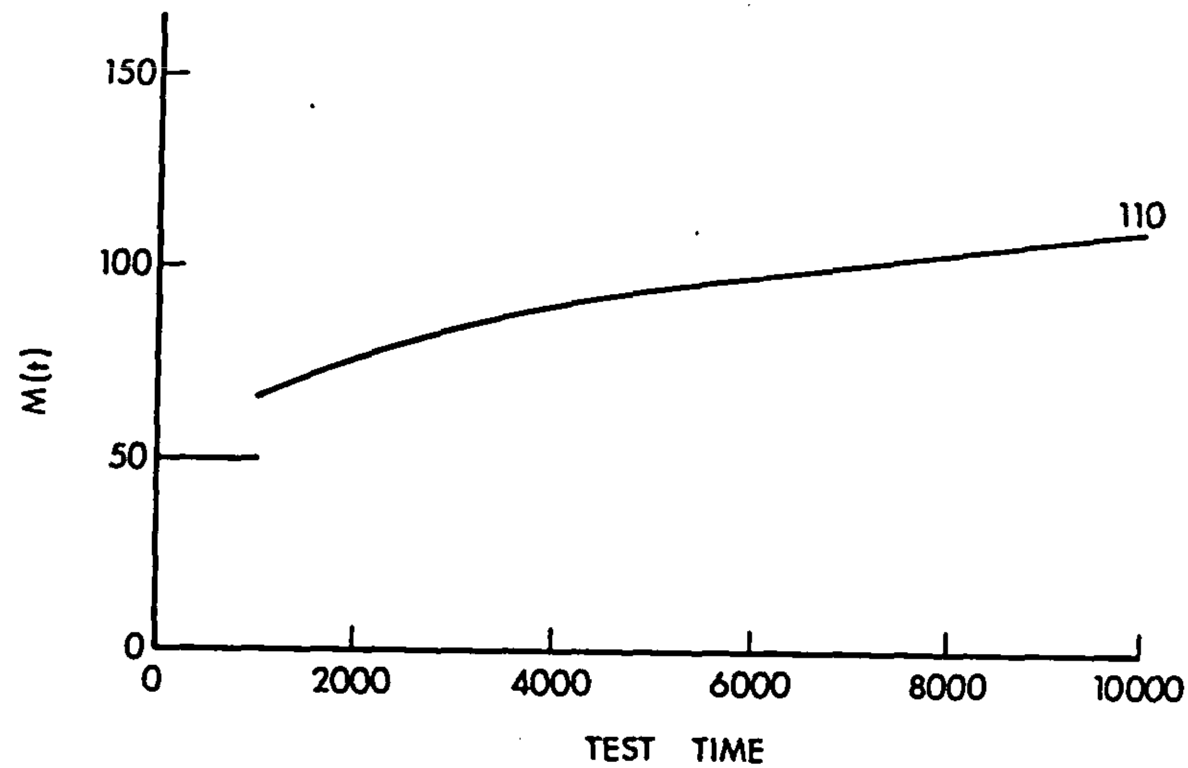


Figure 5.27 Example of Idealized Growth Curve.

5.2.6.2.2 Case 2. How to Determine the MTBF for a Test Phase.

Objective: Determine the average MTBF M_i for the i-th test phase.

Given Conditions: The idealized growth curve

$$M(t) = \begin{cases} M_I & 0<t<t_1 \\ M_I\left(\frac{t}{t_1}\right)^{\alpha}(1-\alpha)^{-1} & t>t_1 \end{cases}$$

is given and the ends of test phases $t_1, t_2, \ldots, t_k$ are known.

<u>Solution for Case 2:</u> The average number of failures for the i-th test phase is determined by $H_i = N(t_i)-N(t_{i-1})$ where

$N(t_i) = \lambda_I \, t_1 \left(\frac{t_i}{t_1}\right)^{1-\alpha}$ and $\lambda_I = 1/M_I$. The average MTBF for the i-th test phase is given by $M_i = (t_i - t_{i-1})/H_i$.

<u>Example of Case 2:</u> In the example in Section 5.2.6.2.1 the first test phase was identified from 0 to 1000 hours. Suppose the program consists of four additional test phases at 1000-2500, 2500-5000, 5000-7000, and 7000-10000 hours. Determine the average MTBF's to be expected over these periods if reliability growth follows the idealized curve

$$M(t) = \begin{cases} 50 & 0<t<1000 \\ \dfrac{50}{.77}\left(\dfrac{t}{1000}\right)^{.23} & t>1000 \end{cases}$$

from the example in Section 5.2.6.2.1.

From the idealized growth curve the parameters are $\lambda_I = .02$ and $\alpha = .23$. Therefore, the average number of failures for the i-th test phase is $H_i = N(t_i)-N(t_{i-1})$ where

$$N(t_i) = .02(1000)\left(\frac{t_i}{1000}\right)^{.77} \text{, for}$$

$t_1 = 1000$, $t_2 = 2500$, $t_3 = 5000$, $t_4 = 7000$, $t_5 = 10000$.

The average number of failures H_i and the average MTBF M_i for each test phase are presented in the table below. The average MTBF's are plotted in Figure 5.28.

Phase i	H_i	$t_i - t_{i-1}$	M_i
1	20.0	1000	50
2	20.5	1500	73
3	28.6	2500	87
4	20.4	2000	98
5	28.3	3000	106

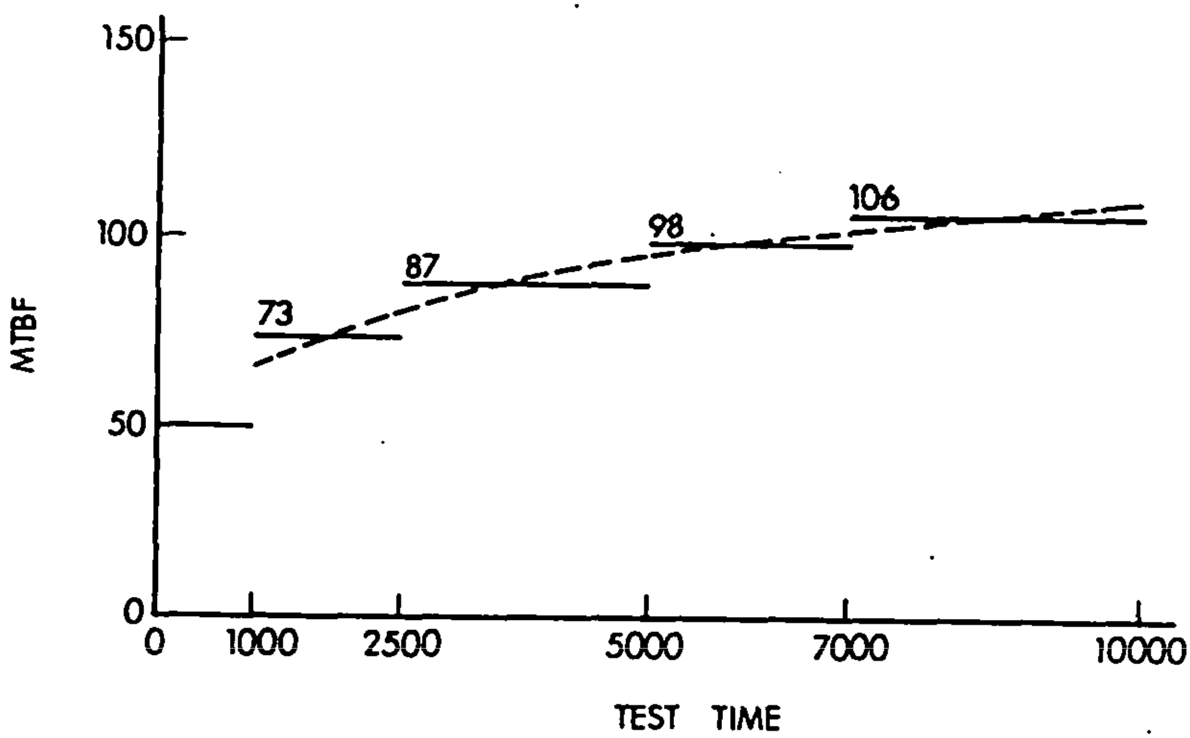


Figure 5.28 Example of Average MTBF's.

5.2.6.2.3 Case 3. How to Determine How Much Test Time is Needed.

Objective: Determine how much test time, T, is needed to attain a final MTBF of M_F.

Given Conditions: The first test phase is from 0 to t_1.

The initial average MTBF is M_I.

The growth parameter is α.

Solution for Case 3: The idealized growth curve at time t is

$$M(t) = M_I\left(\frac{t}{t_1}\right)^{\alpha}(1-\alpha)^{-1}.$$

Find T such that $M(T) = M_F$. That is, find T such that

$$\log T = \log t_1 + \frac{1}{\alpha}\left[\log \frac{M_F}{M_I} + \log (1-\alpha)\right].$$

Example of Case 3: The average MTBF over the first test phase of t_1 = 700 hours is estimated to be 1 hour. With a growth parameter of α = .4 how many test hours are needed to attain a goal of 3 hours MTBF?

From the above, the cumulative test time T necessary to grow from 1 hour MTBF to 3 hours MTBF must satisfy

$$\log T = \log 700 + \frac{1}{.4}\left[\log 3 + \log .6\right] = 8.02.$$

That is, T = 3043 hours.

5.2.7 Test Phase Reliability Growth. Based on the activities and objectives of the program, the reliability growth plan should indicate for each test phase the levels of reliability that are expected to be achieved. Specifically, for each test phase where an assessment will be made, the following points should be clearly expressed by the reliability program plan:

1. Whether the reliability will be held constant over the test phase or reliability growth is planned during the test, i.e., fixes will be introduced into the system during the test phase.

2. If it is planned to hold the reliability constant, then the level of reliability expected during the phase should be specified.

3. If reliability growth is planned during the test phase, then the reliability objective for the system on test at the end of the test phase should be specified.

4. If delayed fixes are planned at the end of the test phase, then the reliability objective for the beginning of the next test phase should be given.

In addressing the test phase reliability objectives it is useful to consider the effectiveness of the test and redesign efforts. A test phase of a given length can be expected to identify a certain number of failure modes. There are three responses that can be made to each identified failure mode:

a. Incorporate a design change during the test phase.

b. Incorporate a design change after the test phase.

c. Incorporate no design change.

5.2.7.1 Design Changes During the Test Phases (Test-fix-test). The planned growth curve should reflect the extent of design changes planned during each test phase; and, of course, implicit in this determination is the extent to which design changes are not planned. Historical information may be useful as well as engineering analysis methods described in Appendix A. The rate of growth during test phases is, of course, primarily dependent upon the extent of design changes that are planned.

5.2.7.2 Design Changes After the Test Phase (Test-find-test). The growth that takes place between test phases is the result of action taken on failure modes discovered during a previous test phase that is not incorporated until the end of the test phase. This growth cannot, however, be verified until some of the next phase of testing is accomplished. Figure 5.29 illustrates the effect of deferring redesign from the test phase to a separate redesign phase.

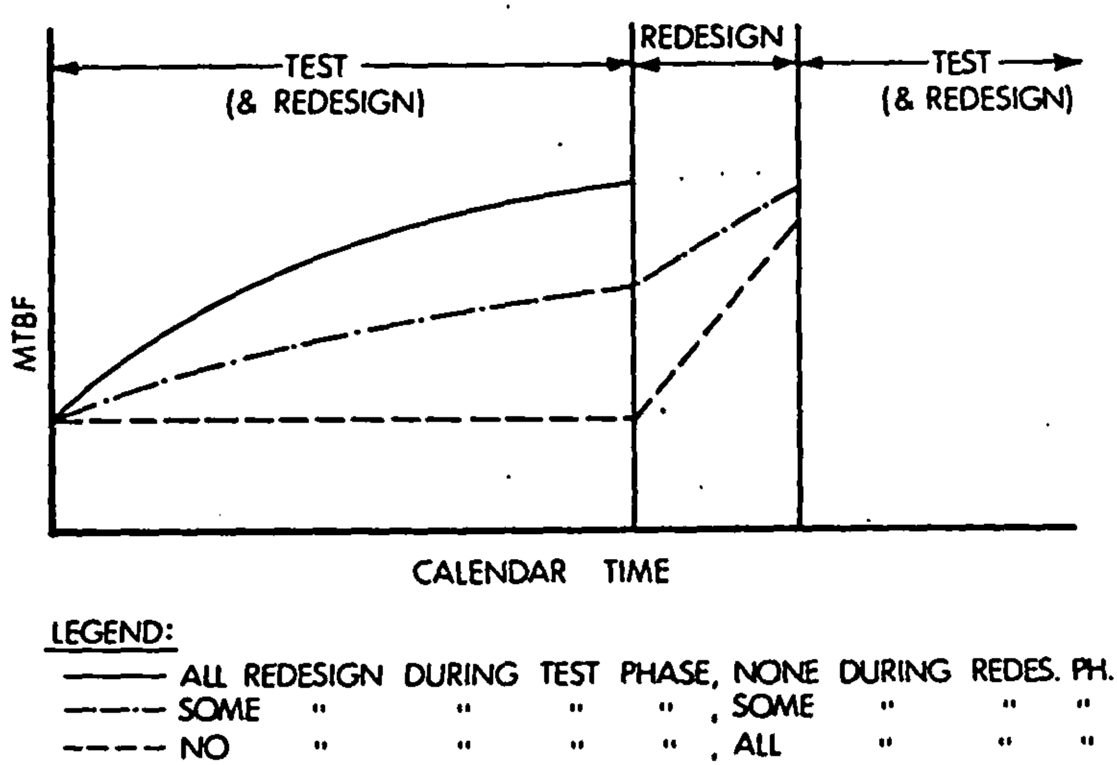


Figure 5.29 Effect of Deferring Redesign.

As more redesign is deferred, the effectiveness is reduced, because of the inability to detect ineffective design changes and newly introduced failure modes. Analytically, then, the redesign phase can be viewed as a delay of design changes that are identified during test, and some allowance should be made for the lesser effectiveness of delayed redesign. When working in terms of test time, a distinct redesign effort will be shown as a vertical jump, similar to that shown in Figure 5.30.2. It must be recognized, however, that a certain amount of calendar time is required to achieve the jump. This calendar time may be completely distinct from the calendar time used for testing, as illustrated in Figure 5.30.1, but more commonly, time constraints require that at least some of the time is concurrent with the previous test phase, as illustrated in Figure 5.30.2. Overlapping redesign and test in this fashion will tend to yield a less effective redesign, since it is started somewhat prematurely. A guide to quantifying the growth between test phases is the computation of the percentage jumps that have been historically observed on similar systems or equipments.

5.2.7.3 Incorporate no Design Change. There will be a certain percentage of failures for which no design changes will be made. There may be an inability to identify appropriate changes, or the identified changes may not be cost effective or may be too time-consuming to pursue.

5.2.8 Examples of Growth Curve Development. The following examples illustrate the development of planned growth curves for two systems.

5.2.8.1 Example of Growth Curve Development for a Fire Control System. The project manager for a fire control system wished to construct a planned growth curve while this system was still in the early stages of an accelerated, competitive development program. The growth curve was needed to assist in scheduling test phases for the program, to use as a reference for evaluating planned growth curves submitted by the competing contractors, and to serve as a baseline for tracking demonstrated reliability during development testing.

Given Conditions: Mission reliability requirements in the Decision Coordinating Papers called for 80 hours MTBF during Development Testing/Operational Testing (DT/OT), 110 hours MTBF during the Follow-on Evaluation (FOE), and 140 hours MTBF during the Initial Production Test (IPT). These reliability requirements were to be demonstrated by fixed configuration testing during the respective test phases. Each test phase was planned to last for 1100 hours. Preceding and following these formal test phases, the contractors were to perform an undetermined amount of inhouse testing and attempt design fixes of any problem failure modes that were discovered.

A mission reliability of 150 hours MTBF was required by the end of the first year of production. Unfortunately, some reliability

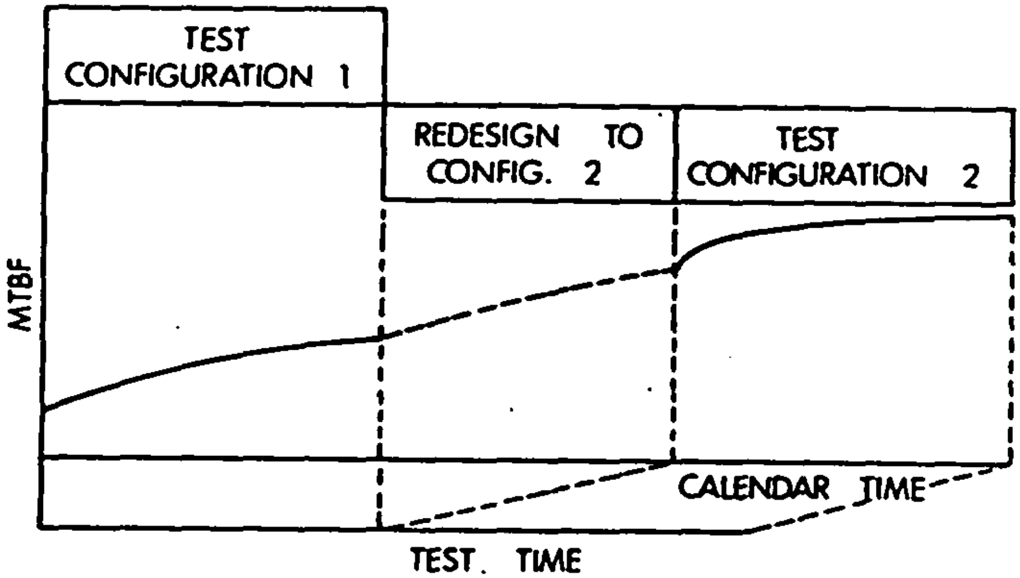


FIG. 5.30.1

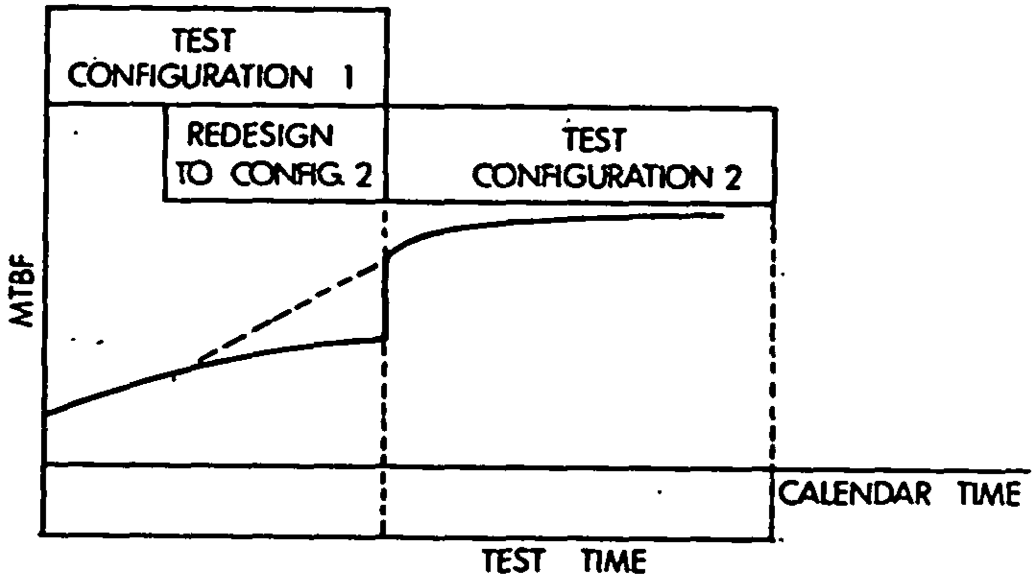


FIG. 5.30.2

Figure 5.30 Accounting for Calendar Time Required for Redesign.

growth had to be planned during the early phases of production. In this instance, however, some failure mode fixes on production items were considered necessary because of the accelerated nature of the program and the 10 months lead time required to implement a fix from the time of discovery of the problem failure mode.

Two further conditions on the development program were the limited number of fire control units and the limited amount of calendar time available for testing. These limitations necessitated a total test time for the formal test phases and contractor in-house testing of no more than 14,000 hours.

Problem

The basic problem of constructing a planned reliability growth curve for the fire control system required decisions about several parameters of the overall test program. The first decision to be made was how much total test time should be planned in order to achieve the final reliability requirement of 150 hours MTBF. Then it had to be decided when the test periods should be scheduled for the three test phases DT/OT, FOE, and IPT. The primary tool for making these decisions was to be the idealized reliability growth curve.

Construction of Idealized Curve

From the results of the initial development testing, it was projected that approximately 34 failures would occur during the first 1700 hours of testing. Since there was not enough calendar time to find, evaluate and fix any failure mode during this initial testing, the MTBF over this period was projected to be a constant equal to 1700/34 = 50.0 hours. Furthermore, it was known that 150 hours MTBF must eventually be achieved and that no more than 14,000 hours of test time was available. It was, therefore, of interest to know what kind of idealized curve would result if the maximum possible test time of 14,000 hours was utilized.

The conditions of this example correspond to the conditions given in Section 5.2.6.1.1 with $T = 14000$, $t_I = 1700$, $M_I = 50.0$, and $M_F = 150.0$. The growth parameter α is obtained by

$$\alpha = -\log\left(\frac{14000}{1700}\right) - 1 + \left(\left(1 + \log\left(\frac{14000}{1700}\right)\right)^2 + 2\log\left(\frac{150}{50}\right)\right)^{1/2} = 0.34.$$

An α value of 0.34 is only moderately high, but it is indicative of a relatively aggressive development program that would require management emphasis on the analysis and fixing of problem failure modes. Using a test time of less than 14,000 hours would result in a projected α greater than 0.34 and would therefore require an even more dynamic reliability growth program. Because such a shortened program would have an increased risk of not achieving the required reliability, the program planners for this fire control system decided to schedule the full 14,000 hours of test time for reliability growth effort. The idealized growth curve for this development program is shown in Figures 5.31 and Figure 5.32.

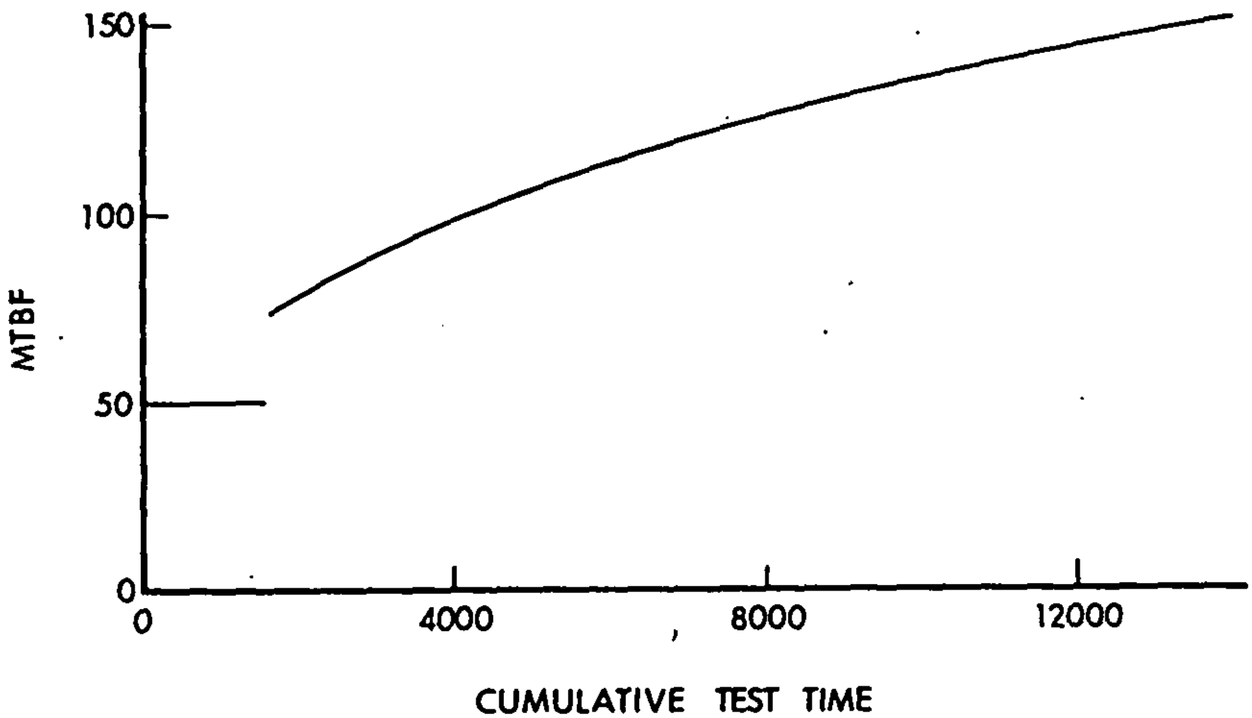


Figure 5.31 Idealized Growth Curve.

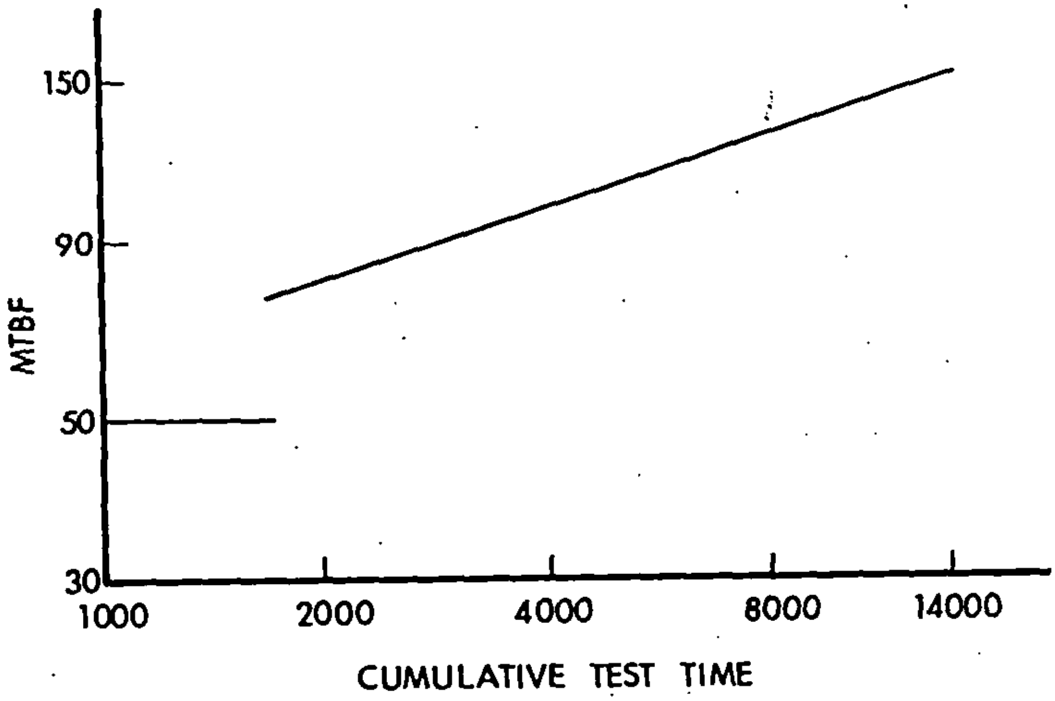


Figure 5.32 Idealized Growth Curve on Log-Log Scale.

Construction of Planned Curve.

Once the idealized curve had been constructed, it was used as a basis for developing a planned growth curve. The three test phases were to be scheduled in the testing program during periods when the corresponding reliability requirements could reasonably be expected to be achieved. An appropriate way of judging what average reliability could be demonstrated during a given test period was to utilize the information contained in the idealized growth curve. In Figure 5.31 the curve reaches 80 hours MTBF at 2100 hours of testing. It is clear, then, that over any test phase which begins at 2100 hours of cumulative test time, the average MTBF should equal or exceed 80 hours. Consequently, DT/OT was scheduled to begin at 2100 hours of cumulative test time.

By the same argument, the FOE was scheduled to begin at 5500 hours of cumulative test time, because the idealized curve in Figure 5.31 showed that the FOE requirement of 110 hours MTBF could be achieved in 5500 hours of testing. The beginning of IPT was scheduled in a similar manner. As stated in the given conditions, these three test phases were to last for 1100 hours each, and the fire control systems undergoing test were to remain in a fixed configuration throughout each test phase. This latter condition implied that the reliability during each test phase should be constant, and the planned growth curve should therefore show a constant reliability during these periods of testing.

After each test phase, the reliability was expected to be increased sharply by the incorporation of delayed fixes. In addition, testing was to be halted after 1700 hours of test time in order to incorporate design fixes into new system prototypes. The planned growth curve had to indicate jumps in reliability at each of these points in the test program. During the test time outside the formal test phases, steady reliability growth was planned because of continual fixing of problem failure modes. The resulting planned growth curve is shown in Figure 5.33.

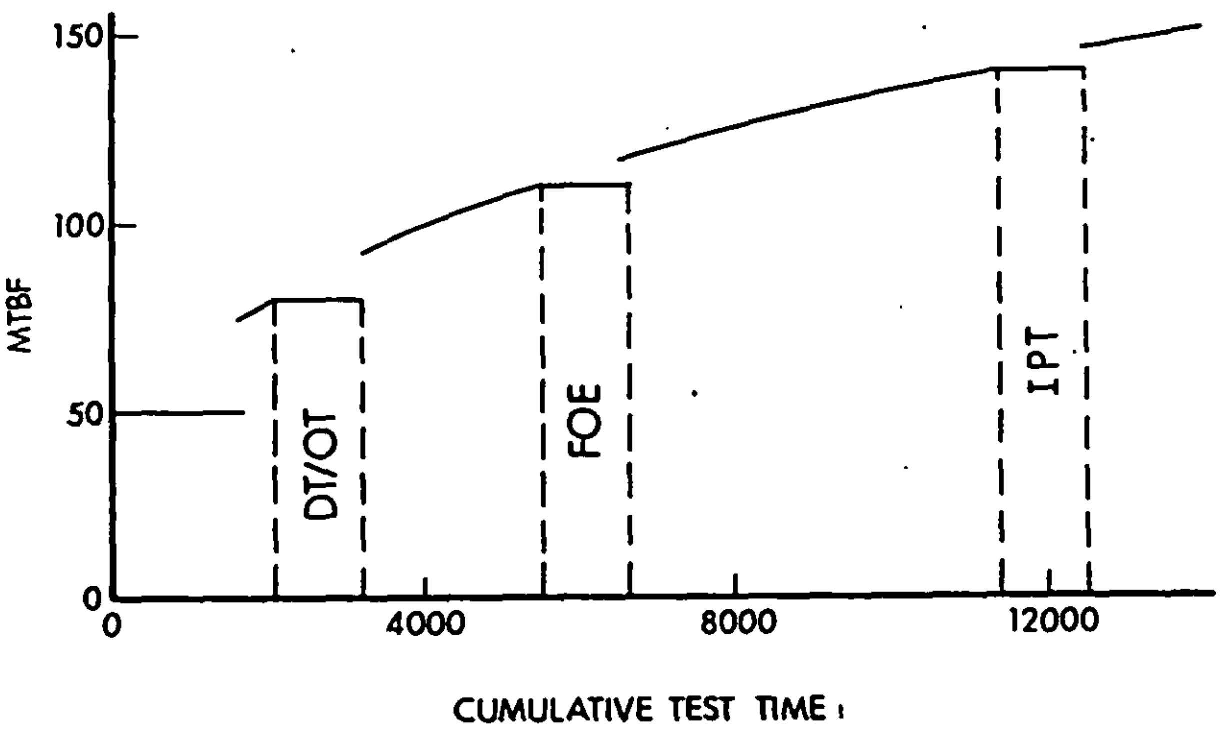


Figure 5.33 Planned Growth Curve.

This planned growth curve shows in graphic terms how the project manager plans to achieve the required mission reliability for the fire control system. The initial portion of the curve indicates how long the testing program should proceed before reliability growth begins and what average reliability is expected over this initial period of testing. The remainder of the curve indicates where in the development program reliability is expected to grow and where it is expected to remain constant. At points where there is a halt in testing and delayed fixes are incorporated, the curve shows how much increase in reliability is expected from the delayed fixes.

5.2.8.2 Example of Growth Curve Development for a Tank. The following discussion illustrates the iterative process employed in the development of the planned growth curve for a tank. Among the factors which were considered are: current policy and guidance, previous experience, program and test constraints, Duane's postulate, and the delay of the incorporation of fixes into actual hardware. The problems which arose, lessons learned, and the uses of the curve are also discussed.

Constraints Identified

During the development of the planned growth curve, several constraints were identified which had to be considered:

a. Overall program schedule,

b. Threshold requirements,

c. Impact to the total development cost,

d. Established test schedule,

e. Delays in incorporating design changes (fixes) into the hardware,

f. Previous experience on other Tank-Automotive hardware, and

g. US Army Materiel Development and Readiness Command policy.

Initial Planned Growth Curve

One of the first things considered was the influence that previous testing has on current testing. Figure 5.34 shows the testing to be accomplished during full scale engineering development (FSED). This testing was divided into five distinct phases. Each separate test was considered for possible impact on the MMBF (mean miles between failure) at the beginning of each phase. This influence is represented by the arrows. The three considerations given to each phase were:

(1) The delay in the introduction of fixes into hardware,

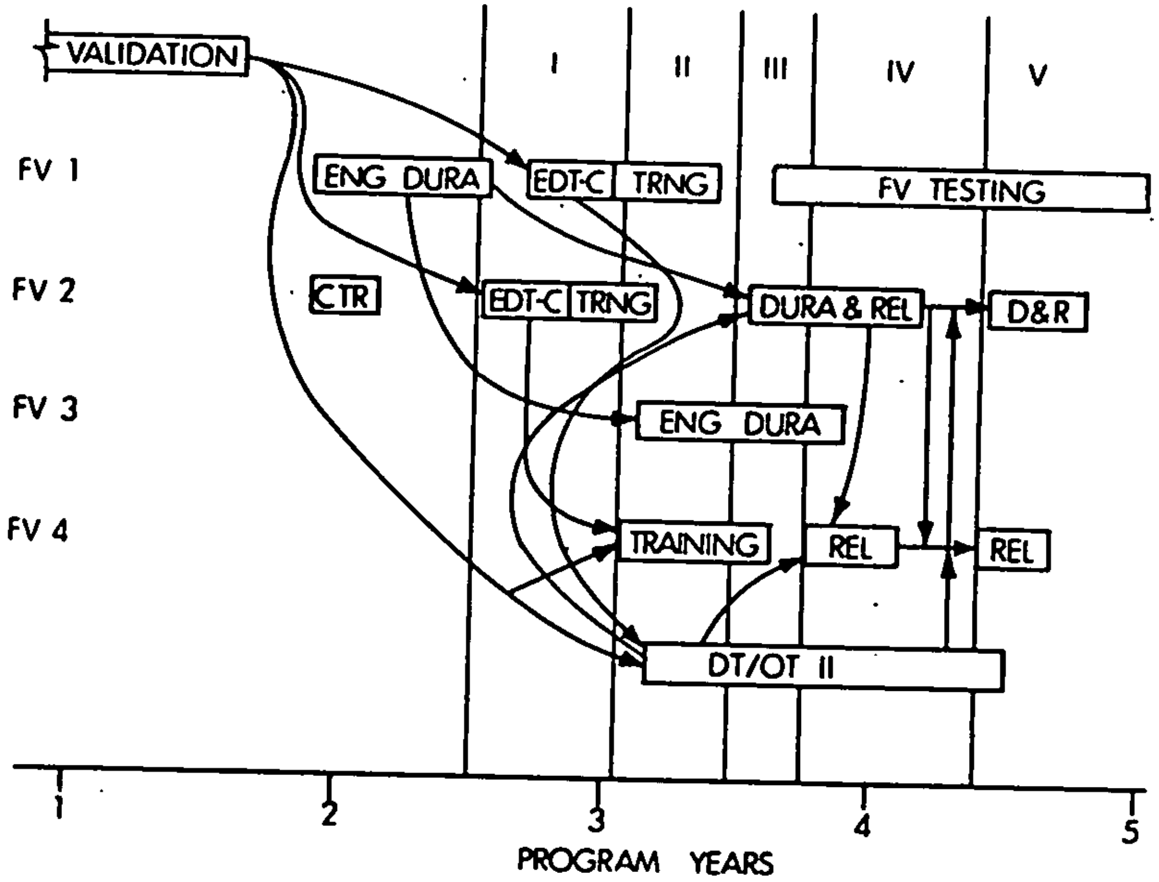


Figure 5.34 Phase Logic.

TABLE 5-I ABBREVIATIONS FOR FIGURE 5.34

FV	Facility Vehicle
Eng Dura	Engine Durability
EDT-C	Engineering Design Test - Contractor
TRNG	Training
CTR	Contractor
Dura & Rel	Durability and Reliability
D & R	Durability and Reliability
Rel	Reliability
DT/OT	Development Test/Operational Test

(2) The inability to measure immediately the impact of these fixes, and

(3) Previous experience (See Table 5-II).

Table 5-II shows some past testing experience with tank automotive hardware and lists the percent of the jump between one test phase and the next. As can be seen, this ranges from about 16% to 56%. The systems shown include engines, trucks, tractor-trailers and tanks.

A hard look at the approximately 10,000 miles of experimental prototype test (EPT) results was also made. The test incidents were divided into the following four categories:

(1) Eliminated (by QC or design),

(2) Nothing being done ("isolated case"),

(3) Redesign considered straghtforward and/or lead time short, and

(4) Redesign considered difficult and/or lead time long

Those incidents which fall into categories 1 and 3 were considered to have an influence on early FSED testing.

TABLE 5-II TEST DEMONSTRATION R JUMPS

System	% Jump From First Test To Second Test	
GOER	16%	(950/820 MMBF)
RISE ENGINE	17%	(6 VS 7 FAILURES)
M274A5	18%	(580/490)
M561	33%	(160/120)
HET	38%	(1120/810)
M551	40%	(700/500)
M60A1E3	56%	(140/90)
(PHASE I, PHASE II)		

The EPT results and the above concepts were used to develop the initial planned growth curve shown in Figure 5.35.

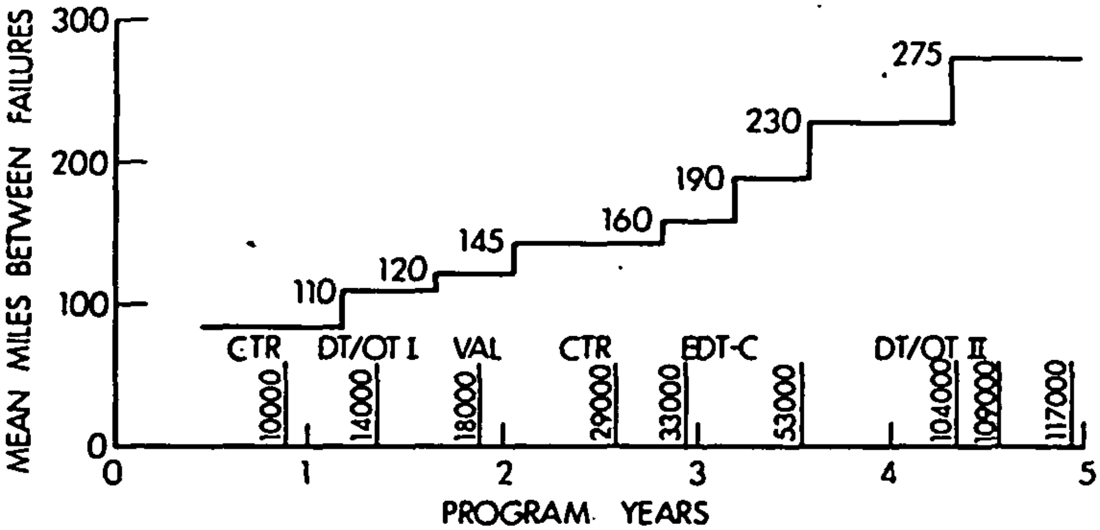


Figure 5.35 Initial PMO Planned Growth.

During periods of no test activity or when testing the same hardware, no growth is planned; however, experience gained can be passed on to succeeding phases (shown later as a jump in MMBF).

The idealized growth curve was checked with a log-log plot (see Figure 5.36). The beginning and ending points were connected and points plotted which correspond to the planned MMBF at the end of each phase. The plotted points fell close to the fitted line which indicates that the idealized growth curve corresponding to this planned curve follows a log-log learning curve pattern. (See Duane's postulate, Appendix B, and Section 5.2.6).

Revised Planned Growth Curve

Because the same hardware was to be used for the first eight months of FSED, the contractor pointed out that no growth would become evident during this period. At the end of this eight month period the test vehicles, however, would be refurbished and would contain several design changes. The contractor and program manager's office (PMO) also agreed that for future planning purposes, the 1500 miles scheduled maintenance periods during development test/operation test (DT/OT II) would be used to incorporate changes into the vehicles. Since most of the tanks would complete 6,000 miles during DT/OT II, three jumps could be expected during actual testing.

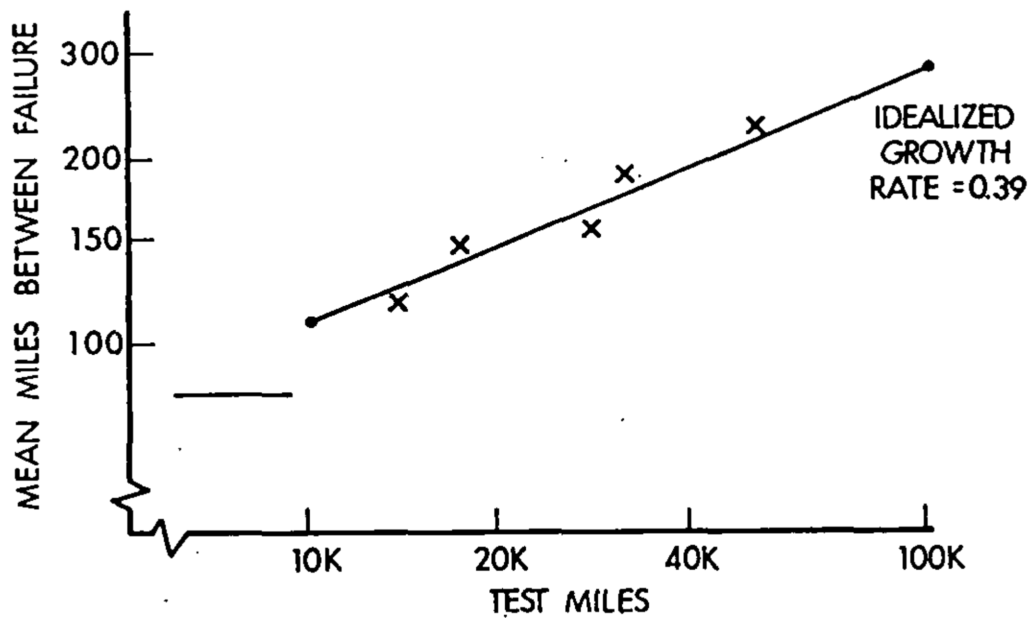


Figure 5.36 Log-Log Plot of Initial PMO Idealized Growth Curve.

The PMO also negotiated for an extended durability program contract which was to develop the engine to a high degree of maturity before full scale production. Slight changes were made to the overall test schedule as part of these negotiations. These schedule changes can be seen by comparing Figure 5.37 with Figure 5.35. The revised planned growth curve shows only a slight improvement between engineering development test (EDT) and the start of DT/OT II because the build-up of the DT/OT II pilots starts before EDT ends. Additionally, to allow for such things as schedule delays, this curve only shows two jumps instead of the expected three during DT/OT II. After DT/OT II, the reliability retest would permit another jump in MMBF. The final jump is planned to occur during DT/OT III.

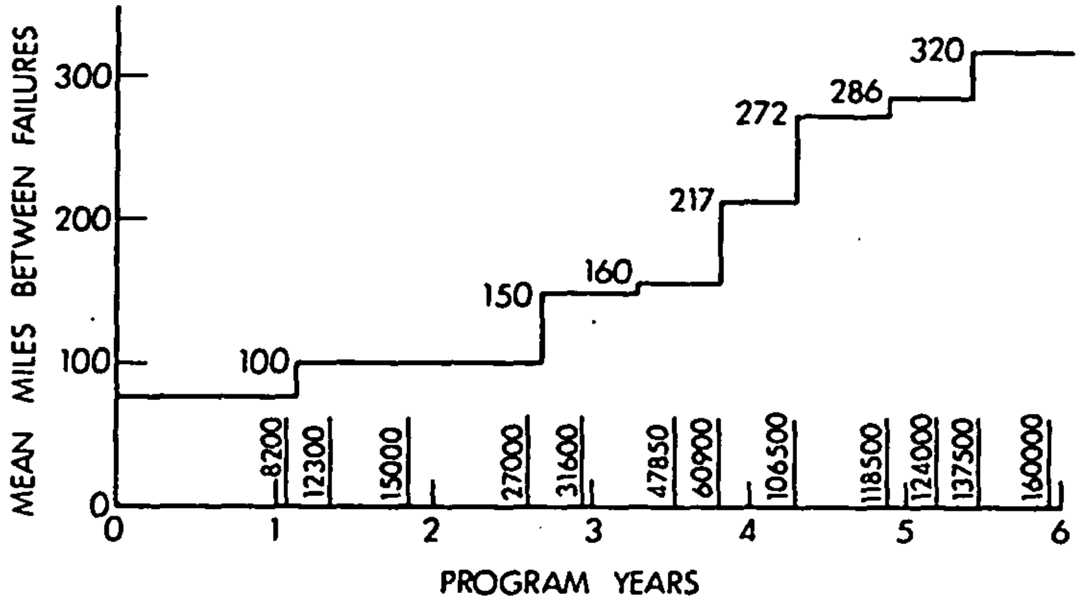


Figure 5.37 Revised PMO Planned Growth Curve.

The revised PMO growth curve was plotted on log-log paper in order to verify that the growth rate for the idealized curve was acceptable. (See Figure 5.38). The growth rate determined was 0.42, which was considered achievable in view of the previously mentioned studies.

Problems Uncovered/Lessons Learned

Throughout the development of the reliability growth curve, some problems were uncovered and also some lessons were learned which will be helpful in the future.

- Iterative process - There are many facets (and extensive negotiations) which have to be addressed throughout the process and the consequences of each step must be carefully considered. One must also keep abreast of all program changes for possible impact on reliability growth.
- Past experience - Data from both similar development programs and recent experience on the system currently being developed must be considered.

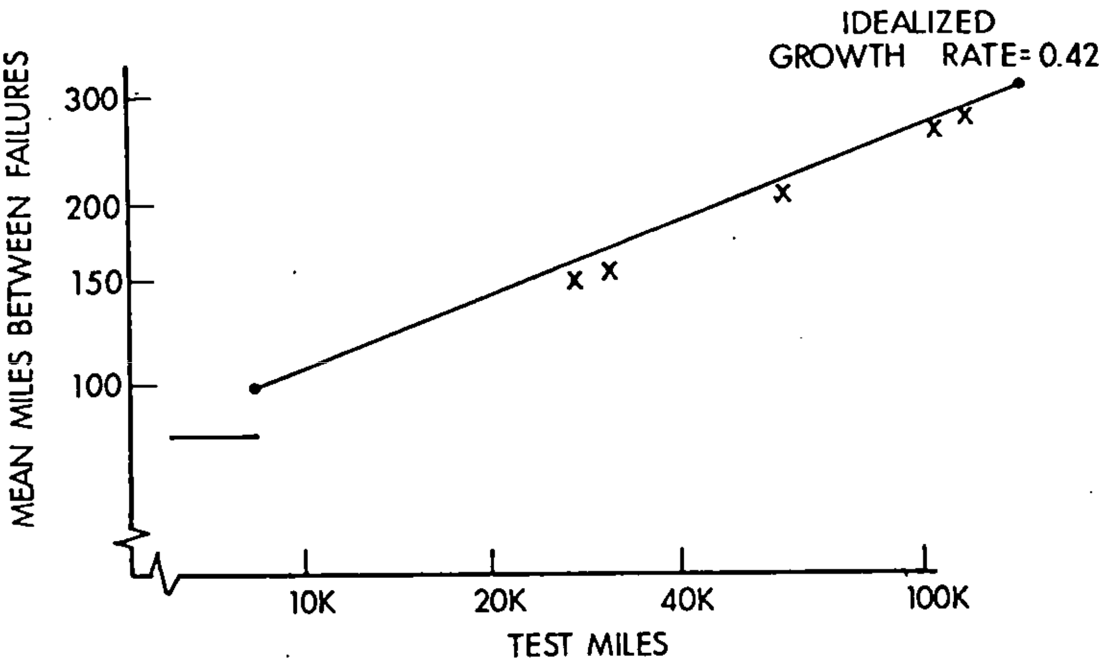


Figure 5.38 Log-Log Plot of Revised PMO Idealized Growth Curve.

• Predictions - If one is relying on previous engineering technology, then this should be reflected in the growth curves. For example, the growth curve should not arbitrarily begin at 10% of the predicted production MMBF.

• Imagination required - The approach to growth must be carefully thought out. It should consider everything that is available, but not rely heavily on any single idea. The portrayal of growth may require novel techniques.

• One Planned Growth Curve - There is in fact only one planned curve for the program which should be agreed to by the PMO and the contractor since both are developing the same hardware.

• Realism - The real world influence of hardware introduction, contractual constraints, contractor apprehension, previous experience, among other things, must be considered.

• Proposal Evaluation - An important lesson was learned here. During contractor proposal evaluation, the growth concepts must be studied very carefully. On the other hand, there is a practical limit on the amount of time that can be spent because of the multitude of activities being performed during evaluation in a relatively short time.

Impact/Uses of Reliability Growth Curve

For this system, the principal use for the reliability growth curve will be to keep the contractor's reliability program aggressive. The contractor has developed a comprehensive method which should assure achievement of an adequate reliability growth rate. The essential elements of this method are:

(1) Product Assurance (PA) Manager sign-off authority on original designs and all design changes.

(2) PA Manager approval of failure analyses and close-out actions on all test incidents as part of the closed loop reporting system.

(3) Reliability Program Plan requirements in all major subcontracts.

The PMO did not want to get into the position of presenting a considerable amount of "paper fixes" in order to "demonstrate" achievement of requirements. One way this will be accomplished is by the periodic fee award reviews. If the contractor is not aggressively getting fixes into the hardware for verification, the PMO can instigate action to set lower than maximum permitted award fee. A second way will be the award of the production contract. Again, the PMO will recommend award based on performance. The third and most immediate way will be through monthly and bi-monthly meetings, reviews, and reports. If the contractor is not performing satisfactorily, this fact will be addressed through these media where it will most assuredly receive the attention of both contractor and PMO top management.

5.3 Reliability Growth Tracking.

This section discusses some of the basic concepts associated with tracking reliability growth during development testing. In general, tracking reliability growth is not simple, and the methods will often have to be tailored for the particular problem under consideration. Moreover, a thorough knowledge of the system and program is necessary to insure that the data analysis is conducted in a manner compatible with the program activities.

The objectives of growth tracking usually include:

a. Determining if growth is occurring and to what degree,

b. Estimating the demonstrated reliability, and

c. Making a projection of the reliability expected at some future time.

The methods discussed in this section are directed toward reliability growth tracking within a major test phase. See Sections 5.1.6 and 5.1.7.

5.3.1 Tracking Within Major Test Phases. As noted previously in Section 5.1, the testing from one test phase to another will often be under different environmental conditions and involve prototypes that may differ substantially in design due to the incorporation of delayed fixes between test phases. Also, the emphasis on reliability growth may vary among the major test phases. These factors will generally affect the level of reliability and the growth rate within each test phase. Therefore, it is usually advantageous to track the reliability growth entirely within each of the major test phases and address demonstrated and projected reliability values.

During a major test phase, the test environment may not be consistent with the environment that the goals are based on. When tracking reliability growth, this should be taken into account with, perhaps, an adjustment in the observed failure times or an adjustment in the milestones for the test phase. This may, for example, involve the application of k factors. (See Section 5.4)

5.3.2 Demonstrated and Projected Reliability Values. It is important that management have realistic demonstrated and projected reliability assessments for the system during development testing. The demonstrated value provides a reliability estimate for the system configuration on test at the end of a phase. This value is determined from an analysis of the actual test results. A projected reliability value estimates the system reliability expected at some future point. A projection can account for the effect of fixes that either may have been introduced into the system very late in the test and as such have not been fully reflected in the test data, or for fixes that are upcoming, such as delayed fixes between test phases.

The demonstrated reliability value may be calculated utilizing the following techniques: (1) reliability growth analysis or (2) engineering analysis. When appropriate, the reliability growth analysis is the preferred method since it provides an objective mathematical assessment of the reliability of the system being tested. The reliability growth method measures the effect that individual fixes that have been incorporated during testing have had on system reliability, and provides credit for fixes in the determined reliability value, if they have in fact proven to be effective. It should be noted that if no fixes are incorporated during testing, then reliability growth procedures would not be necessary and the demonstrated reliability value would be determined by dividing the total test time/miles/rounds, etc., by the number of charged failures. In addition, if a situation arises where the reliability growth procedures cannot be applied to the test data because of data anomalies, then a method known as engineering analysis may be used for determining the demonstrated reliability. The engineering analysis method is subjective, and will, therefore, tend to be less definitive than a data analysis based on reliability growth procedures.

The engineering analysis technique involves using engineering judgment to assess the effectiveness of fixes that have been incorporated during the test program in determining the demonstrated reliability value. In this method, the status of charged failures will be evaluated to determine if their chargeability is changed based on the effectiveness of fixes introduced during the test program. For the chargeability of a failure to change there must be concrete evidence based on test data that the failure rate has been reduced in the operational environment and that it does not create any new failure modes. Useful criteria for determining if the chargeability of a failure has changed are the following:

(a) failure analysis adequacy,

(b) appropriateness of corrective action,

(c) demonstration of corrective action,

(d) verification of effectiveness of corrective action, and

(e) verification of future implementation of corrective action

If the above has been satisfied, i.e., concrete evidence has been presented that a failure mode has been partially or completely eliminated, then the chargeable status of a failure(s) may be changed and a demonstrated value based on revised failure rates for these failure modes may be computed. In most cases a fix will not completely remove a failure mode from the system. If the rate of occurrence of a particular mode has been reduced to a lower rate, but the mode has not been eliminated, then the failure rate estimate for the mode should be adjusted accordingly to a lower value, but not reduced to zero. (See Section 5.4.1.3.5). The adjustments made to the failure rates should be based on fixes that have been verified by test (component, subsystem, and/or system) as effective. If the effectiveness of a fix cannot be verified by test, then any subjective evaluation of the impact of the fix on reliability should be reflected in a projected value but not a demonstrated value.

A projected reliability value is a particularly important consideration when the demonstrated value determined at the end of a phase is found to be below an intermediate requirement and as a result management is very much interested in determining if the final requirements will be achieved with the current test program. In the situation where the demonstrated reliability value is meeting intermediate requirements, it still may be desired to compute a projected value to provide an indication of the system reliability at some future point in the test program.

A projected reliability value may be calculated by extrapolating a growth model or by assessing the impact of fixes (those introduced late in the test phase or those introduced after the end of the test phase). The extrapolation of a growth model to obtain a projected value would normally be conducted if the test-fix-test level of effort in the ensuing test program is going to be about the same as in the past. However, if a substantial amount of fixes are going to be made before the next test phase or if the test-fix-test level of effort is going to increase substantially because the demonstrated reliability estimate is considerably below current thresholds, then the extrapolation technique is unsuitable and an analysis of the impact of the fixes on reliability should be the method utilized.

In determining the impact of fixes on reliability, a number of methods may be employed. These include: (1) determine the effect that previous fixes have had on system reliability and use the fix effectiveness rate determined for assessing the impact of future fixes, (2) from similar systems (e.g., other missiles, tracked vehicles, etc.) establish the fix effectiveness rate and apply this rate to the system under test or (3) assume a varying fix effectiveness rate (e.g., 25%, 50%, 75%) and determine the projected reliability estimate utilizing these rates. It should be emphasized that not all fixes will be effective to the same degree. Some fixes will almost entirely eliminate a failure mode (relatively rare), other fixes will reduce the rate of occurrence (but not to zero), and some fixes may introduce other new failure modes.

5.3.3 Data. If reliability growth is occurring, this will be reflected in the fact that the intervals between successive failure times are tending to increase as development testing continues. Similarly, if negative growth is occurring, these intervals will be getting, on the average, smaller. For no growth the intervals can be expected to be, on the average, the same length. Therefore, to measure the growth trend, early failures as well as late failures are needed for comparison. In general, all failure times, in their chronological order (even those with fixes incorporated into the hardware) are needed for evaluating reliability growth. There should be no purging of the data. (See Section 5.4.) The estimation of the growth rate and system reliability will usually involve the utilization of a reliability growth model.

Clearly, the data must be consistent with the failure definition under consideration. Reliability growth, of course, can be evaluated regardless of the type of failure being evaluated (e.g., mission failures, system failures, etc.). Also major subsystems as well as the entire systems can be tracked. It is often useful to track the individual prototypes separately to determine if any significant differences in performance exist. For example, it is not unusual for some prototypes to receive fixes well before other prototypes or for prototypes to exhibit different failure rates as a result of the test time on the prototypes (e.g., a new prototype might be exhibiting failures associated with burn-in while a prototype that has been under test for an extensive time may no longer be exhibiting burn-in failures).

In general, time to failure data are preferred over data in which the time of each failure is unknown and all that is known is the number of failures that occurred in each period of time (grouped data). Time to failure data will obviously provide more information for estimating system reliability and growth rates.

5.3.4 Data Plots. A plot of the data for the test phase is an initial and basic step in reliability growth analysis. A plot of the data will often indicate a trend, if one exists. Of interest also is whether or not any major jumps in reliability have occurred or if there is a change in the growth rate which may be caused, for example, by different test conditions, the introduction of a new system with a high initial failure rate, or the incorporation of delayed fixes.

A simple plotting method is to calculate average failure rates over the cumulative time on test T. To construct an average failure rate plot, partition the cumulative time T into K subintervals with lengths, $T_1, T_2, \ldots, T_K$. If N_i is the number of failures in the i-th subinterval, then $\lambda_i = N_i/T_i$ is an estimate of the average failure rate over this subinterval. If growth is occurring, then the λ_i's should tend to decrease. If there is a major jump in the reliability due to a design change, this would be reflected in a large difference in an adjoining pair of λ_i's, if sufficient data exist.

Example 1. Consider the following N=46 failure times recorded for a system during a test period of T=3000 hours; 2.4, 24.9, 52.5, 53.4, 54.7, 57.2, 118.6, 140.2, 185.0, 207.6, 293.9, 322.3, 365.9, 366.8, 544.8, 616.8, 627.5, 646.8, 664.0, 738.1, 764.7, 765.1, 779.6, 799.9, 852.9, 1116.3, 1161.1, 1257.1, 1276.3, 1308.9, 1340.3, 1437.3, 1482.0, 1489.9, 1715.1, 1828.9, 1971.5, 2303.4, 2429.7, 2457.4, 2535.2, 2609.9, 2674.2, 2704.8, 2849.6, 2923.5.

We choose to partition the test interval into six subintervals, each of length 500 hours. There were 14 failures in the interval 0-500, 11 failures in the interval 500-1000, 9 failures in the interval 1000-1500, 3 failures in the interval 1500-2000, 3 failures in the interval 2000-2500, and 6 failures in the interval 2500-3000. The average failure rate is the number of failures in each interval divided by 500, the length of the interval. The average failure rates, therefore, for these six intervals are: .028, .022, .018, .006, .006, .012. These are plotted in Figure 5.39, and clearly indicate reliability growth. The number and length of the intervals are, of course, arbitrary, but should be chosen small enough to reflect a trend in failure rate, but large enough to smooth the data.

5.3.5 Statistical Tests for Trend. A plot of the data will usually indicate whether there is no growth (a constant failure rate) or reliability growth (positive or negative). This, however, can be tested statistically. There are a number of tests which can be used to test the null hypothesis of a constant failure rate. The following statistic, which can be used to test this null hypothesis, is sensitive to the alternative hypothesis that growth is occurring according to a learning curve pattern. The test is based on the AMSAA model, which is useful for tracking reliability growth within a test phase.

5.3.5.1 Time Truncated Test. During a test period T suppose that N failures were recorded at times $X_1 < X_2 < \ldots < X_N < T$. The test statistic is

$$\chi^2_{2N} = \frac{2N}{\hat{\beta}}, \text{ where}$$

$$\hat{\beta} = \frac{N}{\sum_{i=1}^{N} \ln(\frac{T}{X_i})}$$

Under the null hypothesis of exponential times between failure (no growth), χ^2_{2N} has a chi-square distribution with 2N degrees of freedom.

The statistic $\hat{\beta}$ estimates the growth parameter β. In the case of no growth β is equal to 1. For reliability growth $\beta < 1$, and negative growth $\beta > 1$. For large or small values of χ^2_{2N}, the null hypothesis of no growth is rejected.

Example 2. For the data in Example 1, $\hat{\beta}$ is .616, indicating reliability growth. To test the null hypothesis of no growth, the statistic χ^2_{2N} can be used. Under the null hypothesis, this statistic is chi-square with 2N = 92 degrees of freedom. At the 10 percent significance level, the appropriate critical values, found in a table of chi-square percentiles for 92 d.f., are CV1 = 70.9, CV2 = 115.4. The test statistic is $\chi^2_{92} = 149.3$. Since $\chi^2_{92} >$ CV2, the null hypothesis of no growth is rejected at the 10 percent significant level. Since $\hat{\beta} < 1$, and the null hypothesis is rejected, there is strong evidence of reliability growth.

5.3.5.2 Failure Truncated Test. If the data are failure truncated at X_N instead of time truncated at T, then the test statistic is

$$\chi^2_{2(N-1)} = \frac{2N}{\hat{\beta}}, \text{ where } \hat{\beta} = \frac{N}{\sum_{i=1}^{N-1} \ln(\frac{X_N}{X_i})} .$$

This statistic is chi-square distributed with 2(N-1) d.f. when the null hypothesis is true.

5.3.5.3 Grouped Data. There is also a chi-square test for trend which does not require that actual failure times be known. Divide the test time into K intervals with lengths $T_1, T_2, \ldots, T_K$ in such a way that $NT_i/T > 5$ for $i=1, 2, \ldots, K$. Let N_i be the number of failures in the ith interval. For this type of data the statistic

$$\chi^2_{(K-1)} = \sum_{i=1}^{K} \frac{(N_i - NP_i)^2}{NP_i},$$

where $P_i = T_i/T$, $N = \sum_{i=1}^{K} N_i$, $T = \sum_{i=1}^{K} T_i$, is approximately chi-square distributed with K-1 degrees of freedom, when the null hypothesis of exponential times between failure is true. The lengths $T_1, T_2, \ldots, T_K$ of the K intervals do not have to be equal to apply this test statistic, but the requirement that $NT_i/T > 5$ for $i=1,2,\ldots, K$ is recommended,. The null hypothesis is accepted for small values of $\chi^2_{(K-1)}$, and rejected for large values of $\chi^2_{(K-1)}$.

Example 3. Consider again the data in Examples 1 and 2. Since there are 46 failures altogether, no more than 9 intervals should be used. The total test time T = 3000 hours which suggests intervals of about 325 hours. Let us make $T_1 = T_2 = \ldots = T_8 = 330$ hours and $T_9 = 360$ hours, so that $NP_1 = NP_2 = \ldots = NP_8 = 5.06$ and $NP_9 = 5.52$. The numbers of failures, N_i, $i=1,2,\ldots,9$ are respectively: 12,6,7,5,4,3,1,4,4. For this example the statistic $\sum_{i=1}^{9} \frac{(N_i - NP_i)^2}{NP_i}$ has approximately a chi-square distribution with eight degrees of freedom under the hypothesis of exponential times between failure. The observed value of the statistic for these data is 15.4 and the critical value at the 0.10 level of significance is 13.3. Since 15.4 > 13.3 we reject the hypothesis of exponential times between failure.

Suppose we use 6 intervals of length 500 hours each. In this case $NP_1 = NP_2 = \ldots = NP_6 = 7.67$ and the observed frequencies are respectively: 14,11,9,3,3,6. The value of the chi-square statistic is 12.5 and the critical value at the 0.10 level of significance is 9.2. Once again since 12.5 > 9.2 we reject the hypothesis of exponential times between failures.

5.3.6 Fitting Growth Models to Data. Appendix B is a discussion of reliability growth models and should serve as a useful guide in selecting a particular model for application. Generally speaking, the simplest model which is realistic, and justifiable from previous experience, engineering considerations, goodness of fit, etc., will probably be a good choice.

In many cases the data may suggest a model or an approach. For example, if the cumulative failure rate plots linearly on log-log scale within a test phase, then this may suggest the AMSAA (Army Materiel Systems Analysis Activity) model. The model chosen should always be compatible with the average failure rate plots. If the average failure rates indicate, for example, that a particular design change resulted in a significant jump in the MTBF, then this may suggest that a single smooth model fitted to the data may not be realistic. An alternative approach may be to use the data prior to the design change to fit a curve and to use the data after the design change to fit another curve.

If a goodness of fit test is available for the model chosen, it should be applied to determine statistically if the model is justified. If a model is rejected by a goodness of fit test, the next step may not be to select another model but instead to examine the data to rationalize why the model did not fit. A significant jump in the MTBF may be a possible reason or a change in the reliability growth trend may be another possibility.

Example 4. In this example the AMSAA reliability growth model is fitted to the data of Example 1. For the AMSAA model a goodness of fit test exists to test if the model and data are compatible. Under this model the failure rate is given by $r(t) = \lambda\beta t^{\beta-1}$ where t is cumulative test time. The MTBF is then expressed as $m(t) = [r(t)]^{-1}$.

In example 2 the estimate of β was calculated to be $\hat{\beta} = .616$. From Appendix C, the estimate of λ is $\hat{\lambda} = N/T^{\hat{\beta}} = .332$. The failure rate at time t is then estimated as $\hat{r}(t) = \hat{\lambda}\hat{\beta}t^{\hat{\beta}-1}$. For instance, the failure rate estimate at 3000 hours is .009. In Figure 5.39, the failure rate function $\hat{r}(t) = \hat{\lambda}\hat{\beta}t^{\hat{\beta}-1}$ is plotted with the average failure rates given in Example 1. One can see that the plot and curve are compatible.

The MTBF function $\hat{m}(t) = [\hat{r}(t)]^{-1}$ is plotted in Figure 5.40. At 3000 hours the current MTBF is estimated by $[.0094]^{-1} = 106$.

The Cramer-von Mises statistic, given in Appendix C can be used to test if this model is compatible with the data. This statistic is indexed as m=N, for time truncated data and is expressed as

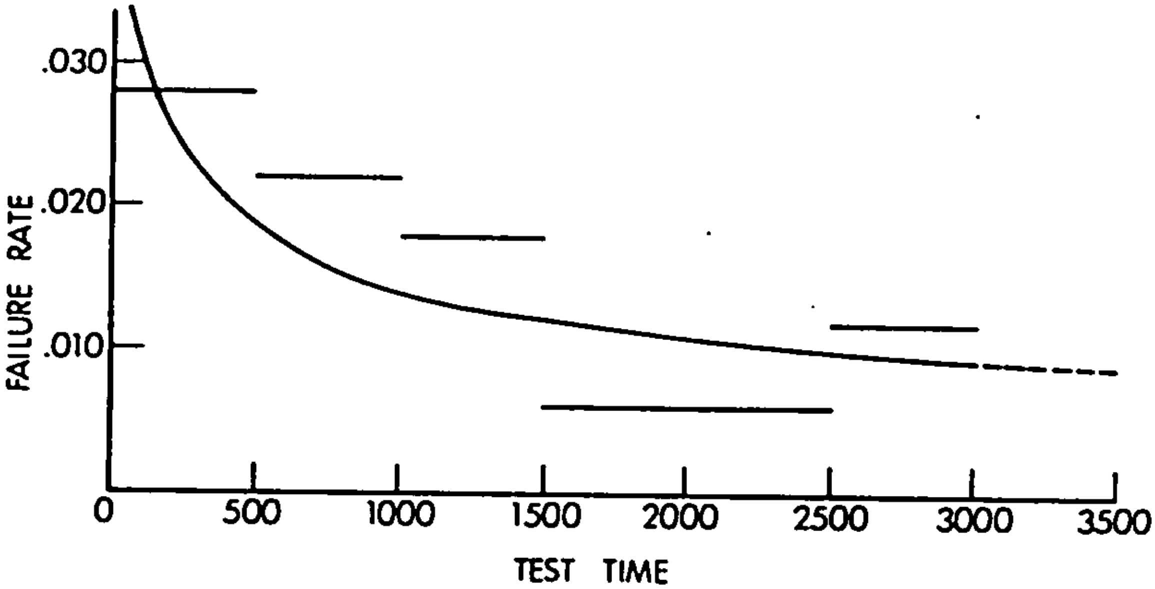


Figure 5.39 Estimated Failure Rate Function with Average Failure Rate Plots.

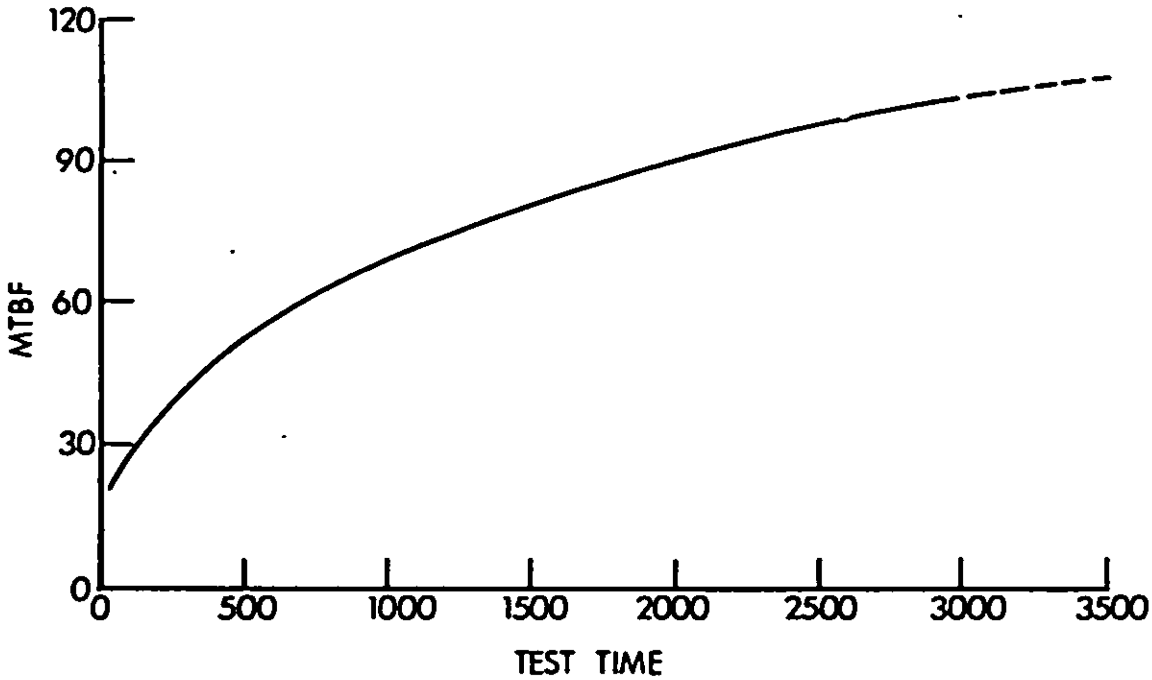


Figure 5.40 Estimated MTBF Function.

$$C_m^2 = \frac{1}{12m} + \sum_{j=1}^{m} \left[\left(\frac{X_j}{T} \right)^{\bar{\beta}} - \frac{2j-1}{2m} \right]^2$$

where $\bar{\beta} = \frac{N-1}{N} \hat{\beta}$. A table of critical value for this statistic is also given in Appendix C. The null hypothesis that the data follow the AMSAA model is accepted for small values of C_m^2 and rejected for large values.

For the data considered in this example C_m^2 is calculated to be .043. At the 10 percent significance level the appropriate critical value for m=46 is .172. Since .043 < .172, the AMSAA model is accepted as being compatible with the data.

5.3.7 Tracking One-shot Systems. Continuous growth models can be used as a good approximation for tracking the reliability of one-shot systems, provided the number of trials within each test phase is relatively large and the reliability relatively high. If these conditions are not met, discrete models may be required. Since these models have not been sufficiently evaluated in regard to their application and properties, no guidance on their use will be given.

Example 5. The following example discusses a reliability growth study of a missile system conducted by the Army. The purpose of the study was to use historic data on the first 801 valid flight tests to determine the growth curve and also to ascertain in retrospect how these data could have been used to track and project system reliability during development.

This exercise involved looking back on time and making predictions. Although all the data exist to confirm these predictions, this case history shows that a program manager can determine from test data the current system reliability status, estimate the rate of growth and obtain projections of future expected reliability. In this manner he may evaluate the system throughout the program to determine whether or not the reliability is growing at a sufficient rate to meet the required goals and allocate available resources accordingly.

The system is defined as the round less the warhead. The data included flight results from firings of successive designs for the round, starting with the R&D program and progressing to limited production. The format of the data routinely received identified the missile flights by round serial number and date of firings. Each flight was evaluated using the equipment scoring criteria established by the missile scoring criteria committee. This evaluation placed each flight attempt into one of five categories:

a. The Missile and Tracker were reliable,

b. The Missile was not reliable,

c. The Tracker was not reliable,

d. The flight was not a proper test of reliability, or

e. The information on the flight was insufficient to determine reliability.

Because interest was in tracking the growth of the round only, the first two categories were used as a basis for estimating the reliability growth. Therefore, the rounds which were scored either c, d, or e were not used in this study. This resulted in a data base of 801 valid flights.

In reliability growth considerations, it is configuration changes on the system which are of prime importance. Consequently in this study, these 801 valid flights were chronologically ordered by date of manufacture. Since the valid flights were identified as either success or failure, and they are ordered according to manufacturing date, this should reflect the sequence and consequence of system change during development. In this form, the data provided an acceptable base for reliability growth evaluation. The AMSAA reliability growth model was used in this study to track reliability since the number of trials was large. For application to discrete data, it was assumed that the failure probability for the i-th missile produced is $f_i = \lambda\beta i^{\beta-1}$.

The first step in this analysis was to estimate the average failure probability for 100 flight intervals. These are shown as horizontal lines in Figure 5.41. Each horizontal line is obtained by dividing the number of failures by 100 flights. The data were plotted to gain some insight in the form of the relationship of the data of manufacture (flight number in this case) and number of failures. These failure rate plots are useful for visualizing the system failure rate trend.

Using the failure results for the 801 flights and the estimation procedures given in Appendix C, the failure probability curve based on the AMSAA model was determined. This is shown in Figure 5.42 with the average failure rate plots. The goodness of fit statistic was then calculated to determine if the estimated failure probability curve and data were compatible. The value of the statistic was highly significant (i.e., very large) indicating that the curve did not reasonably represent the data. This is evident from the large discrepancy between the actual data and the fitted curve. The way the data fall either on one side or the other of the fitted curve indicates that there appears to be two distinct groups of data; i.e., the first two hundred rounds and the remaining group. This implies that a single, smooth, failure probability curve would not reflect the reliability growth of this system.

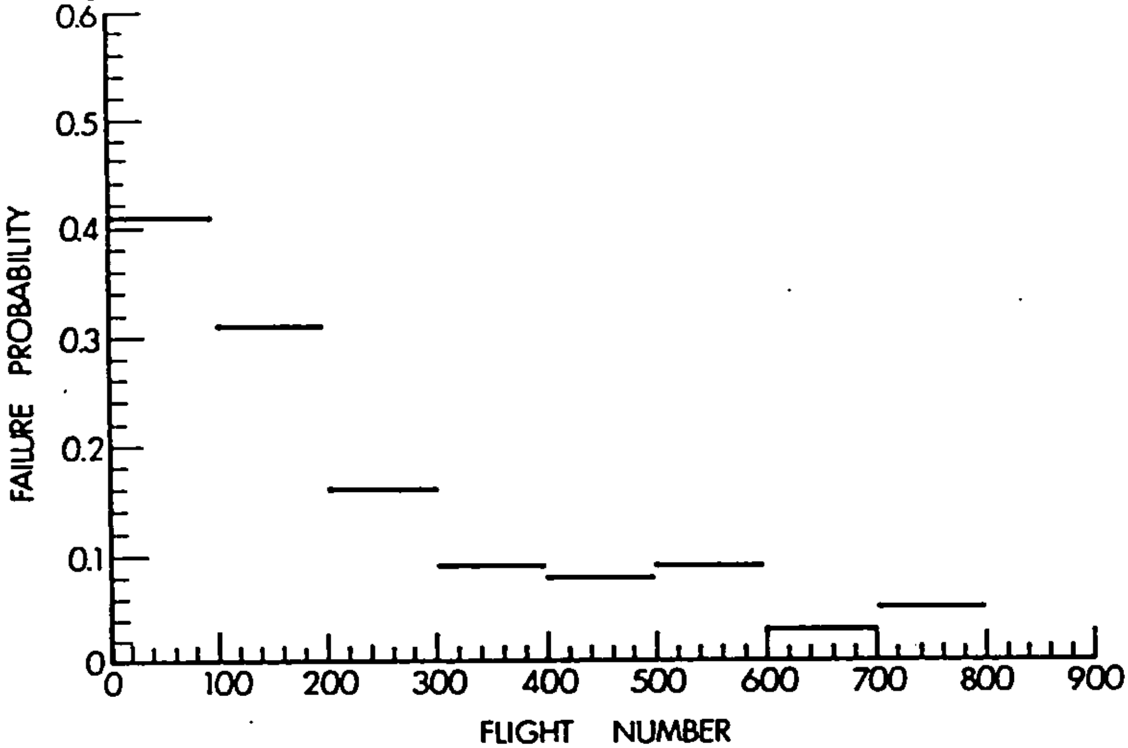


Figure 5.41 Average Flight Failure Probability by 100 Flight Intervals.

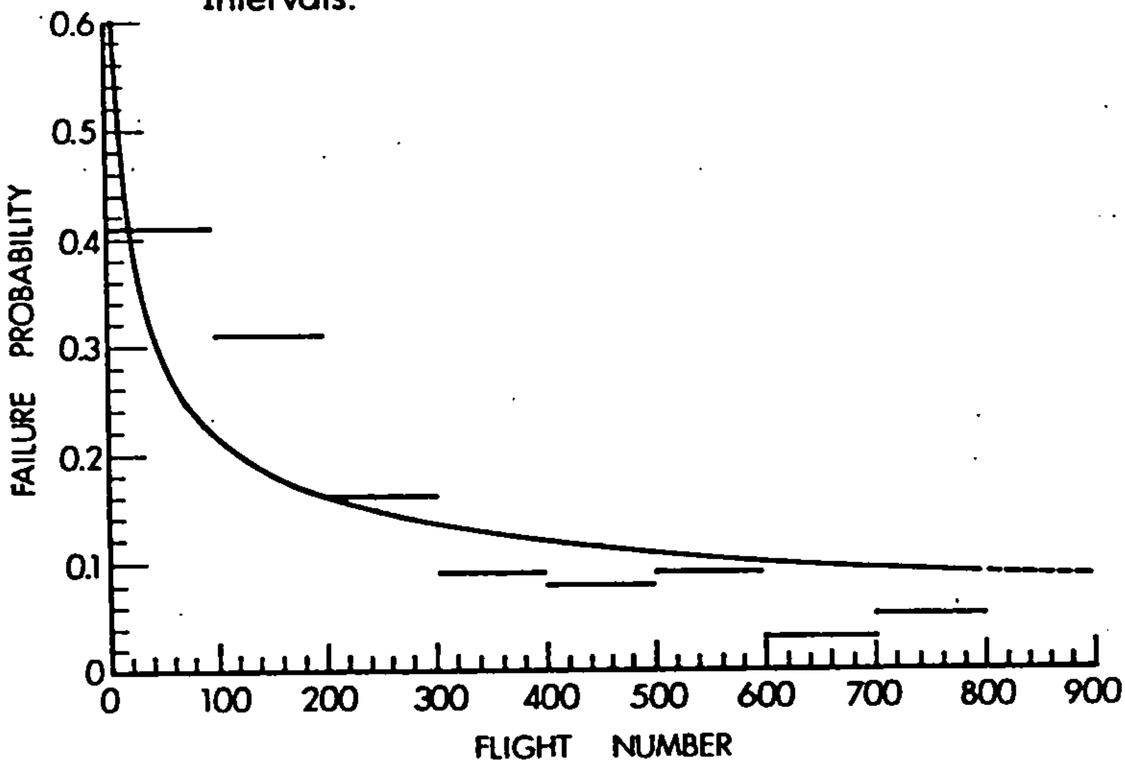


Figure 5.42 Initial Estimate of Failure Probability for 801 Valid Flights.

Further investigation revealed that the development program experienced a major re-emphasis on reliability improvement after the 200-th flight. This re-emphasis included major design changes being incorporated into the system at about the 200-th flight which resulted in a significant jump in the missile reliability. Therefore, there was a physical justification for breaking the data into two sets at the 200-th flight. The failure probability was estimated separately for the first 200 flights and again for the remaining 601 flights. In both cases, the goodness of fit test to the data was acceptable and also the relative positions of the curves to the plotted averages improved as shown in Figure 5.43.

The reliability, i.e., 1 minus the failure probability, was estimated for each curve. This is shown in Figure 5.44. This is the estimated reliability growth curve for the missile.

The exact lower 90 percent confidence bound at the 801-st flight was computed using all the data on flights 201 through 801. The resulting lower bound was .93. Similar lower confidence bounds can be computed periodically to determine when the system has sufficiently demonstrated the required reliability. Once this has been established, emphasis in the development of a system can then be directed to other areas.

We next considered how growth tracking could have been used to project system reliability during development. If reliability growth had continued in the direction it was obviously going in the first 200 rounds, the system would have been in trouble. The project office realized this and made a concerted effort to improve the situation. If the AMSAA model had been used for tracking, then at 200 flights the reliability estimate would have been .68, and a projected estimate to 800 flights would have been .73. This projection would indicate to management that the reliability requirement of .95 would not be met with the present development effort (see Figure 5.45).

There was, of course, a major re-emphasis on reliability after the 200-th flight, and based on the next 100 flights (201-300), the reliability estimate at 300 would have been .89, and a projection of the reliability at the 800-th flight would have been .94. The estimated rate of reliability growth would have indicated that the requirement could be met. This is shown in Figure 5.46.

Table 5-III shows the prediction capability of the AMSAA growth model as the data are increased by increments of 100 flights. The projected reliability changed very little with the added data base and would have indicated as early as the 300-th flight that management could expect to meet the reliability requirement with the present development effort.

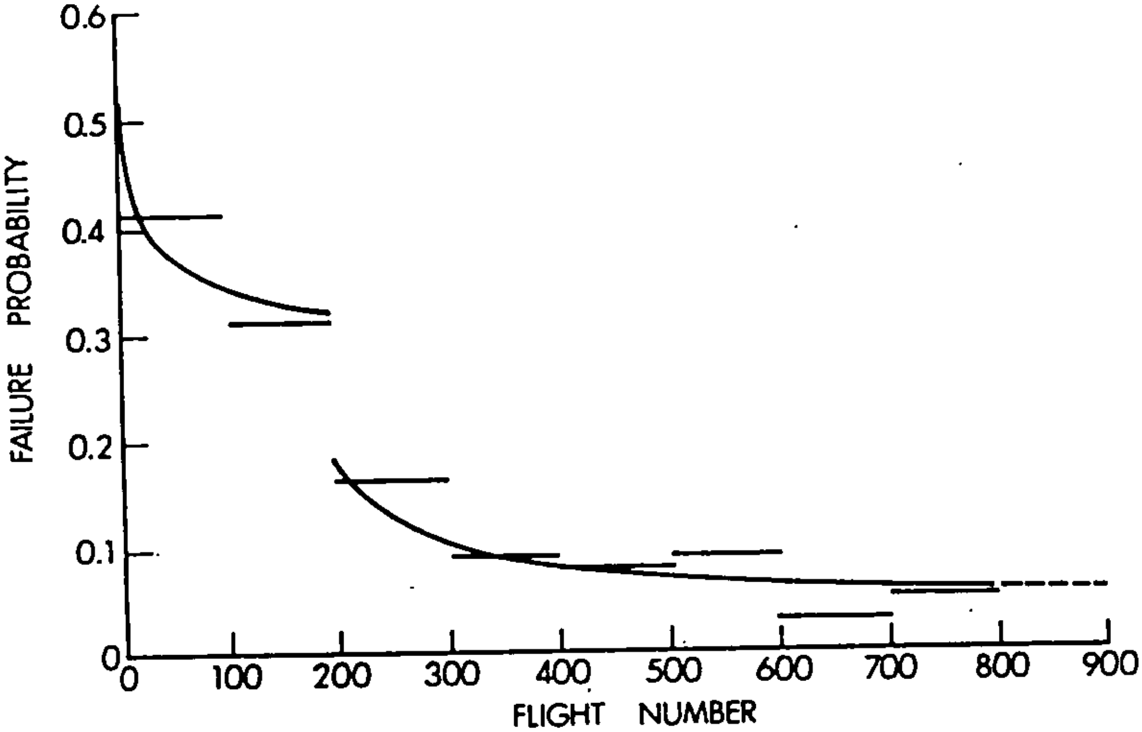


Figure 5.43 Estimate of Failure Probability Based on AMSAA Model.

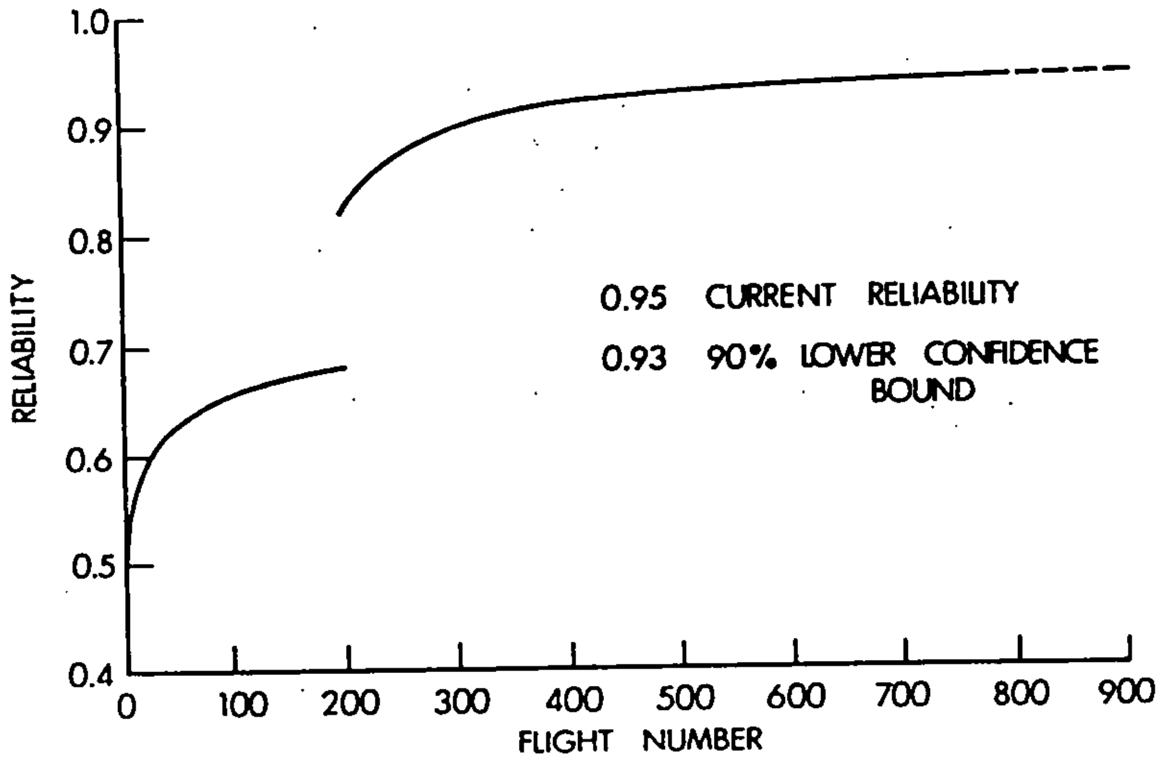


Figure 5.44 Estimate of Reliability Based on AMSAA Model.

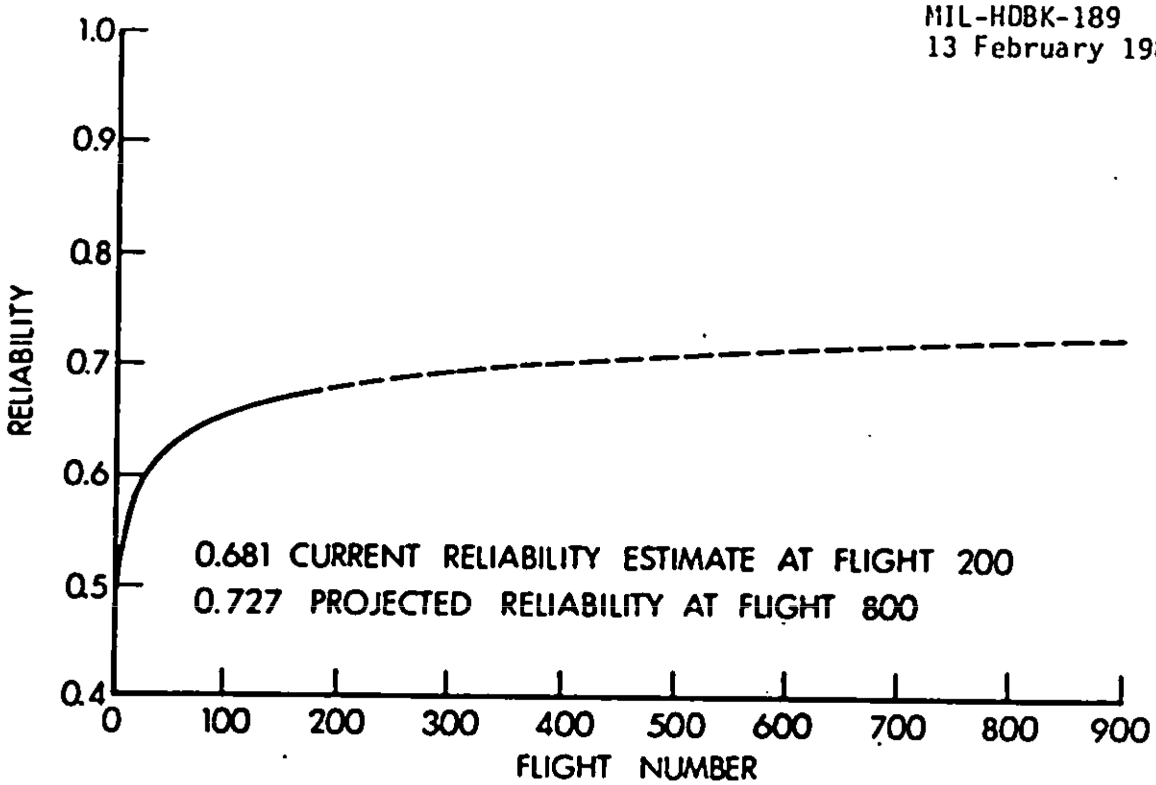


Figure 5.45 Reliability Estimates Based on First 200 Valid Flights.

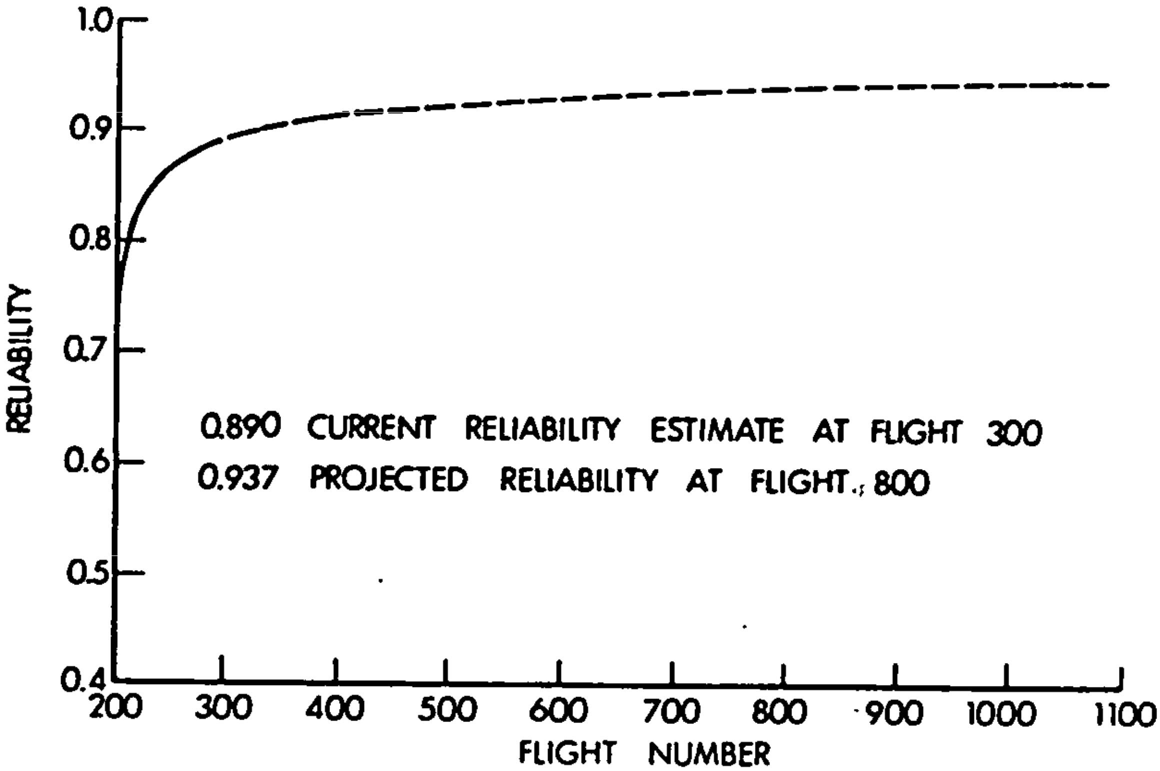


Figure 5.46 Reliability Estimates Based on Flights 201 to 300.

TABLE 5-III

Flight Number	Estimated Reliability	Projected Reliability At Flight 800
200	.68	.73
300	.89	.94
400	.92	.94
500	.93	.94
600	.93	.94
700	.94	.95
800	.95	.95

5.3.8 Tracking Systems with High MTBF. Systems with a high MTBF relative to the test time may be difficult to track entirely within each major test phase because of the small number of observed failures. A reasonable approach to this problem is to structure the program essentially as one major test-fix-test phase and combine all test data to fit a growth curve. This procedure would usually require that no significant jumps in reliability occur and that the test environment and development effort be held fixed. Because of the few problems that are observed for systems with high MTBF, fixes may be incorporated into the system during the test instead of delayed until the end of testing. In this case the program is of the test-fix-test type and may be considered as one test phase.

5.3.9 Reliability Growth Projections. Reliability projections can be made by extrapolating a fitted growth curve and by engineering analyses. The next two examples illustrate these methods.

Example 6. In Example 4, the AMSAA reliability growth model was fitted to a set of failure data resulting in estimates of $\hat{\lambda} = .332$, $\hat{\beta} = .616$ for the parameters. Under this model the MTBF at time t = 3000 is given by $\hat{m}(t)=[\hat{r}(t)]^{-1} = 106$. This is the current assessment. A projection to time 3500, which may, for example, be the end of the test phase, is determined by evaluating $\hat{m}(t)$ for t = 3500. This gives $\hat{m}(3500) = 112$, which is the expected reliability to be attained with 500 additional hours of development testing, if the present growth rate continues. This is shown in Figure 5.40.

Example 7. Analysis of reliability data collected during the development of an Army helicopter indicated that the observed reliability growth was insufficient. Projections on the tracking curve indicated that the reliability requirement would not be met. A major milestone had arrived where a high level decision needed to be made regarding the future of the aircraft development program.

It was known that the contractor had many delayed fixes which he planned to implement in the future and that the projection of the tracking curve could not anticipate the effect of these fixes. It was, therefore, determined that an engineering assessment of these fixes would be required in order to predict more accurately the future reliability of the aircraft. Therefore, a study which was unique at that time was conducted by a special team of engineers and analysts.

Primary emphasis during the initial investigation was on discussions with contractor technical personnel about specific failure modes which had occurred during testing, especially those failure modes where design modifications had been incorporated or were planned. In all, 120 Failed Item Analysis Reports (FIAR), grouped into 81 failure modes, were reviewed. After discussing each failure mode with contractor personnel, team members recorded a personal estimate of the effectiveness of contractor corrective actions to reduce failures. Effectiveness estimates (k factors) were scored as the proportion of the failures estimated to be eliminated if the corrective action had been available for the flying period covered. The personal estimates of corrective action effectiveness provided by the 11 Army team members were averaged for each major subsystem. These averages, or k factors, are summarized in Table 5-IV.

For example, a k factor of .60 assigned to a fix for a particular failure mode over a certain time interval would indicate that after the fix is incorporated, 60 percent of the number of failures of that particular mode over the same time interval would not be expected to occur; that is, only 40 percent of the number of failures would be expected to occur over the same time interval if the fix was incorporated.

Thus, if N represents the number of failures of a particular mode occurring over a certain time period, and N_E represents the expected number of failures over the same time period after the fix with effectiveness factor k is incorporated, then $N_E = N(1-k)$. For example, if 5 failures of a particular mode occur over a certain time period, and a fix with k factor .60 is identified, then $N_E = 5(1-.60) = 2$. That is, after the fix is incorporated, only 2 failures would be expected to occur over the same time period.

The analysis of existing failure modes was based on failures charged by Army personnel during 239 hours of flight testing. Of the 81 failure modes mentioned above, the contractor had identified fixes for only 66. The failure rates associated with the remaining 15 failure modes, therefore, would not be expected to change and since 25 failures were associated with these modes each of these 25 would be fully counted in the assessment process.

However, the 66 modes for which the contractor had identified fixes would need to be analyzed by the evaluation team to estimate the effectiveness of each. The 95 associated failures would be adjusted accordingly to reflect the reduced failure rates that may be expected.

TABLE 5-IV EXPECTED TOTAL FAILURE RATE, NON-WEAROUT MODES

Subsystem	Number Failure Modes	Number of Failures	Average k Factor*	Expected N_E Number of Failures After k Factors Are Applied**	Expected Failure Rate Based upon 239 Hours
Rotor	10	16	.708	4.88	0.020
Transmission	7	7	.642	2.51	0.011
Propulsion	14	19	.576	8.49	0.036
Electrical	6	14	.584	6.00	0.025
Avionics	8	9	.362	5.93	0.025
Airframe	6	7	.598	3.10	0.013
Electronic Controls	3	3	.689	0.93	0.004
Hydraulics/Flight Control	12	20	.551	9.08	0.038
TOTAL	66	95		40.92	0.172

*These are averages over subsystems; therefore, N_E cannot be obtained directly from these numbers.

**Obtained by applying to each failure the average k factor assigned by the Government Evaluation Team.

The results of this analysis are summarized in Table IV. The 66 failure modes have been classified by major helicopter subsystem. Therefore, the k factor listed for each subsystem represents only the average factor for the modes occurring within the subsystem; therefore, the expected number of failures (N_E) cannot be determined exactly from the average k factors. (Space would not permit listing all failure modes.)

N_E represents the evaluation team's best judgment of the number of failures that would be expected to occur during an equivalent period of test time (239 hours) after the contractor has installed his proposed fixes. The final column shows the expected failure rate after the fixes.

The estimate of mean time between failure (MTBF) based upon the original test was 239 hours/120 failures = 2.0 hours. The expected number of failures after the fixes, however, is now 40.92 plus the 25 failures that resulted from failure modes for which fixes were not identified. Therefore, the new estimate of MTBF would be 239/(40.92 + 25) = 3.6 hours. The Army consequently used this number to estimate the reliability of the mature helicopter system.

5.4 <u>Contracting for reliability growth</u>. Growth procedures and concepts must be clearly and effectively translated to any contract for developing a system. Because contracts and contracting procedures vary greatly both within and among the services, it will be necessary to reduce these procedures to the basic structure which any contracting procedure must follow.

First, prospective contractors must be solicited, and a detailed accounting of what is needed must be given to each. For most military contractors this is called the Request for Proposal (RFP).

Second, each contractor must respond to the RFP with a statement as to what each believes he can deliver.

Third, after proposal evaluation and some possible negotiations, a contractor is selected and a contract is signed.

This section will, therefore, be divided into three parts: 5.4.1 Request for Proposal (RFP); 5.4.2 Evaluation of Proposal; and 5.4.3 The Contract. Primary emphasis, however, will be placed upon the RFP since the subsequent evaluation of the proposal and the contract are primarily a matter for negotiation.

5.4.1 Request for proposal (RFP). The RFP must clearly define what is expected in the contractor's proposal regarding his reliability growth program. It should basically consist of four areas of discussion:

Reliability requirements (interim and final)

Planned growth curves

Testing

Tracking reliability growth

5.4.1.1 Reliability requirements. Realistic requirements are, of course, basic to the entire reliability program and are not directly associated with the growth program. However, it is important to understand that the growth program is dependent upon realistic requirements since the reliability must, at some time grow to equal or exceed the requirements. Basically the requirements must reflect a need and must reflect the state-of-the-art within constraints on cost. There are two types of requirements that must be considered for reliability growth purposes:

5.4.1.1.1 Final requirements. At some point in time the reliability must equal or exceed some pre-determined goal. The point in time may vary with programs (production, fielding, etc.), but each program is required to specify a final reliability requirement or goal.

5.4.1.1.2 Interim reliability requirements. These are reliability requirements imposed at specific milestones during the development cycle. The interim requirements must lead to the final requirement and must set a standard by which the progress of the program may be judged. For this reason, interim requirements must be specified for the same hardware and should be determined by the planned growth curve.

Interim requirements must also be located at times during the program when sufficient data are available to make reasonable inferences regarding the reliability of the system. Moreover, highly reliable systems will normally require a large amount of test time before a requirement can be verified based on data.

5.4.1.2 Planned growth curve. A planned growth curve must be required by the RFP. It should be developed as described elsewhere in this document. However, the following specific requirements regarding this curve should be specified in the RFP.

5.4.1.2.1 Application. It is important that the specific equipment to which the planned growth curve applies be clearly defined. For instance, if the curve applies to a system to be developed, does this system include government furnished equipment, equipment being developed by other contractors, etc.? Normally the requirement, as far as the contract is concerned, applies only to the equipment for which the contractor is responsible. In any case, however, the planned growth curve must apply to the same equipment to which the final reliability requirement applies.

There is no need to provide a planned growth curve for subsystems or components if there is no contractually binding reliability requirement for that subsystem or component. (Such curves may in some cases be of value for analysis or other purposes, but they need not be part of the contract). This point is mentioned because some contractors have, in the past, included hundreds of meaningless planned growth curves in their reliability program plan, indicating a lack of understanding of the purpose of the growth concept.

5.4.1.2.2 Prototypes. If several prototypes are to be developed under different conditions, then the RFP must either require a planned growth curve for each condition, must specify an average condition that should lead to the requirement, or must specify the prototype(s) to which the planned growth curve applies. The following are examples of a number of situations that could occur and what should be done in each case.

Example 1.

Two prototypes are to be developed. One is to receive fixes as they are developed; the other is to receive few or no fixes until late in the program.

Solution: A growth curve should be submitted for each prototype. The curves should meet at a point later in the program when both prototypes have the same configuration. The item receiving the fixes as they are generated should be considered the lead or control prototype and the program evaluation should be based upon it until such time as the other prototype receives the same fixes.

Example 2. Several prototypes are to be developed at different points in time. Newer prototypes will include design changes already incorporated in the older prototypes and will, therefore, be similar in configuration to the older prototypes at the same time that testing is begun on the newer prototypes.

Solution: It will be assumed that at time, t, on the growth curve scale, each prototype will be of approximately the same configuration age although their chronological ages may differ. One planned curve should be sumitted which represents an average reliability for the prototypes as a function of development time.

Example 3. One or more prototypes are to be tested under either more severe environmental conditions or less severe environmental conditions than the reliability requirement specifies.

Solution: The planned growth curves normally should not apply to these prototypes. If absolutely necessary, however, K-factors can be applied to the data to compensate for unusual environments. This procedure is not recommended except under most unusual circumstances. The K-factor problem will be discussed later in more detail (para. 5.4.1.3.3).

5.4.1.2.3 Historical information. As mentioned earlier, all available data and/or information should be utilized in the construction of the planned growth curves. The RFP must require that such data and/or information be used. The following should be required:

Source of data and/or information

Applicability of data and/or information

Degree to which judgment has been used

Method of analyzing data and/or information

Rationale for determining the starting point for the curve

5.4.1.2.4 Milestones. All proposed milestones must be identified in the contractor's proposal. These milestones must be associated with specific points on the planned growth curve.

The RFP must specify that milestones shall correspond to points where viable decisions can be made. For instance, if numerous fixes are to be incorporated at one time, it would not be appropriate to schedule a milestone at or near this point in time. The reason is that the effectiveness of the fixes cannot be determined until some testing has been conducted subsequent to the fixes. (See Failure Purging, para. 5.4.1.3.5.)

5.4.1.2.5 Relationship to final requirement. The planned growth curve must lead to the final requirement at an acceptable point in time. Normally this point should occur at or before full production of the item. A planned growth curve that does not lead to this requirement may be interpreted as a prediction that the requirement cannot be met.

5.4.1.3 Testing. The RFP must require the contractor to propose a program or plan for testing the prototypes under development and for reporting the test results. So far as reliability growth is concerned, test results need only to be reported on those systems, subsystems and/or components for which planned growth curves are submitted, i.e., lower level test results should not normally be used. The amount of reliability demonstration testing may be reduced since the techniques outlined in this document permit the use of development test data for reliability assessment purposes. The following information should be required in the RFP.

5.4.1.3.1 Type of test program. The contractor must specify the type of testing he plans to conduct, e.g.,

Will fixes be incorporated as they are developed? (test-fix-test)

Will testing be interrupted for periods of time for the incorporation of many fixes? (test-find-test)

Will a combination of the above be employed? (test-fix-test with delayed fixes)

In any event the type of testing will affect the planned growth curve and the RFP must require that the type of testing and the planned growth curve be compatible.

5.4.1.3.2 Environments. The RFP must require that proposed environmental test conditions be defined. Only testing conducted in environments defined for the requirements are directly applicable to reliability growth estimation. For instance, a field requirement cannot normally be demonstrated in a laboratory environment unless provisions are made for simulating the field environment. Every effort must be made to assure that sufficient testing is conducted under the environment specified for the final requirement.

If testing under non-representative environments must be used for reliability growth purposes, then appropriate adjustments must be made. One such adjustment is the use of K-factors discussed next.

5.4.1.3.3 K-Factors. When testing is to be conducted under either more severe environmental stress or less severe environmental stress than the field reliability requirement specifies, the use of K-factors may represent an acceptable means of transforming the test results from one environment to another if test results under environments specified for the requirement are impossible to obtain. For example, if a government approved analysis indicates that one hour of environment stress is the equivalent of two hours of normal stress, then the K-factor would be two. The times during which the stress was applied could then be multiplied by two and the failure times occurring during that time period could be appropriately adjusted.

Example: Ten hours of testing is to be conducted of which all but the fourth and fifth hour are to be conducted under normal stress. The fourth and fifth hour are to be conducted under severe stress, and a K-factor of two was found to represent an appropriate correction. The failures occurred at the following times:

Actual Failure Times (t_A) (hrs.)
1.1
2.7
3.2
4.1
4.6
5.0
5.3
5.7
6.8
8.2
9.1

To obtain the equivalent failure times (t_E) use the following transformations:

Actual hours	Transformation
$0 \le t_A \le t_{A_0} = 4$	$t_E = t_A$
$t_{A_0} = 4 < t_A < t_{A_0} + T = 6$	$t_E = K(t_A - t_{A_0}) + t_{A_0}$
$t_{A_0} + T = 6 \le t_A \le 10$	$t_E = t_A + (K - 1)T$

where t_{A_0} equals the actual time that the stress testing begins and T equals the number of hours of stress testing. For this case $t_{A_0} = 4$ and $T = 2$. The following equivalent test times are, therefore, obtained.

Equivalent Failure Times (t_E) (hrs.)
1.1
2.7
3.2
4.2
5.2
6.0
6.6
7.4
8.8
10.2
11.1

The use of K-factors should not be recommended in the RFP unless there is a valid and specific need to use them and unless the contractor can propose a realistic means of arriving at a suitable number. (See Section 5.4.2.3 Evaluation of Proposal). K-factors are, at best, very subjective.

5.4.1.3.4 Incident Classification and Reporting. An incident is defined for this handbook as an occurrence during test which may or may not be classified as a failure. For contract purposes only those incidents classified as failures chargeable to the contractor are included for the reliability purposes. For instance, the contractor cannot be held responsible for failures of government furnished equipment (GFE).

The RFP must require the contractor to propose a means for detecting, reporting and evaluating incidents that may be related to failures. He should first specify what information he proposes to record. This information should include the following as a minimum:

Description of incident

Chargeability. Should the incident be charged to the contractor as a relevent reliability failure, and how should this be determined? A means for handling disagreements between the government and contractors should be specified.

Identification of failed item

Time of failure in terms of the time scale for the planned growth curve

Classification of failure in terms of its effect upon the mission and/or other criteria such as cost

The specific form of classification is left to the discretion of the project manager since it is highly dependent upon the particular type of system being developed.

All incidents must be reported and the RFP should clearly prohibit the elimination during reliability assessment of any failures on the basis of design fixes.

For specific instructions regarding data ordering and classification failures, see MIL-STD-781 and MIL-STD-785.

5.4.1.3.5 Failure purging. The RFP should clearly state that failure purging as a result of design fixes is an unnecessary and unacceptable procedure when applied to determining the demonstrated reliability value. It is unnecessary because of the recently developed statistical procedures to analyze data whose failure rate is changing. It is unacceptable for the following reasons:

a. The design fix must be assumed to have reduced the probability of a particular failure to zero. This is seldom, if ever, true.

Usually a fix will only reduce the probability of occurrence; and in some cases, fixes have been known to actually increase the probability of a failure occurring.

b. It must be assumed that the design fix will not interact with other components and/or failure modes. Fixes have frequently been known to cause an increase in the failure rate of other components and/or failure modes.

The hard fact is that fixes do not always fix; and, therefore, the attitude of the government must be to defer judgment until further testing is conducted. However, even after the effectiveness of a design fix has been established, failures associated with eliminated failure modes should not be purged. The reason is - if there has been sufficient testing to establish the effectiveness of a design fix, then an appropriate reliability model will, by then, have sufficient data to reflect the effect of the fix in the current reliability estimate.

The above discussion, of course, applies to the demonstrated reliability values. It may, however, be necessary to weight the effectiveness of proposed fixes for the purpose of projecting reliability. However, the difference between assessments and projections must be clearly delineated.

5.4.1.4 Tracking reliability growth. The RFP must require a means of tracking the reliability growth of the developmental item and of comparing the tracked growth with the predicted growth given by the planned growth curve.

5.4.1.4.1 Analytical methodology. It would not be appropriate for the RFP to specify precisely what methodology will be used. A primary responsibility of this handbook is to convey the message that an acceptable analysis cannot be standardized. The test data itself must ultimately determine the analytical methodology.

Past experiences with military systems under development have provided convincing evidence that the AMSAA model is the most versatile procedure for tracking reliability growth. Other models may on particular occasions fit some sets of data better, but the AMSAA model has been found to fit nearly all reliability data. Furthermore, it has been found through experience that in those cases where the AMSAA model does not fit, other models usually do not fit either.

For this reason it is recommended that the RFP require that the AMSAA model be applied first to determine its applicability. If a poor fit is obtained, steps should be taken to determine any physical reason for a lack of fit. A common reason, for instance, has been that for periods of time only part of the system was being exercised. During these times, of course, the failure rate would be expected to be lower than during those times that the full system was being exercised. Frequently procedures of this sort are not reported.

Another common reason for a poor fit is a sudden and unexpected change in the course of a program. Such a change could be the result of a management change, procedural change, failure criteria change, etc. The model can frequently serve as a tool for detecting such changes and/or their effect on the reliability estimate.

A poor fit, therefore, is not necessarily a valid reason for discarding the AMSAA model since frequently such an occurrence reveals information that may have gone unnoticed. Furthermore, the model will usually fit the data in segments as described elsewhere in this document.

The purpose of requiring the AMSAA model is not to impose a particular methodology whether or not it is appropriate. If it can be shown, from the hard data, that the model is inappropriate then other alternatives should be proposed. The purpose of the AMSAA model is to impose some firmness or strength into the contract. This will, in turn, provide a measure of standardization into military contracts and will prevent the use of complex models that do not reflect the real world and are usually difficult to understand or refute without an extensive study.

5.4.1.4.2 Application. The RFP should clearly indicate to the contractor that the reliability growth curve (tracking curve) will be used to assess the progress of the development effort in comparison to its predicted progress as defined by the planned growth curve. During the course of the development program, the contractor, therefore, will be held accountable for significant departures from the predicted reliability as defined by the planned growth curve.

This comparison may be made by computing confidence limits for the tracking curve and noting the location of the planned curve with respect to the confidence limits. For example, if the planned curve lies below the lower limit of the tracking curves, the reliability growth program may be considered ahead of schedule. On the other hand, if the planned curve lies above the upper confidence limit, the reliability growth program may be considered behind schedule. The growth program, of course, should be considered on schedule if the planned growth curve lies between the limits.

Although primary consideration should be given to the status of the program at the point in time that the latest data were available, it is possible that the program will at that point be on or ahead of schedule but that the projection will not lead to the requirement. This will usually be caused by a planned jump or other rapid increase in reliability of some future time. If such is the case then the projection computed before the jump would not be expected to project to the requirement since the projection is based upon past performance and cannot predict jumps or other rapid increases. This must be understood and discussed in the contractor's proposal.

5.4.1.4.3 Planned course of action. The RFP must require the contractor to submit a planned course of action based upon the comparison of the planned growth curve with the tracking curve. If the reliability growth program is either on or ahead of schedule, then no course of action is required. However, if the reliability growth program is behind schedule the contractor must be prepared to take action.

The course of action should include a reevaluation of the reliability program. He should, for instance, determine whether or not the problem is associated with a particular component or sub-assembly. If this should be the case, he should conduct an engineering study to determine what can be done to improve the reliability of that component or sub-assembly. If nothing can be done he should look at the possibility of improving the reliability of other components or sub-assemblies in order to compensate for the lower reliability of the component or subassembly in question.

If the system reliability problem cannot be associated with a particular component or sub-assembly, the contractor should conduct an engineering study of the most prevalent failure modes to determine what can be done to reduce the rate of occurrence of failures to the point where the reliability program will be back on track.

Although the above engineering studies should be conducted in conjunction with government hardware and reliability experts, a detailed report of the findings should be forwarded to the project manager and other appropriate elements of the government.

5.4.2 Evaluation of proposal. The contractor must respond to the Request for Proposal (RFP) with his proposal. The proposal must address all information requested in the RFP. The contractor must justify any proposed variances from what has been requested in the RFP. The Government in turn must evaluate the contractor's proposal. This section provides guidelines for that evaluation.

5.4.2.1 Reliability requirements. Although the requirement applies to the overall reliability program, for purposes of growth management it must be determined at what point in time the contractor plans to meet the requirement and if he has adequately supported any predictions relative to meeting the requirement either early or late (as opposed to the normal schedules where the requirements are expected to be met at some time near initial production).

5.4.2.2 Interim reliability requirements and planned growth curve. Interim reliability requirements should coincide with the planned growth at any point in time. The following questions may be used as guidelines in the Government evaluation of the contractor's proposal for planned growth curves.

5.4.2.2.1 Have planned growth curves been submitted and do they comply with instructions given elsewhere in this document?

5.4.2.2.2 Do the planned growth curves apply to systems, subsystems, and/or components for which there are contractually binding reliability requirements?

5.4.2.2.3 Are prototypes being developed under different conditions and/or during different intervals of time?

5.4.2.2.4 If the answer to the previous question is "yes", have the various conditions and/or time intervals been accounted for in the construction of the planned growth curves?

5.4.2.2.5 Has historical information been used; and, if so, does it include the information required in paragraph 5.4.1.2.3?

5.4.2.2.6 Have appropriate milestones been identified? If so, are they located at points in time where sufficient information will be available to provide current inferences regarding the progress of the growth program? (Keep in mind that the effect of design fixes cannot be determined until an adequate amount of system testing has been performed).

5.4.2.2.7 Do the planned growth curves lead to the requirement at an acceptable point in time? If the curves lead to the requirement either early or after the requirement should have been met, has supporting rationale been provided? Is the rationale acceptable to the Government?

5.4.2.3 Testing. Testing should be discussed as it affects the reliability growth program. The following questions should, as a minimum, be answered by the Government in its evaluation?

5.4.2.3.1 Has the type of testing been specified as to diagnostic time, fixes, and implementation of fixes?

5.4.2.3.2 Is the type of testing compatible with the planned growth curve? For instance, if many fixes are to be incorporated over a short period of time between testing, does the planned growth curve show an appropriate "jump" in reliability?

5.4.2.3.3 Is non-standard environmental testing to be conducted? Non-standard environmental testing is defined as testing performed under conditions that do not simulate the conditions under which the requirement is based.

5.4.2.3.4 If the answer to the previous question is "yes", have appropriate measures been taken to either eliminate the results of non-standard tests from the data or to adjust the results to make them compatible with the requirement? (See paragraphs 5.4.1.4.2 and 5.4.1.4.3).

5.4.2.3.5 If the use of k-factors is proposed, has the contractor established a realistic means of arriving at an appropriate number?

5.4.2.3.6 Is the contractor clearly committed to recording all incidents that may be related to a failure?

5.4.2.3.7 Will the incident be fully discussed and described?

5.4.2.3.8 Does the contractor propose a means of classifying incidents, e.g., chargeable failures, mission failures, non-relevant, etc.?

5.4.2.3.9 Does the contractor clearly state that all chargeable failures will be considered and that assessed failures will not be purged on the basis of design fixes?

5.4.2.4 Tracking reliability growth. The contractors proposal must realistically discuss his proposed tracking procedures and his planned course of action, should he fall below his predicted reliability given by the planned growth curve. The following questions should be addressed regarding tracking.

5.4.2.4.1 Does the contractor propose to use the AMSAA model?

5.4.2.4.2 If he proposed a different model, has he provided sound justification for its use?

5.4.2.4.3 Is his model easily understood?

5.4.2.4.4 Does he produce empirical evidence that his data will fit the model (i.e., evidence similar to the numerous examples available for the AMSAA model)?

5.4.2.4.5 Does the contractor understand how his progress will be tracked by the government (paragraph 5.4.1.4.2)?

5.4.2.4.6 Has the contractor proposed an appropriate course of action should his progress fall behind schedule?

5.4.3 The contract. When a contractor has been selected, the procedures for reliability growth management should be included in the contract. Each development program is different, and responses to the RFP will vary in many ways. It is, therefore, impossible for this document to prescribe systematically what must be included in all contracts. Ideally, however, everything that the RFP required should be included in the contract. The contractor's response should determine how much of the RFP is attainable subject to the government's evaluation of his rationale. Differences, of course, must be negotiated. However, it should be understood that a viable reliability growth management program will have limited impact if it is not contractually binding.

CUSTODIANS:
ARMY- CR
NAVY- EC
AIR FORCE- 17

PREPARING ACTIVITY
ARMY- CR

REVIEW ACTIVITIES:
NAVY-EC SH AS

PROJECT NUMBER-RELI-0002

Next page is blank.

This page intentionally left blank.

APPENDIX A

ENGINEERING ANALYSIS

10. SCOPE

10.1 Purpose. The majority of reliability growth data analyses are statistical analyses. Statistical analyses view growth as being the result of a smooth, continuous process. In fact, reliability growth occurs in a series of finite steps corresponding to discrete design changes. Mathematical models describe the smooth expectation of this discrete process. Rather than being concerned about whether specific design changes are effected rapidly or slowly--or whether they are very effective, not effective, or even detrimental--the statistical models work with the overall trend. In most situations, this is a desirable feature as it focuses attention on longterm progress rather than on day-to-day problems and fixes. The application of statistical analyses relies on analogy. For example, the growth pattern observed for program A may be used as a planned growth model for program B, because the programs are similar. As another example, the growth pattern observed early in program B may be extrapolated to project the growth expected later in the program because of similarities between the early and later portions of the program. The difficulty that occurs in applying the analogy approach is that perfectly analogous situations rarely exist in practice. The engineering analyses described in this section rely on synthesis. That is, they build up estimates based on a set of specific circumstances. There is still, however, reliance on analogy; but the analogies are applied to the parts of the problem rather than to the whole. Although synthesis may be used to provide a complete buildup of an estimate, it is simpler and more common to use synthesis to account for the differences, or lack of perfect analogy, between the baseline situation and the situation being analyzed.

10.2 Application. The general approach to growth planning and long-term projection is similar to that used for assessment and short-term projection purposes. The main difference is that for planning and long-term projection purposes, attention must be directed to program characteristics and general hardware characteristics, since specific design changes are unknown at the time of program planning. For assessment and short-term projection purposes, attention must be directed to the specific hardware changes made or anticipated. For the most part, the program and general hardware characteristics can be ignored, since they have already played their role in determining the specific hardware changes. The only differences between assessment and short-term projection is whether a change has been incorporated in the hardware or not. The analysis is the same in both cases except that recent test results may be incorporated in the assessment. It should also be noted that the type of assessment described in this section, because of the judgment involved in arriving at it, is particularly suitable for use within an organization. For inter-organization use, completely objective demonstrated values, computed by a means acceptable to the organizations concerned, are usually necessary.

20. ASSESSMENT AND SHORT-TERM PROJECTION

20.1 Application. At times, it is desirable to assess or project reliability growth by means of engineering analysis rather than statistical analysis. This detailed look is usually desirable in the following situations:

a. When, near the end of a test phase, design changes have been, or will be incorporated without adequate demonstration. It is highly desirable to analyze these unverified "fixes" separately on their unique merits, rather than treating the "fixes" as average ones with a statistical model.

b. When a major design change is made, or will be made, in the future. Such a change often causes a jump in reliability that is unrelated to the growth process prior to the change, since it represents a departure from a pure "find and fix" routine.

c. When there are few distinct test and fix phases. In this case growth projections by statistical extrapolation may not be appropriate.

d. When it is desired to evaluate possible courses of design improvement. By considering the failure modes observed and possible corrective actions available, a desirable course of design improvement can be determined. For example, it can be determined if correction of the single-worst problem will bring the system reliability up to an acceptable level.

20.2 Objective. When a failure mode is observed on test, it becomes desirable to anticipate the improvement that can be expected in a system if that failure mode is subjected to design improvement. The ultimate improvement possible is to completely remove the failure mode or reduce its rate of occurrence to zero. The practical lower limit on the failure rate is limited by the state-of-the-art, and even this value can be attained only under perfect conditions. The failure rate actually attained will usually be somewhat higher than the state-of-the-art limit because of unforeseen minor faults in the design and the failure rates of the parts involved.

20.3 Design changes. Although this appendix emphasizes reliability analysis of design changes for reliability improvement, all design changes should be analyzed in this manner, since every design change has a potential for enhancing or degrading system reliability. This requires that the reliability management system be linked to the configuration management system and other pertinent programs such as for maintainability and producibility.

20.4 Significant factors. Some of the factors affecting the expected effectiveness of a design change for reliability are listed below. For convenience in application, these are categorized as factors that create reference values and factors that influence estimates.

a. Factors that create reference values:

(1) What is the failure rate being experienced in similar applications?

(2) What is the failure rate of components to be left unchanged?

(3) What is the analytically predicted failure rate?

(4) What failure rate is suggested by laboratory or bench tests?

(5) How successful has the design group involved been in previous redesign efforts?

b. Factors that influence estimates:

(1) Is the failure cause known?

(2) Is the likelihood of introducing or enhancing other failure modes small?

(3) Are there other failure modes in direct competition with the failure mode under consideration?

(4) Have there been previous unsuccessful design changes for the failure mode under consideration?

(5) Is the design change evolutionary, rather than revolutionary?

(6) Does the design group have confidence in the redesign effort?

20.5 Explanation of factors.

20.5.1 What is the failure rate being experienced in similar applications? The failure rate that a component experiences in similar applications serves as an objective reference point indicative of what may resonably be expected of that component.

20.5.2 What is the failure rate of components to be left unchanged? Since it is usually unreasonable to expect one of the worst components in a system to be among the best as the result of a design change, the average failure rate of components to be left unchanged can be used as a rough optimistic limit. Although the guidance provided by this reference value is not very firm and may easily be overridden by other factors, there are three reasons to encourage its use. First, it raises the general question of over-optimism. Second, it is a valid and common approach to reliability improvement to bring problem components into conformance with the other components in the system. Third, this reference value is among the easiest to determine.

20.5.3 What is the analytically predicted failure rate? The failure rate for the failure mode under consideration may, in some cases, be analytically predicted using techniques such as probabilistic design analysis. As an analysis of this type cannot consider unforeseen peculiarities in the design or application, such a value should be viewed as an optimistic limit.

20.5.4 What failure rate is suggested by laboratory or bench tests? Laboratory or bench test indications must be viewed with some scepticism. Specifically, an attempt should be made to judge whether test conditions are yielding an opimistic or pessimistic comparison.

20.5.5 How successful has the design group involved been in previous redesign efforts? The success rate of the design group provides another objective point of reference. For example, one organization has found that corrective actions are normally not more than 80 percent effective. Usually, this index is evaluated as the proportion of design changes that result in elimination (essentially) of the failure mode, or it is evaluated as the average proportion of failure rate reduction. In both of these cases, the range of failure rate values under consideration is between the current value and zero. The effectiveness of the design group may also be determined by the average proportion of the predicted improvement that is attained. In this case, the range of failure rate values under consideration is between the current value and the predicted value. This measure of effectiveness is more precise, but also more cumbersome, to work with. If this measure is used, it must be treated as an influence rather than a reference value.

20.5.6 Is the failure cause known? Knowledge of the failure cause relies heavily on the ability to perform a failed part analysis. Only when the failure cause and the precise failure mechanism are known can a design change be expected to be fully effective. At the other end of the spectrum are problems that must be attacked by trial and error because the failure case is (at least, initially) unknown. In this case, the expected effectiveness will be close to zero. Nevertheless, this type of a change may be used to gain insights that will give higher expectations in future changes.

20.5.7 Is the likelihood of introducing or enhancing other failure modes small? The likelihood of other failure modes being affected by a design change can usually be evaluated by use of failure mode and effect analysis. Attention should be directed to components that are adjacent to the affected one in either a functional or physical sense.

20.5.8 Are there other failure modes in direct competition with the failure mode under consideration? It is a special, particularly difficult situation when a component or assembly has other failure modes in direct competition with the failure mode under consideration. These are usually characterized by opposite failure mode descriptions such as tight, loose; or high, low. In a situation like this, there is no single, conservative direction, and avoiding one failure mode often results in backing into another. Seals on rotating shafts are an example of this type of problem. An application may initially have a leakage problem. Going to a tighter seal often results in a wear problem, and changing to multiple seals often causes the outer seals to run dry. The optimism solution in a case like this is usually a less-than-satisfactory compromise. And it is not unheard of to end up eventually with the original design.

20.5.9 Have there been previous unsuccessful design changes for the failure mode under consideration? Each unsuccessful design change for a specific failure mode will, in itself, lead to lower expectations for the effectiveness of further changes. This is caused by selecting the most promising alternatives first. However, previous unsuccessful changes may have provided sufficient information on the failure mechanism to outweigh this factor.

20.5.10 Is the design change evolutionary rather than revolutionary? Idealistically, an evolutionary change involves a single, small deviation from previous practice. Increases in either the magnitude or number of deviations make the change more revolutionary. When a design is refined in an evolutionary manner, the expectation is for improvement to occur with each iteration. A revolutionary design change is, however, virtually the same as a new design fresh from the drawing board (for the subsystem and components concerned). Thus, the redesigned part of the system may have an initial MTBF only, say, 10 or 20 percent of the predicted value. The revolutionary change may, however, have a potential inherently higher than the original design.

20.5.11 Does the design group have confidence in the redesign effort? Although subjective and intuitive, the confidence of the design group should reflect all of the factors previously discussed. Because of this, any analysis of reliability growth expectations should be compared against this intuitive feel; and, of course, the two opinions should compare well. As with any kind of cross-checking, the objective is to ferret out any errors and oversights. The main point is that an adequate analysis of reliability growth expectations cannot be accomplished without input from the design group.

30. METHODOLOGY

There are two major steps involved in estimating the effect of a design improvement. The first step involves using any reference values that can be determined to roughly define the range within which the new reliability value is expected to be. The second step involves considering the effect of the various influencing factors to narrow down to a likely point within this range. It must be emphasized that this methodology is a thought-process guide rather than an explicit procedure to be followed blindly. Some of the listed factors may be meaningless or inappropriate for a given design change. Some may be overshadowed by other factors. And some combinations of factors may have a net effect that is not consistent with the linearly additive relationship suggested in the example to follow. Special cases, such as a component with acceptable reliability that is to be modified for other reasons, will require adaptation of the basic procedure.

40. EXAMPLE

40.1 Objective. This example is intended to illustrate a general methodology that may be used to predict the effectiveness of design changes. This may be used as a method of assessment for design changes incorporated in the hardware, but not adequately tested. It may also be used to make short-term projections. This example considers just a single design change. It must be emphasized that the methodology is intended as a guide to reasoning, and no quantitative precision is implied.

40.2 Problem statement. The failure mode under consideration is weld cracking in a travel lock of a howitzer. The design change to be incorporated is an increased weld fillet size.

40.3 Analysis.

40.3.1 Determination of reference values. The first step is to determine any reference values that are obtainable as shown in Figure A-1.

Current Failure Rate	.0005 Failures per Round, as Demonstrated by Test
Analytical Prediction	None
Test Results	Lab tests (accelerated) show about a 4 to 1 improvement, suggesting a failure rate of about .00012 is attainable.
Failure rate of similar components in similar applications	None sufficiently comparable.
Success ratio of the design group	In general, they have been capable of removing 60% of the failure rate, implying .0002 as an expected failure rate.
Average failure rate of unchanged components	The system failure rate is .004, and there are roughly 300 active, or failure-prone components. .004/300 = .000013

FIGURE A-1. Reference Values

40.3.2 Design change features. The second step is to determine features of the design change that would influence the failure rate to be attained as shown in Figure A-2.

Is the failure cause known?	Moderately well. Analysis of broken welds showed no significant flaws; thus ruling out a quality problem. The level of forces encountered is not well known, and there is question about the stress concentration in the vicinity of the weld
Is there a likelihood of introducing other failure modes?	No other related failure modes are foreseen.
Are there competing failure modes?	No
Is the design change evolutionary?	Yes. This is a single, relatively minor change.
Have there been previous unsuccessful design changes for the failure mode under consideration?	Yes. This is the second change. The first change increased the cross-section of the stop. This caused some improvement, but the same type of cracking persists. Further increase in cross-section is impossible without a major design change.
Does the design group have confidence in this change?	Their confidence is moderate.

FIGURE A-2. Design Change Features

40.3.3 Defining and refining estimates. The third and fourth steps in the process involve defining the region of interest in terms of reference values and then refining estimates within (or perhaps slightly beyond) this region by consideration of the influencing factors. This process is shown graphically for illustrative purposes in Figure A-3. Point A represents a likely failure value, ignoring the influencing consideration. In this case, the lab test results were felt to be realistic and considerably more concrete than the general expectation, although the two values are in reasonably good agreement. The failure rate of other components does little more in this case than to provide assurance that the failure rate is only being brought into "reasonable conformance" to the rest of the system, rather than surpassing it. Line A-B represents the detrimental influence expected from some lack of knowledge of the failure cause. Since the failure cause is not known exactly, the lab testing may not have adequately reproduced the failure cause. Line B-C represents the influence expected from other failure modes that may be aggravated by the change. No influence is expected. Line C-D represents the influence expected from other competing failure modes. No influence is expected. Line D-E represents the influence of the evolutionary versus revolutionary nature of the design change. Since this is an evolutionary change, no effect is expected. Line E-F represents the detrimental influence expected from this being a second design correction attempt. Line F-G takes into consideration the confidence that the design group has in this change. Since their feelings are consistent with the analysis up to this point, no effect is shown. This analysis, then predicts a failure rate of about .00025 after the design change. Similar analyses for other design changes may then be combined to estimate the effect at the system level. Finally, it must be emphasized again that this type of estimator is highly subjective.

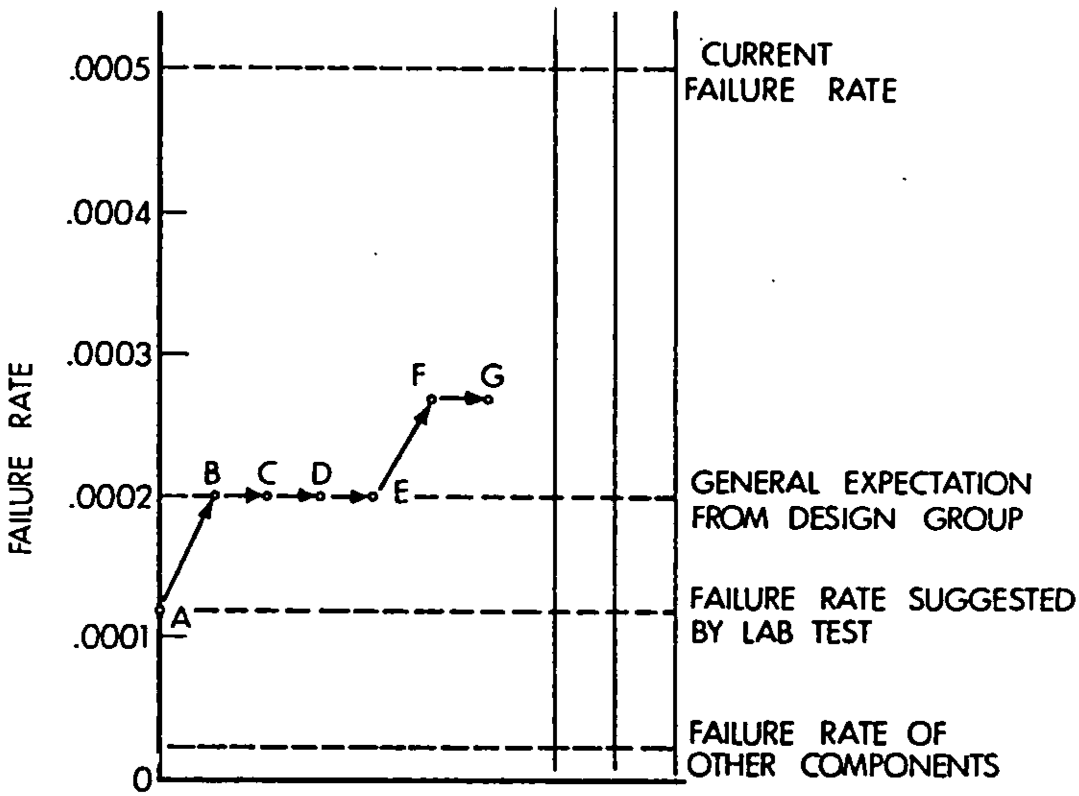


FIGURE A-3. Defining and Refining Estimates

50. PLANNING AND LONG TERM PROJECTION

50.1 Purpose. From an academic standpoint, growth planning and longrange projection have as their purpose the determination of the reliability growth that can be expected for a given set of program alternatives. From a more practical standpoint, a set of such analyses enables the program planner to evaluate the benefits and drawbacks of various program alternatives.

50.2 Approach. Basically, growth planning and long-range projection consider program constraints, activities, and sequencing to judge whether they will encourage or deter growth and to what extent. The three main variables of interest are the number of failure sources identified, the time required to perform the various activities, and the effectiveness of redesign efforts. Particular care must be taken when evaluating these variables to ensure that the sequencing of events is properly accounted for.

50.3 Organization or program characteristics. The basic reliability growth feedback model discussed in paragraph 4.2.1 will be used as a means of organizing and assimilating program characteristics. Because of the significance of hardware fabrication time, the fabrication of hardware element is included in the model as illustrated in Figure A-4.

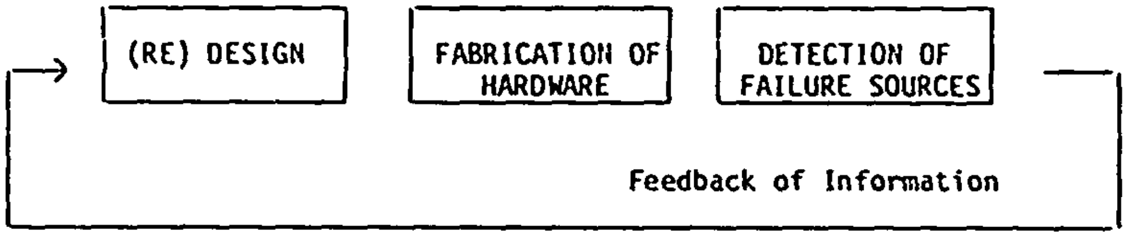


FIGURE A-4. Feedback Model

50.4 Program-related questions. The four major elements of the reliability growth feedback model can be further broken down to a set of specific program-related questions. In the following list of questions, T is used to indicate time-related questions, # is used to indicate questions related to the number of identified failure modes, and E is used to indicate questions related to the effectiveness of corrective actions.

a. Detection of Failure Sources

(1) Are the test durations and the number of systems on test adequate or excessive? T,#

If the amount of testing is too small, the number of failure modes identified will be too small to properly guide redesign effort. On the other hand, once the redesign direction is well established, but changes are not incorporated in the test hardware, not all of the newly identifie+3failure modes will be useful. In effect, we are testing "yesterday's" design once it has served its purpose of providing design guidance.

(2) To what extent can and will failed part analysis be performed to determine what failed and why it failed? T,#

For most types of equipment, this is a minimal problem, and the time required may be negligible. However, missiles and munitions (as examples) often require special instrumentation to determine what failed, and the determination of what failed may be a time-consuming process.

(3) Will early tests investigate the later life characteristics of the system? T,#

Frequently, early tests are relatively short. When longer tests are run later in the development phase, new failure modes associated with wearout may be observed. It is important that they are observed early enough in the program to allow for corrective action and verification.

b. Feedback of Information

(1) Is the feedback system responsive? and T

(2) Can information be lost by the feedback system? #

A well-designed information feedback system should experience no problems in either of these areas, but these questions must be addressed since flaws in the feedback system are as critical as flaws elsewhere in the loop and are more easily corrected.

(3) Can failures find a home in the organization? T

A significant amount of time may be expended determining the responsibility for a given failure mode.

c. Redesign Effort Based on Problems Identified (and nonreliability reasons).

(1) What general emphasis is to be placed on initiating corrective action? T

In an aggressive reliability program, each failure mode will be analyzed and corrective action at least considered. Less aggressive programs may wait for pattern failures to occur before investigating a failure mode.

(2) How severe are other design constraints? E

As other design constraints become more severe, the number of design alternatives becomes more limited. As an example, on one type of equipment approximately 30% of the design changes for reliability have involved some weight increase. This suggests that if a program for equipment of this type is severely weight constrained, approximately 30% of the usual design alternatives must be ruled out.

(3) What design changes for non-reliability reasons can be anticipated? E

This is very closely related to the above question, but it is convenient to view the restriction of reliability growth and the (possible) introduction of reliability problems separately. This question concerns the introduction of reliability problems when design changes are made for other reasons.

One approach that has been used is to treat design changes for non-reliability reasons the same as changes for reliability reasons. For example: if 40% of all design changes for reliability reasons were "unsuccessful," in that the failure mode was not essentially removed or another was introduced, we may estimate that 40% of all design changes for non-reliability reasons would cause reliability problems.

(4) Have allowances been made in terms of dollars and time for problems which will surface late in development? T,E

If a program has been planned for success at each stage, there is no margin for error; and the unexpected, yet inevitable, problems are difficult to accommodate. In the early program stages, there are usually enough variables in the program to accommodate problems. However, near the end of a development program, there may be nothing left to trade off. When planning for reliability growth, it must be recognized that it is possible to approach the end of a development effort with an identified problem, an identified "fix," but insufficient time or money to incorporate the fix.

(5) What is the strength of the design team, and what amount of design support will it receive from the reliability function? T,E

The main interests are the time required to effect design changes (on paper) and the effectiveness of the changes. These will be affected by the size and competence of the design team and also by the support it is given and the disciplines that are imposed. In general, design principles, such as the use of proven components, or the conduct of a failure mode and effects analysis increase design effectiveness at the expense of time and money.

d. Fabrication of Hardware

(1) What intervals of time can be expected between the time that component design changes are finalized and the time that the components are ready to be tested? T

Within a given system, this time can easily range from nearly zero in cases where off-the-shelf components can be used; to many months, in cases where special tooling is required. As a minimum, the longest leadtime components should be identified and from these a probable longest leadtime determined. This provides a rough estimate of the minimum leadtime required before a new design configuration can be placed on test. All leadtimes will have some impact on the practical attainment of reliability growth; but as a first cut, the long leadtime components yield the most information. It is also worthwhile noting that identification of a reliability problem in a long leadtime component may be a signal of a reliability growth problem that is not otherwise identified.

(2) What provisions are there to replace or repair components that fail on test? T,E

Ideally, replacement and repair procedures during test should duplicate those planned for the fielded equipment. However, since there may be no, or few, spares for the prototypes on test, some compromises may be necessary. Testing delays may be necessary while replacement parts are fabricated, or extraordinary repairs may be made to keep the equipment on test. When extraordinary repairs are made, the validity of some subsequently discovered failure modes may be questionable. For example, a casting that is cracked by testing may be repaired by welding, instead of being replaced as it would be in field use. If cracking subsequently occurs in another area of the casting, there may be a question whether the cracking is a result of a design deficiency or a result of residual stresses caused by welding. This doubt effectively reduces the number of identified failure modes.

50.5 Synthesis. The above questions can be used as a guide to program characteristics that will influence reliability growth. The program characteristics can then be used to synthesize the growth expectations for the program.

60. EXAMPLE

60.1 Objective. This example is intended to illustrate the general type of reasoning used to synthesize growth expectations. It does not cover a complete program and is somewhat simplified, but additional details will vary greatly from one program to the next. It considers a development of a weapon for which the majority of design changes will occur between tests. It must be emphasized that, in spite of the apparent mathematical precision, the estimates should be viewed as just ballpark figures.

60.2 Problem Statement. The first prototype weapon is to be tested for 10,000 rounds. An MRBF of 200 is anticipated, implying that 50 failures are expected during the test. From experience with similar systems in early stages of development, it is expected that the 50 failures will be in about 20 different modes. The average failure rate in a mode is expected to be

$$\frac{1}{(200)(20)} = .00025.$$

60.3 <u>Analysis of improvement in existing failure modes. What results can be expected when the second prototype is tested?</u> First, of the 20 modes expected, it is anticipated that about 18 will have design corrections attempted, and the changes are expected to reduce the failure rates by 60%. Thus, the combined failure rate expected for these modes is (18) (.40) (.00025) = .0018. For the other two failure modes, no design correction will have been made. One is expected to be a long leadtime change which won't be reflected until the third prototype, and the other is expected to be impossible to improve without exceeding the weight constraint. Thus, for these two modes, the combined failure rate is expected to be 2(.00025) = .0005. Or, for the entire system, a failure rate of .0018 + .0005 = .0023 can be expected, implying an MRBF of 1/.0023 ≅ 435, provided no new failure modes are introduced.

60.4 <u>Analysis of new failure modes anticipated. To take into consideration any new failure modes, a calculation will</u> first be made of the residual failure modes otherwise expected when testing the second prototype. The planned test duration for the second prototype is 15,000 rounds. With an MRBF of 435, about 34 failures are expected which, based on previous experience, suggests that about 15 modes will be found. Because some wearout characteristics are expected, it is anticipated that the later life test experience beyond 10,000 rounds will expose 2 new failure modes. Furthermore, an additional 2 new failure modes are expected from the dozen or so design changes motivated by non-reliability considerations. With about 15 + 2 + 2 = 19 modes expected, previous experience suggests that about 46 failures can be expected. And the expected MRBF is therefore 15000/46 ≅ 326 MRBF.

Next page is blank.

This page intentionally left blank.

APPENDIX B

GROWTH MODELS

10. INTRODUCTION

10.1 Scope

The intent of this appendix is to provide an overview of various mathematical models for reliability growth that have been proposed in the literature. This listing may be used as a guideline for choosing a candidate model for a particular application. Technical references are given for each of these models where a more complete discussion of the model may be found.

10.2 Types of models

The growth models are distinguished according to two major types as follows:

- Discrete Growth Models
- Continuous Growth Models

20. DISCRETE RELIABILITY GROWTH MODELS

20.1 General

This section describes a number of discrete reliability growth models which are currently available. Each model is briefly described including the basic assumptions that were made in deriving the models.

20.2 Model 1

Lloyd and Lipow (18) introduced a reliability growth model for a system which has only one failure mode. For each trial the probability that the system will fail if the failure mode has not been previously eliminated is assumed to be a constant. If the system does not fail, no corrective action is performed before the next trial. If the system fails, then an attempt is made to remove the failure mode from the system. The probability of successfully removing the failure mode is also assumed to be a constant for each attempt. They show that the system reliability, R_n, on the n-th trial is

$$R_n = 1-Ae^{-C(n-1)}$$

where A and C are parameters.

20.3 Model 2

Another reliability growth model was considered by Lloyd and Lipow (18) where the development program is conducted in K stages and on the i-th stage a certain number of systems are tested. The reliability growth function considered was

$$R_i = R_\infty - (\alpha/i),$$

where R_i is the system reliability during the i-th stage, R_∞ is the ultimate reliability as $i \to \infty$ and $\alpha > 0$ is a parameter. Maximum likelihood and least squares estimates of R_∞ and α are given by Lloyd and Lipow along with a lower confidence limit for R_K.

20.4 Model 3

Wolman (25) considered a situation where the system failures are classified according to two types. The first type is termed "inherent cause" and the second type is termed "assignable cause". Inherent cause failures reflect the state-of-the-art and may occur on any trial while assignable cause failures may be eliminated by corrective action, never to appear again. Wolman assumed that the number of original assignable cause failures is known and that whenever one of these modes contribute a failure, the mode is removed permanently from the system. Wolman uses a Markov-chain approach to derive the reliability of the system at the n-th trial when the failure probabilities are known.

20.5 Model 4

Barlow and Scheuer (4) considered a nonparametric model for estimating the reliability of a system during a development program. They assumed that the design and engineering changes do not decrease the system's reliability, but, unlike some other models, they do not fit a prescribed functional form to the reliablity growth. Their model is similar to Wolman's in that each failure must be classified either as inherent or assignable cause.

It is further assumed that the development program is conducted in K stages, with similar systems being tested within each stage. For each stage, the number of inherent failures, the number of assignable cause failures and the number of successes are recorded. In addition, they assumed that the probability of an inherent failure, q_0, remains the same throughout the development program and that the probability of an assignable cause failure, q_i, in the i-th stage does not increase from stage to stage of the development program. The authors obtained the maximum likelihood estimates of q_0 and of the q_i's subject to the condition that they be nonincreasing. A conservative lower confidence bound for the reliability of the system in its final configuration was also given.

20.6 Model 5

Virene (24) considered the suitability of the Gompertz equation

$$R = ab^{c^t},$$

$0 < a < 1$, $0 < b < 1$, $0 < c < 1$, for reliability growth modeling. In this equation a is the upper limit approached by the reliability R for a fixed time period as the development time $t \to \infty$. The parameters a, b and c are unknown. Virene gave estimates of these parameters and demonstrated by examples the application of this model.

20.7 Model 6

Barlow, Proschan and Scheuer (3) considered a reliability growth model which assumes that a system is being modified at successive stages of development. At stage i the system reliability (probability of success) is p_i. The model of reliability growth under which one obtains the maximum likelihood estimates of $p_1, p_2, \ldots, p_K$ assumes that

$$p_1 \leq p_2 \leq \cdots \leq p_K.$$

That is, it is required that the system reliability not be degraded from stage to stage of development. No particular mathematical form of growth is imposed on the reliability. In order to obtain a conservative lower confidence bound on p_K, it suffices to require only that

$$p_K \geq \max_{i<K} p_i.$$

That is, it is only necessary that the reliability in the latest stage of development be at least as high as that achieved earlier in the development program.

Data consist of x_i successes in n_i trials in stage i, $i=1, \ldots, K$.

A variation of this model is treated in Barlow and Scheuer. (See Section 20.5.) In that model two types of failure, inherent and assignable cause, are distinguished.

20.8 Model 7

Another reliability growth model considered by Barlow, Proschan and Scheuer (3) assumed that at stage i of development the distribution of system life length is F_i. The model of reliability growth under which the maximum likelihood estimates of $F_1(t), F_2(t), \ldots, F_K(t)$ are obtained, writing

$$\bar{F}_i(t) = 1-F_i(t)$$

is

$$\overline{F}_1(t) \leq \overline{F}_2(t) \leq \ldots \leq \overline{F}_K(t)$$

for a fixed $t \geq 0$. In order to obtain a conservative upper confidence curve on $F_K(t)$ and thereby, a conservative lower confidence curve on $\overline{F}_K(t)$ for all non-negative values on t, it suffices only to require that

$$\overline{F}_K(t) > \max_{i<K} \overline{F}_i(t)$$

for all $t \geq 0$. That is, the probability of system survival beyond any time t in the latest stage of development is at least as high as that achieved earlier in the development program.

20.9 Model 8

Singpurwalla (23) considered an approach to reliability growth analysis of discrete data involving the use of time series methods. Since a time series can be defined simply as, "...a set of observations generated sequentially in time" it is straightforward to formulate the growth process as the following time series problem: on a complex system which is undergoing successive developmental changes, tests are performed to monitor progress and to determine whether reliability requirements are being met. The outcome of each test is judged to be either a success or a failure. In particular, at the end of the j-th stage, n_j independent tests are conducted of which v_j are deemed to be successful. If we denote the reliability of the system at the end of the j-th stage by p_j, then v_j is binomially distributed with parameters n_j and p_j. Let $\hat{p}_j$ be an estimator of p_j, j = 1,2, ..., M. Given estimates for p_j, j = 1,2, ... M, we can apply time series methods, (1) to determine whether p_j is increasing with j, (2) to obtain a good estimate of the probability of success at the present stage of testing (p_M), and (3) to obtain forecasts of p at future stages, M + 1, M + 2 ... .

In particular, the methods proposed by Box and Jenkins (6) have been found to be powerful and flexible enough for application to many fields. Singpurwalla (23) is a specific application of this approach to reliability growth problems. The Box-Jenkins Autoregressive-Integrated Moving Average (ARIMA) model/approach has the following major advantages:

(a) No specific model need be selected in advance. The data themselves lead to selection of a specific model within the very broad and flexible class of ARIMA models.

(b) Models with either deterministic or stochastic indications of growth can be fitted to data. Normally the deterministic model should be used only in cases where the growth process is well understood and controlled. This is particularly true if the model is being used to forecast future reliability.

(c) The Box-Jenkins methodology has a built-in theory of forecasting, as well as techniques to obtain numerical forecasts.

It must be recognized that his approach has some disdvantages as well. For example, data from a relatively large number of stages must be available, i.e., M should be of the order of 20 or so before meaningful conclusions can be drawn in most cases. If the process is a complex one, it is possible that $M \geq 50$ will be required. Another disadvantage is that the methodology cannot be applied in a cookbook fashion. Considerable judgment is required and it is possible to derive very inappropriate conclusions.

30. CONTINUOUS RELIABILITY GROWTH MODELS

30.1 General

The previous section discussed situations where a device or system either operated successfully when called upon or failed to perform its mission, i.e., a go/no-go situation. The other broad category which must be considered is the repairable system which must operate successfully over periods of time which cannot be regarded as fixed and hence, cannot be divided into a go/no-go categorization. In this case, we must be concerned with the sequence of successive times-between-failures of the system. If the system is improving (as a result of design fixes, debugging of bad parts, better repair procedures, or any other reason) then the successive times-between-failures (inter-failure times) will tend to increase. Reversals will occur for many reasons, including inappropriate design fixes, damage caused by previous repairs, changing environmental stresses, or even sampling variability. Hence, it may not be obvious that growth is occurring without some sort of analysis. Moreover, even if the presence of growth can be verified by inspection, it usually will be necessary to use some systematic technique(s) to estimate the rate at which growth is occurring or to forecast future changes in reliability. Some of the following models are based on the non-homogeneous Poisson process which is described in 30.1.1. The discussion for models 13-17 are from reference 20.

30.1.1 Poisson Processes

A stochastic process $\{N(t), t \geq 0\}$ is said to be a counting process if $N(t)$ represents the total number of events which have occurred in the interval $(0,t)$. The counting process $\{N(t), t \geq 0\}$ is said to be a homogeneous Poisson process (HPP) if

(1) $N(0) = 0$,

(2) $\{N(t), t \geq 0\}$ has independent increments, and

(3) The number of events (in our context, failures) in any interval of length $t_2 - t_1$ has a Poisson distribution with mean $\rho(t_2 - t_1)$.

That is, for all $t_2 > t_1 \geq 0$,

$$P\ \{N(t_2) - N(t_1) = n\} = \frac{e^{-\rho(t_2-t_1)}\ \{\rho(t_2-t_1)\}^n}{n!}$$

for $n \geq 0$.

From condition (3) it follows that

$$E\ \{N(t_2 - t_1)\} = \rho(t_2 - t_1)$$

where the constant, ρ, is the rate of occurrence of failures. It can be shown that the successive times-between-failures of the HPP defined above are independent and identically distributed exponential random variables.

The non-homogeneous Poisson process (NHPP) differs from the homogeneous Poisson process (HPP) only in that the intensity function varies with time rather than being a constant. That is, conditions (1) and (2) are retained and condition (3) is modified to be:

(3a) The number of failures in any interval (t_1, t_2) has a Poisson distribution with mean $\int_{t_1}^{t_2} \rho(t)\, dt$

That is, for all $t_2 > t_1 \geq 0$

$$P\ \{N(t_2) - N(t_1) = n\} = \frac{\left(\int_{t_1}^{t_2} \rho(t)\, dt\right)^n e^{-\int_{t_1}^{t_2} \rho(t) dt}}{n!}$$

for $n \geq 0$.

From (3a) it follows that

$$E\ \{N(t_2) - N(t_1)\} = \int_{t_1}^{t_2} \rho(t)\, dt$$

30.2 Model 9

Duane (11) analyzed data for several systems developed by General Electric in an effort to determine if any systematic changes in reliability improvement occurred during development for these systems. His analysis revealed that for these systems, the cumulative failure rate fell close to a straight line when plotted on log-log scale.

Let $N(t)$ denote the number of system failures by time t, $t > 0$. The observed cumulative failure rate $C(t)$ is $C(t) = N(t)/t$. The log-log plots imply that $\log C(t)$ is approximately a straight line. That is, $\log C(t) = \delta - \alpha \log t$, or $C(t) = \gamma t^{-\alpha}$, where $\gamma = e^{\delta}$. It follows also that $N(t) = \delta t^{1-\alpha}$.

The change per unit time of $N(t)$, $r(t) = \frac{d}{dt} N(t) = \gamma(1-\alpha)t^{-\alpha}$.

Duane interpreted this as the current failure rate. In this context, the reciprocal of $r(t)$, $m(t) = [\gamma(1-\alpha)t^{-\alpha}]^{-1}$, may be interpreted as the current or instantaneous MTBF. This is Duane's postulate which is a deterministic learning curve formulation of reliability growth.

When the test time t is the cumulative test time for the program, then the log-log property of the cumulative failure rate, $C(t)$, indicates an overall trend for reliability growth or an idealized type pattern. Section 5.2.6 provides appropriate methods for construction and interpretation of the idealized growth curve and test phase reliability when $C(t)$ is linear on log-log scale.

30.3 Model 10

Crow (9) considered a model (called the AMSAA model) which can be used for tracking reliability growth within test phases. This approach assumes that within a test phase, reliability growth can be modeled as a NHPP. It also assumes that based on the failures and test time within a test phase, the cumulative failure rate is linear on log-log scale. This is a local, within test phase pattern for reliability growth comparable to the global pattern noted by Duane (11). Let t be the test time from the beginning of the test phase and let $N(t)$ denote the number of system failures by time t. It follows that the expected value of $N(t)$ can be written as $E[N(t)] = \lambda t^{\beta}$.

The AMSAA model assumes that the test phase reliability growth follows the NHPP with mean value function $\theta(t) = \lambda t^{\beta}$ and intensity function $\rho(t) = \lambda\beta t^{\beta-1}$. This model allows for the development of rigorous statistical procedures useful for reliability growth tracking. The AMSAA model is thoroughly considered in Appendix C.

30.4 Model 11

A NHPP model proposed by Cox and Lewis (7) is $\rho(t) = e^{\alpha+\gamma t}$. The parameters α and ν can be estimated from test data and a goodness of fit test applied for this model. For additional details and background information, see Ascher and Feingold (2).

30.5 Model 12

Lewis and Shedler (17) extended the Cox-Lewis model (Model 11) by developing estimation techniques for the exponential polynomial model for powers up to 10, i.e., for models of the form

$$\rho(t) = \exp\,(\alpha_0 + \alpha_1 t + \ldots + \alpha_{10} t^{10}).$$

30.6 Model 13

The IBM model, Rosner (21), assumes explicitly that: (1) there are random (constant intensity function) failures occurring at rate λ, and (2) there are a fixed but unknown, number of non-random design, manufacturing and workmanship defects present in the system at the beginning of testing. Let $N(t)$ be the number of non-random type defects remaining at time $t \geq 0$. This model makes the intuitively plausible assumption that the rate of change of $N(t)$ with respect to time is proportional to the number of non-random defects remaining at t. That is,

$$d\,N(t)/dt = -K_2 N(t)$$

and hence

$$N(t) = e^{-K_2 t+c}$$

Now if we denote the unkown number of non-random failures present at $t = 0$ by K_1 then

$$N(t) = K_1 e^{-K_2 t} \quad t > 0,\ K_1,\ K_2 > 0.$$

Defining $V(t)$ to be the expected cumulative number of failures up to time t then

$$V(t) = \lambda t + K_1\,(1-e^{-K_2 t}). \qquad (1)$$

Thus, the expected cumulative number of failures by time t is the expected number of random failures by time t plus the expected number of non-random failures removed by time t. It should be noted that $V(0) = 0$ as expected. Moreover as $t \to \infty$, $V(t) \to \lambda t + K_1 + \lambda t \to \infty$, as expected.

Because of the non-linearity of the model (1) the estimation of λ, K_1 and K_2 must be accomplished by iterative means.

In addition to this model being "plausible" the most interesting feature is the ability of the model to predict the time when the system/equipment is "q" fraction debugged (i.e., q fraction of the original K_1 nonrandom failures have been removed, $0 < q < 1$). The number of non-random defects removed by time t is clearly

$$N(0) - N(t) = K_1 - K_1 e^{-K_2 t}$$

and hence the fraction (of K_1 initial non-random defects) removed by time t is

$$q = \frac{K_1 - K_1 e^{-K_2 t}}{K_1} = 1 - e^{-K_2 t} \qquad (2)$$

Thus having estimated K_2, say $\hat{K}_2$, we can find the time at which $q = 0.95$ of the non-random defects have been removed by solving (2) for $t_{0.95}$. That is,

$$t_{0.95} = \frac{-\ln 0.05}{\hat{K}_2} .$$

In general, for arbitrary q, $0 < q < 1$ the time by which the system/equipment is q fraction debugged is

$$t_q = \frac{-\ln (1-q)}{\hat{K}_2} . \qquad (3)$$

Equation (3) is a powerful tool because it can be used to help determine the length of development testing, or, the debugging period.

Another important feature of this model is that the number of nonrandom failures remaining at time t can be estimated and of course is $\hat{K}_1 e^{-\hat{K}_2 t}$. The estimate of λ, say $\hat{\lambda}$, gives the estimate of the long-run achievable MTBF.

In the above model the dependent variable was the expected cumulative number of failures by time t. In all of the following models the dependent variable is the cumulative mean time between failures, Y(t) where

$$Y(t) = \frac{t}{\text{Total No. of Failures in } (0,t)}$$

30.7 Model 14

Suppose that K is used to denote the limiting value of Y(t) as $t \to \infty$ and suppose the rate of growth dY(t)/dt is jointly proportional to the remaining growth (namely K-Y(t)) and some growth function g(t). Thus

$$dY(t)/dt = [K-Y(t)]\ g(t).$$

Taking g(t), the growth function, to be a constant, say $K_2 > 0$, then the solution to the differential equation is easily seen to be

$$Y(t) = K\ (1-K_1 e^{-K_2 t}),\ t > 0.$$

This may be referred to as the exponential-single term power series model.

Here $K_1 > 0$ is an intercept parameter arising as a constant of integration.

The growth rate (i.e., dY(t)/dt) is largest at t = 0 is monotonically decreasing in t so that

$$\lim_{t \to \infty} [dY(t)/dt] = 0$$

It is entirely plausible that the growth rate is largest at t = 0 and decreases to 0 as $t \to \infty$. This model is also extremely flexible because it has three parameters

K: The limit of cumulative MTBF.

K_1: When t = 0, $Y(0) = K\ (1-K_1)$. Thus $K\ (1-K_1)$ may be thought of as the initial MTBF of the system/equipment when $0 < K_1 < 1$. K_1 may also be thought of as the growth potential.

K_2: The growth function; constant in this case.

The disadvantage of this model is clear enough. Like the IBM model it has three parameters and is non-linear in t; nor can it be transformed to a linear function of t. Thus the least squares estimates of K, K_1, and K_2 must be obtained by iterative procedures. More details on this model can be found in Perkowski and Hartvigsen (19).

30.8 Model 15

A model proposed by Lloyd and Lipow (18) supposes that the growth rate is inversely proportional to the square of time t, i.e.,

$$dY(t)/dt = K^2/t^2,\ K_2 > 0.$$

Then clearly,

$$Y(t) = K - K_2/t.$$

Here K is a constant of integration but it should be noticed that

$$\lim_{t \to 0} Y(t) = K$$

and thus K is the limiting value of cumulative MTBF.

The parameter K_2 is a growth rate parameter which also affects the location of the curve. Since Y(t) cannot be negative and

$$\lim_{t \to 0} Y(t) = -\infty$$

we must define

$$Y(t) = 0,\ 0 \leq t < K_2K.$$

This definition provides a time period (0, K_2/K) when the cumulative MTBF is 0. This may be realistic for some systems.

By making the change of variable t' = 1/t we see that

$$Y(t') = K - K_2t'$$

and thus Y(t') is linear in t' with slope K_2 and intercept K which means the parameters K and K_2 can be easily estimated by the usual least squares methods.

30.9 Model 16

Aroef (1) assumed that the growth rate is jointly proportional to the growth achieved at t, i.e., Y(t), a constant multiplier (growth rate parameter) K_2 and inversely proportional to t^2. That is,

$$dY(t)/dt = K_2\ Y(t)/t^2.$$

This differential equation has the solution

$$Y(t) = K e^{-K_2/t}$$

Since $\lim_{t \to 0} Y(t) = K$ the reliability growth limit in cumulative MTBF is K. Also

$$\lim_{t \to 0} Y(t) = 0.$$

Since

$$\ln Y(t) = \ln K - K_2/t,$$

letting

$$t' = 1/t,$$

$$\ln Y(t') = \ln K - K_2 t'$$

and usual linear least squares methods can be used to estimate the constants K and K_2.

30.10 Model 17

The last model considered is the simple exponential model:

$$Y(t) = K e^{K_2 t}, \; K > 0,\; K_2 > 0.$$

$Y(0) = K$ which is the "initial" cumulative MTBF. Since $\ln Y(t) = \ln K + K_2 t$ then the linear least square method can be used to fit the constants.

APPENDIX C

THE AMSAA RELIABILITY GROWTH MODEL

10. MODEL DESCRIPTION

10.1 Introduction

The US Army Materiel Systems Analysis Activity (AMSAA) employs the Weibull process to model reliability growth during a development test phase. This model adequately represents the improvement in reliability during development for a wide variety of systems. It is applicable to systems for which usage is measured on a continuous scale, for example, time in hours or distance in miles. Also, for high reliability and a large number of trials, the model may be used for one-shot systems.

Development test programs are generally conducted on a phase by phase basis. For each test phase it is typical for the test data to be compiled and a reliability evaluation made. It is important to note that the AMSAA Reliability Growth Model is designed for tracking the reliability within a test phase and not across test phases. This model evaluates the reliability growth that results from the introduction of design fixes into the system during test and not the reliability growth that may occur at the end of a test phase due to delayed fixes.

10.2 Basis of the Model

Figure C.1 illustrates reliability growth on a phase by phase basis. The AMSAA Model addresses the reliability growth within a particular test phase.

The beginning of a test phase will be denoted as time $t = 0$. Within the test phase let $0 < S_1 < S_2 < \ldots < S_k$ denote the cumulative test times on the system when design modifications are made. See Figure C.2. Between the times when design changes are made on the system, the failure rate can generally be assumed to be constant. Let λ_i denote the constant failure rate during the i-th time period $[S_{i-1}, S_i)$ between modifications. See Figure C.3.

Based on the constant failure rate assumption, the number of failures N_i during the i-th time period has the Poisson distribution with mean $\lambda_i(S_i-S_{i-1})$. That is,

$$\text{Prob } [N_i=n] = \frac{[\lambda_i(S_i-S_{i-1})]^n \, e^{-\lambda_i(S_i-S_{i-1})}}{n!}$$

$n = 0, 1, 2 \ldots$. Also, the constant failure rate assumption during $[S_{i-1}, S_i)$ implies that for this interval the times between successive

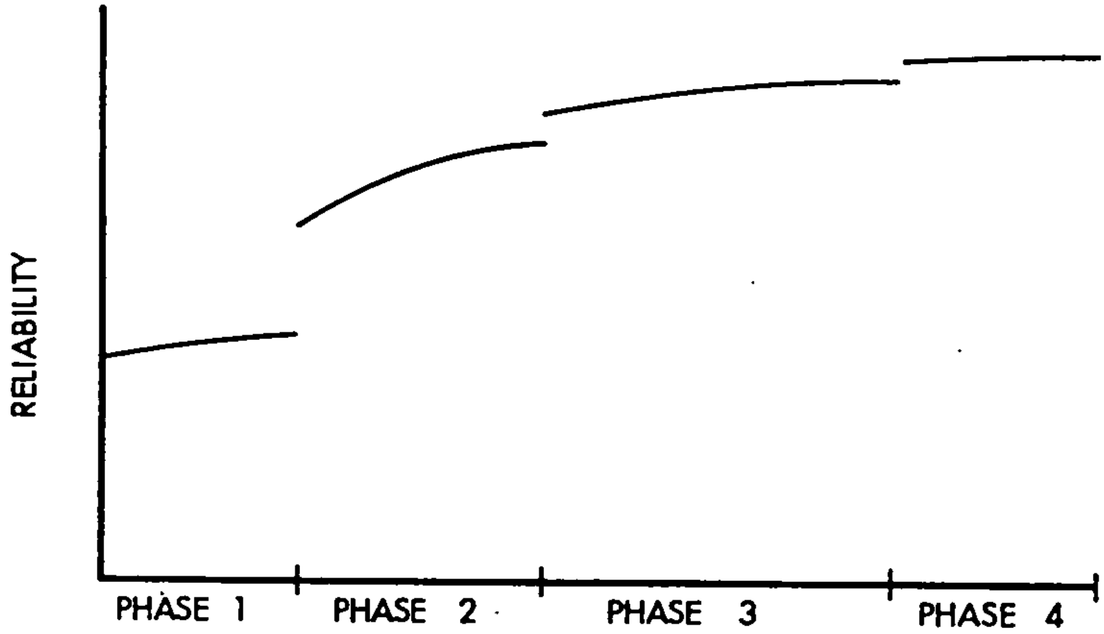


Figure C-1 Phase-by-Phase Reliability Growth.

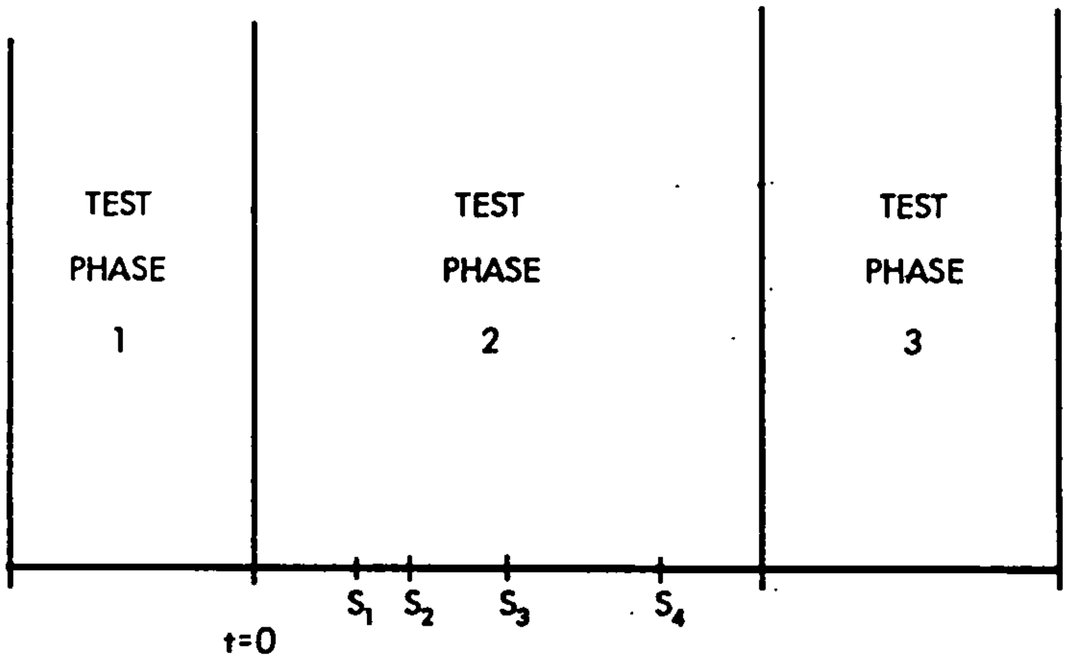


Figure C-2 Times of Design Modifications for Test Phase 2.

failures follow the exponential distribution

$$F(x) = 1 - e^{-\lambda_i x}, \quad x > 0.$$

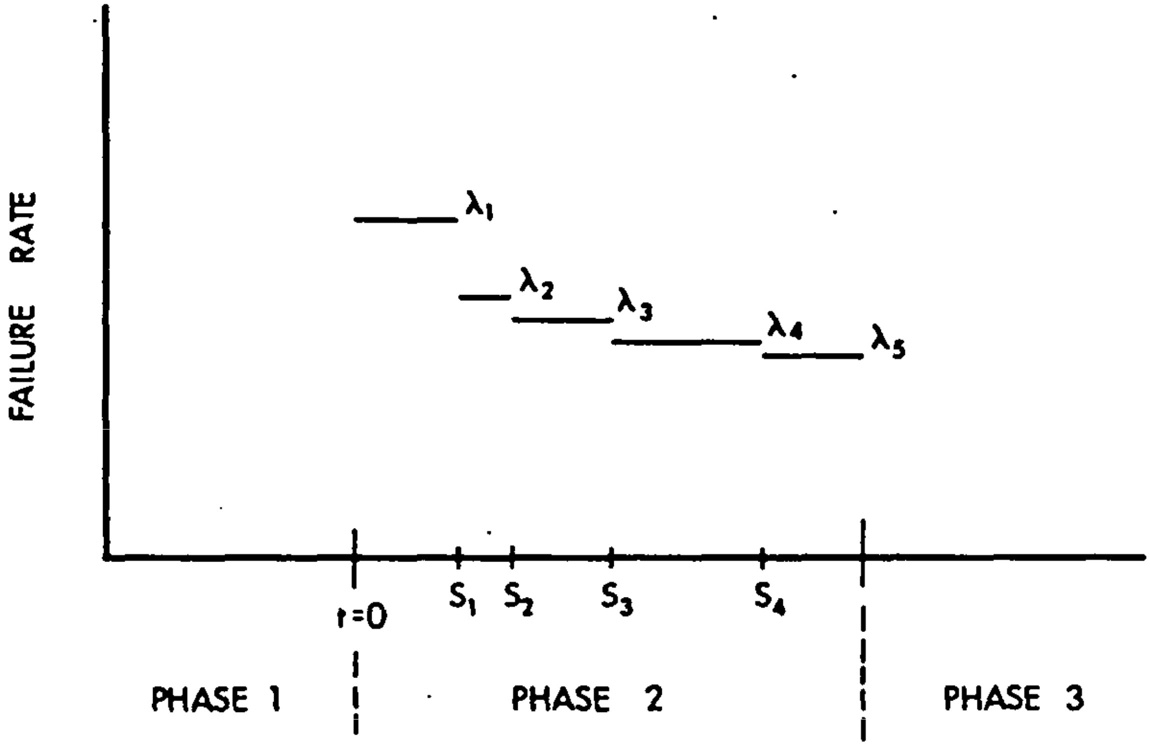
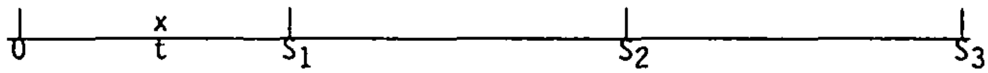


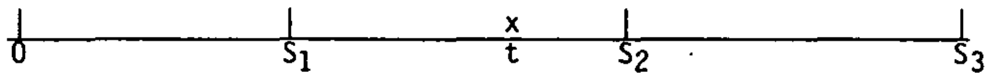
Figure C-3 Failure Rates Between Modifications.

During developmental programs, more than one prototype is often tested. If the prototypes have the same basic configuration between modifications, then under the constant failure rate assumption, the times S_i may be considered as cumulative test time to the i-th modification. Also, on a cumulative time scale, N_i is the total number of failures experienced by all systems during $[S_{i-1}, S_i)$.

Let t denote cumulative test time and let N(t) be the total number of system failures by time t. If t is in the first interval, then N(t) has the



Poisson distribution with mean $\lambda_1 t$. Suppose t is in the second interval. In



this case N(t) is the number of failures N_1 in the first interval plus the number of failures in the second interval between S_1 and t. The failure rate for the first interval is λ_1 and the failure rate for the second interval is λ_2. Therefore, the mean of N(t) is the mean of N_1 which is $\lambda_1 S_1$ plus the mean number of failures from S_1 to t, which is $\lambda_2(t-S_1)$. That is, N(t) has mean $\theta(t) = \lambda_1 S_1 + \lambda_2 (t-S_1)$.

When the failure rate is constant (homogeneous) over a test interval, then N(t) is said to follow a homogeneous Poisson process with mean of the form λt. When the failure rates change with time (nonhomogeneous), e.g., from interval 1 to interval 2, then under certain conditions, N(t) is said to follow a nonhomogeneous Poisson process. For the situation under consideration, N(t) would follow the nonhomogeneous Poisson process with mean value function

$$\theta(t) = \int_0^t \rho(y)dy, \text{ where}$$

$\rho(y) = \lambda_i$, $y \in [S_{i-1}, S_i)$. That is, for any t,

$$\text{Prob}[N(t) = n] = \frac{[\theta(t)]^n e^{-\theta(t)}}{n!} ,$$

n = 0, 1, 2, ... For example, when t is in the first test phase, $\theta(t) = \lambda_1 t$. When t is in the second test phase, $\theta(t) = \lambda_1 S_1 + \lambda_2(t-S_1)$, etc.

The integer-valued process $\{N(t), t > 0\}$ is called a nonhomogeneous Poisson process with intensity function $\rho(t)$. The physical interpretation of $\rho(t)$ is that for Δt infinitesimally small, $\rho(t)\Delta t$ is approximately the probability of a system failure in the interval $(t, t+\Delta t)$. If $\rho(t) = \lambda$, a constant for all t, then this probability is not changing with test time. If $\rho(t)$ is decreasing $(\lambda_1 > \lambda_2 > \lambda_3 \ldots)$, then the failure probability $\rho(t)\Delta t$ is decreasing, implying reliability growth. For $\rho(t)$ increasing $(\lambda_1 < \lambda_2 < \lambda_3 < \ldots)$, system reliability is deteriorating.

From a practical point of view it is advantageous to approximate the intensity function $\rho(t)$ by a continuous, parametric function since all test data during a test phase may then be pooled to estimate these parameters.

The AMSAA Model assumes that $\rho(t)$ may be approximated by the parametric form $\rho(t) = \lambda\beta t^{\beta-1}$, $t > 0$, $\lambda > 0$, $\beta > 0$, which is recognized as being the Weibull failure rate function. See Figure C.4. This implies that the mean number of failures by time t is $\theta(t) = \lambda t^{\beta}$. A motivation for this form of $\rho(t)$, which is based on a learning curve pattern for the cumulative failure rate is given below.

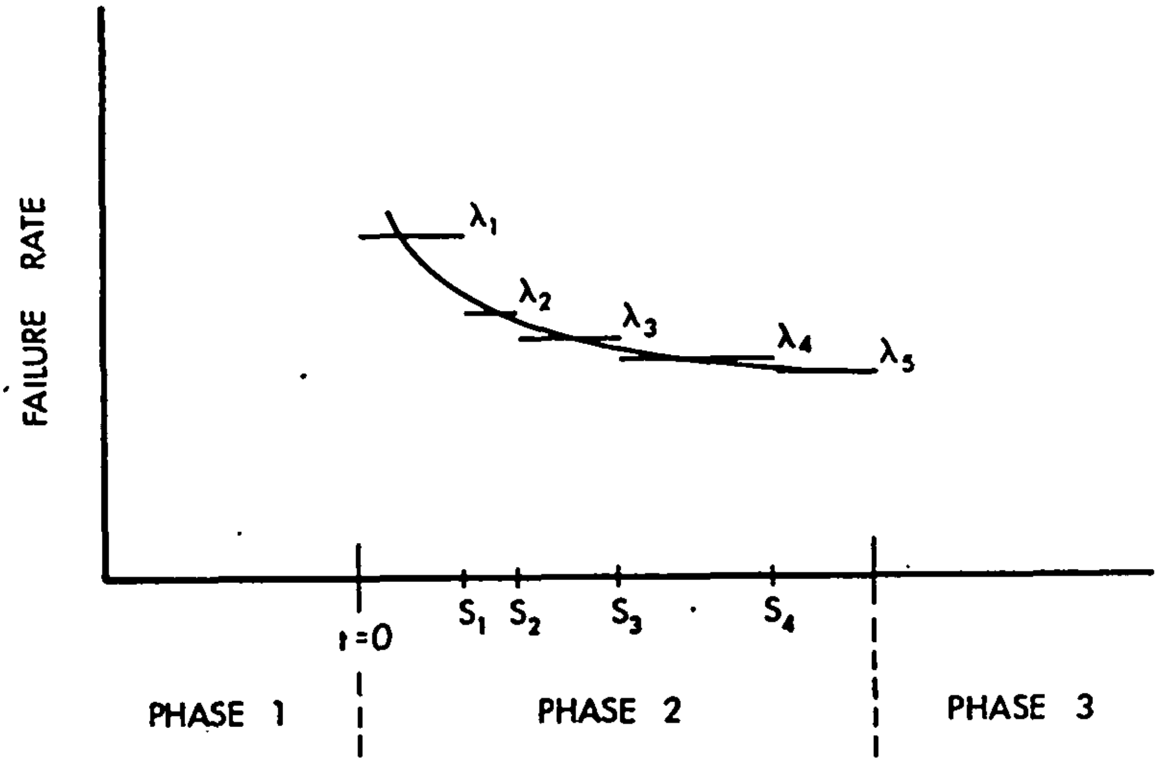


Figure C-4 Parametric Approximation to Failure Rates Between Modifications.

The observed cumulative failure rate $C(t)$ is defined as $C(t) = N(t)/t$. Suppose that within a test phase $C(t)$ is linear with respect to t on a log-log scale. This local pattern for reliability growth within a test phase is analogous to the global, idealized growth pattern observed by Duane (reference 11) for systems during their development programs. See Figure C.5. Equating $C(t)$ to its expected value and assuming an exact linear relationship on log-log scale, it follows that $E[C(t)] = \lambda t^{\kappa}$ where κ represents the slope of the local pattern on a log-log scale. Hence, $E[N(t)] = \lambda t^{\beta}$, for $\beta = \kappa + 1$. Thus, within the test phase the expected number of system failures by time t is λt. The instantaneous failure rate, $\rho(t)$, for the system is the change per unit time of $E[N(t)]$. Therefore, $\rho(t) = \frac{d}{dt} E[N(t)] = \lambda\beta t^{\beta-1}$.

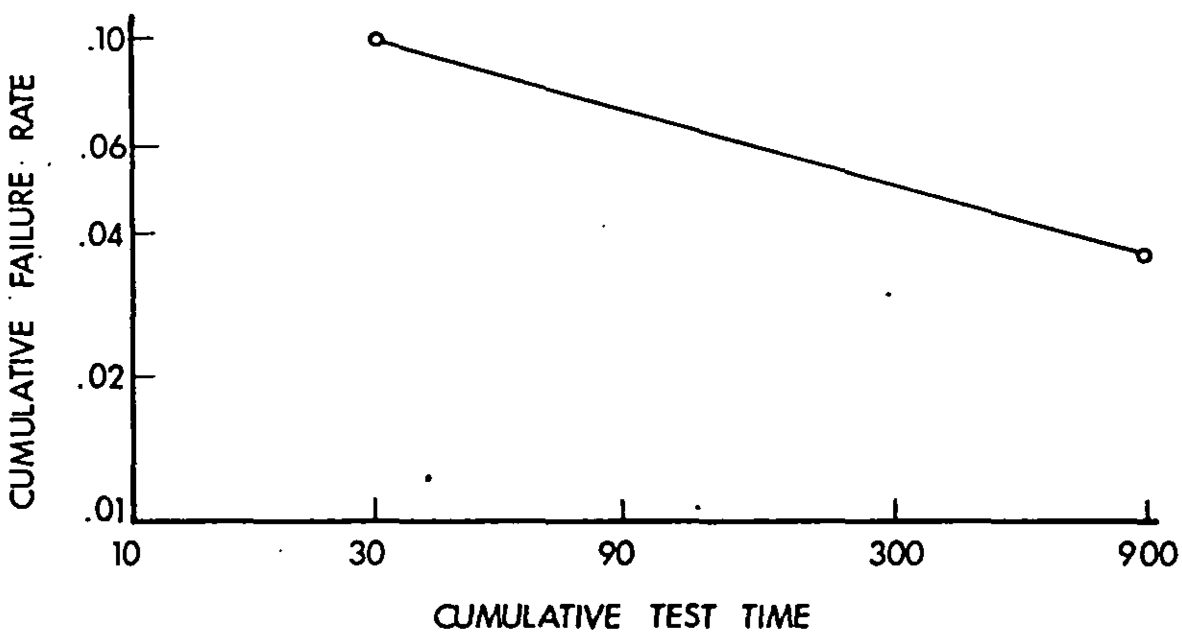


Figure C-5 Log-Log Plot Within Test Phase.

The AMSAA Reliability Growth Model assumes that system failures during a development testing phase follow the nonhomogeneous Poisson process with Weibull intensity $(t) = \lambda\beta t^{\beta-1}$, where $\lambda > 0$, $\beta > 0$. For $\beta = 1$, $\rho(t) = \lambda$, the homogeneous Poisson or exponential case. For $\beta < 1$, $\rho(t)$ is decreasing, implying reliability growth. When $\beta > 1$, $\rho(t)$ is increasing, indicating a deterioration in system reliability. It is important to note that the model assumes a Poisson process with Weibull intensity function $\rho(t) = \lambda\beta t^{\beta-1}$, and not the Weibull distribution. Therefore, statistical procedures for the Weibull distribution do not apply for this model.

With a failure rate or intensity function that may change with test time, the nonhomogeneous Poisson process provides a basis for describing the reliability growth process within a test phase. With the AMSAA Model estimates can be made for assessment purposes, confidence bounds can be found, and the data can be subjected to an objective goodness-of-fit test.

10.3 The Model

The AMSAA Reliability Growth Model assumes that within a test phase failures are occurring according to a nonhomogeneous Poisson process. It is further assumed that the failure rate or intensity of failures during the test phase can be represented by the Weibull function $\rho(t) = \lambda\beta t^{\beta-1}$ where $\lambda > 0$, $\beta > 0$ are parameters and t is cumulative test time. Under this model the function $m(t) = [\lambda\beta t^{\beta-1}]^{-1}$ is interpreted as the instantaneous MTBF of the system at time t. When t corresponds to the total cumulative time for the system, then m(t) is the demonstrated MTBF or the MTBF of the system in its present configuration. See Figure C.6. Crow formulated this process as a model to describe the pattern of reliability growth in reference 8. Other references on this process include Kempthorne and Folks (15), Englehardt and Bain (12), Bassin (5), Crow (9) (10), Finkelstein (13), and Lee and Lee (16).

10.3.1 Cumulative Number of Failures. The total number of failures, N(t), accumulated on all test items in cumulative test time t is a random variable with the Poisson distribution. The probability that exactly n failures occur between the initiation of testing and total test time t is

$$P\,(N(t) = n) = \frac{[\theta(t)]^n e^{-\theta(t)}}{n!} .$$

in which $\theta(t)$ is the mean value function; that is, the expected number of failures expressed as a function of test time. To describe the reliability growth process this function is of the form $\theta(t) = \lambda t^{\beta}$ in which λ and β are positive parameters.

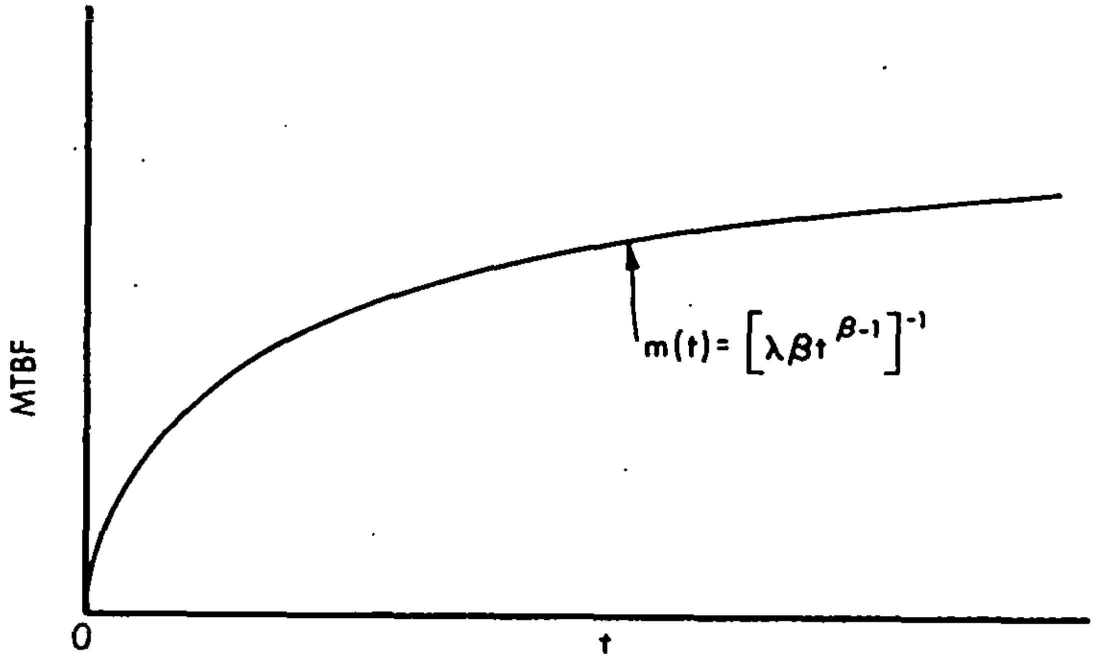


Figure C-6 Test Phase Reliability Growth Based on AMSAA Model.

10.3.2 Number of Failures in an Interval. The number of failures occurring in the interval from test time a until test time b is a random variable having the Poisson distribution with mean

$$\theta(b)-\theta(a) = \lambda(b^{\beta}-a^{\beta})$$

The number of failures occurring in any interval is statistically independent of the number of failures in any interval which does not overlap the first interval. Only one failure can occur at any instant according to the Weibull process model.

10.3.3 Intensity Function. For the reliability growth process the intensity function is

$$\rho(t) = \lambda\beta t^{\beta-1}.$$

The probability of the occurrence of a failure between time t and time t + h is approximately $\rho(t)h$ if the increment h is sufficiently small. The intensity function is sometimes called the failure rate; however, this concept is different from that of the failure rate or hazard rate of a life distribution. Caution should be exercised so that the two ideas are not confused. The parameter λ is called a scale parameter because it depends upon the unit of measurement chosen for t. The parameter β is of prime importance because it characterizes the shape of the graph of the intensity function. If β is equal to one, the intensity function is constant. In that case the reliability of the system is not changing since the times between successive failures are independent identically distributed random variables with an exponential distribution with mean λ^{-1}. If β is not equal to one, the times between successive failures are not identically distributed and do not have exponential distributions. For a development process during which the system reliability improves, the shape parameter β is less than one. In this case the expected number of failures in an interval of fixed length decreases as its starting point increases. In a poorly managed reliability program improper design changes can result in degradation of system reliability. This situation is characterized by values of the shape parameter β greater than one. This indicates that the number of failures expected in a fixed increment of time is increasing with time.

10.3.4 Mean Time Between Failures. At time t_0 the intensity of failure is $\rho(t_0) = \lambda\beta t_0^{\beta-1}$. In practice it is generally assumed that if no improvements are incorporated into the system after time t_0, then failures would continue at the constant rate $\rho(t_0) = \lambda\beta t_0^{\beta-1}$ with further testing. That is, if no additional modifications are made on the system after time t_0, then future failures would follow an exponential distribution with mean $m(t_0) = [\lambda\beta t_0^{\beta-1}]^{-1}$. The function $m(t) = [\lambda\beta t^{\beta-1}]^{-1}$ is interpreted as the instantaneous MTBF of the system at time t and represents the system reliability growth under this model.

20. RELIABILITY GROWTH ASSESSMENT

20.1 Graphical estimation. Plots derived from the failure data provide a graphic description of test results. They furnish the analyst a means to examine the nature of the data. Graphical methods can also be used to obtain rough estimates of the reliability parameters of interest in the reliability growth process. Two types of graphs are described below. The first tells the analyst if growth is obviously demonstrated by the data. The second method goes further since it provides rough estimates of the two parameters in the expression for the intensity function.

20.1.1 Average failure rate plots. Construction of a plot of the average failure rates observed during testing yields a crude approximation of the intensity function. To construct such a plot divide the elapsed test time into at least three nonoverlapping intervals. These nonoverlapping intervals can be of unequal length. Next calculate the frequency of occurrence of failures within each interval by dividing the number of failures in the interval by its length. Plot the failure frequency as a horizontal line at the appropriate ordinate. The line should extend over the abscissas corresponding to time within the interval. Any significant increasing or decreasing trend in the intensity function should be apparent from this plot.

20.1.2 Cumulative failure plots. A graph of the observed cumulative number of failures plotted against cumulative test time on full logarithmic paper furnishes crude estimates of the parameters which describe the intensity function. Taking logarithms in the expression for the mean value function yields the result

$$\log m(t) = \log \lambda + \beta \log t.$$

Therefore, the expression for the mean value function is represented by a straight line on full logarithmic paper. A line drawn to fit the data points representing the cumulative number of failures at the time of each failure occurrence is a suitable approximation of the true line. The ordinate of the point on the line corresponding to t equal to one is an estimate of λ. The actual slope of the line as measured with a ruler yields an estimate of the shape parameter β. Alternative methods include the plotting of the cumulative numbers of failures divided by cumulative test time or the reciprocal of that quantity. If either of those methods are used, the method for estimating the parameters is slightly more complicated.

20.1.3 Examples of graphical estimation. The following data are used to demonstrate the graphical estimation procedures. Two prototypes of a mechanical system are tested concurrently with the incorporation of design changes. The first system runs 132.4 hours, and the second runs 167.6 hours. The time on each system and the cumulative test time at each failure are listed below. An asterisk denotes the failed system.

	#1 Hours	#2 Hours	Cumulative Hours		#1 Hours	#2 Hours	Cumulative Hours
1	2.6*	.0	2.6	15	60.5	37.6*	98.1
2	16.5*	.0	16.5	16	61.9*	39.1	101.1
3	16.5*	.0	16.5	17	76.6*	55.4	132.0
4	17.0*	.0	17.0	18	81.1	61.1*	142.2
5	20.5	.9*	21.4	19	84.1*	63.6	147.7
6	25.3	3.8*	29.1	20	84.7*	64.3	149.0
7	28.7	4.6*	33.3	21	94.6*	72.6	167.2
8	41.8*	14.7	56.5	22	104.8	85.9*	190.7
9	45.5*	17.6	63.1	23	105.9	87.1*	193.0
10	48.6	22.0*	70.6	24	108.8*	89.9	198.7
11	49.6	23.4*	73.0	25	132.4	119.5*	251.9
12	51.4*	26.3	77.7	26	132.4	150.1*	282.5
13	58.2*	35.7	93.9	27	132.4	153.7*	286.1
14	59.0	36.5*	95.5	End	132.4	167.6	300.0

Although the occurrence of two failures at exactly 16.5 hours is not possible under the assumption of the model, such data can result from rounding in order to construct an average failure rate plot. The test time is divided into 50 hour intervals. The number of failures in each interval is divided by the length of the interval to give the average failure rates shown in Figure C-7. The data indicate that the intensity function is decreasing. The cumulative number of failures is plotted as a function of test time in Figure C-8. The line in the figure is an approximation to the data points. The ordinate of the line corresponding to 1 hour is .49. This is an estimate of λ. The slope is measured as 7.3/10.0. Thus, .73 is an estimate of β. This procedure thus quantifies the trend of reliability improvement. The estimate of the intensity function at 300 hours is $(.49)(.73)(300)^{-.27}$ or .077 failures per hour. If no further changes are made, the estimated mean time between failures is 1/.077 or 13 hours. These estimates are satisfactory for a quick analysis of the data; however, the statistical estimates described in 20.2 provide a more precise description of the growth process.

20.2 Statistical estimation. Modeling reliability growth as a nonhomogeneous Poisson process permits the assessment of the demonstrated reliability performance by statistical procedures. The method of maximum likelihood provides estimates of the scale parameter λ and the shape parameter β. These estimates are used in estimation of the intensity function. In accordance with 10.2.4, the reciprocal of the current value of the intensity function is the mean time between failures that the system would exhibit in the absence of further improvements. Procedures for point estimation and interval estimation of mean time between failures are described below. The data employed in the estimation consist of failure times from testing terminated at a given time or from testing terminated at the occurrence of a specified number of

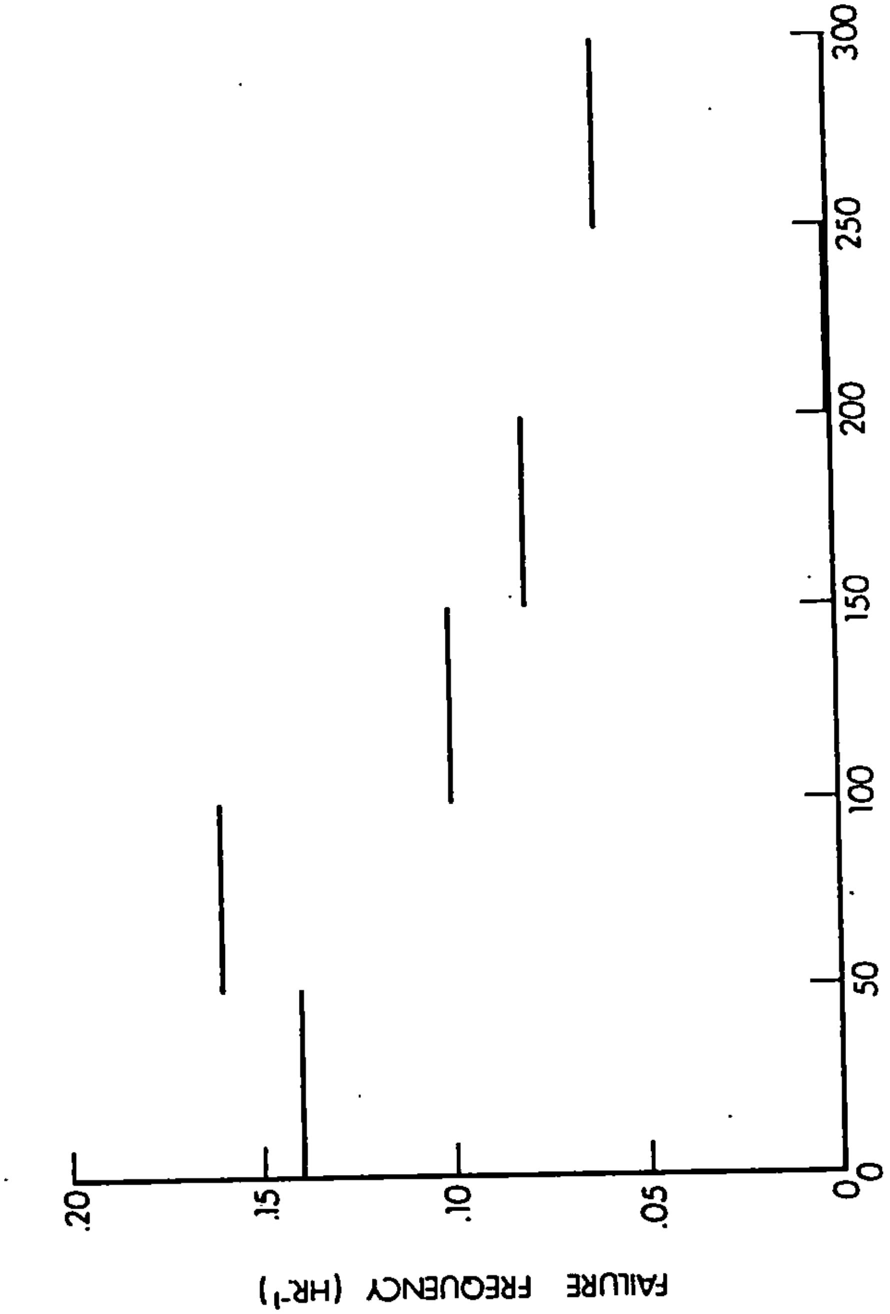


Figure C-7 Average Failure Frequency Plot.

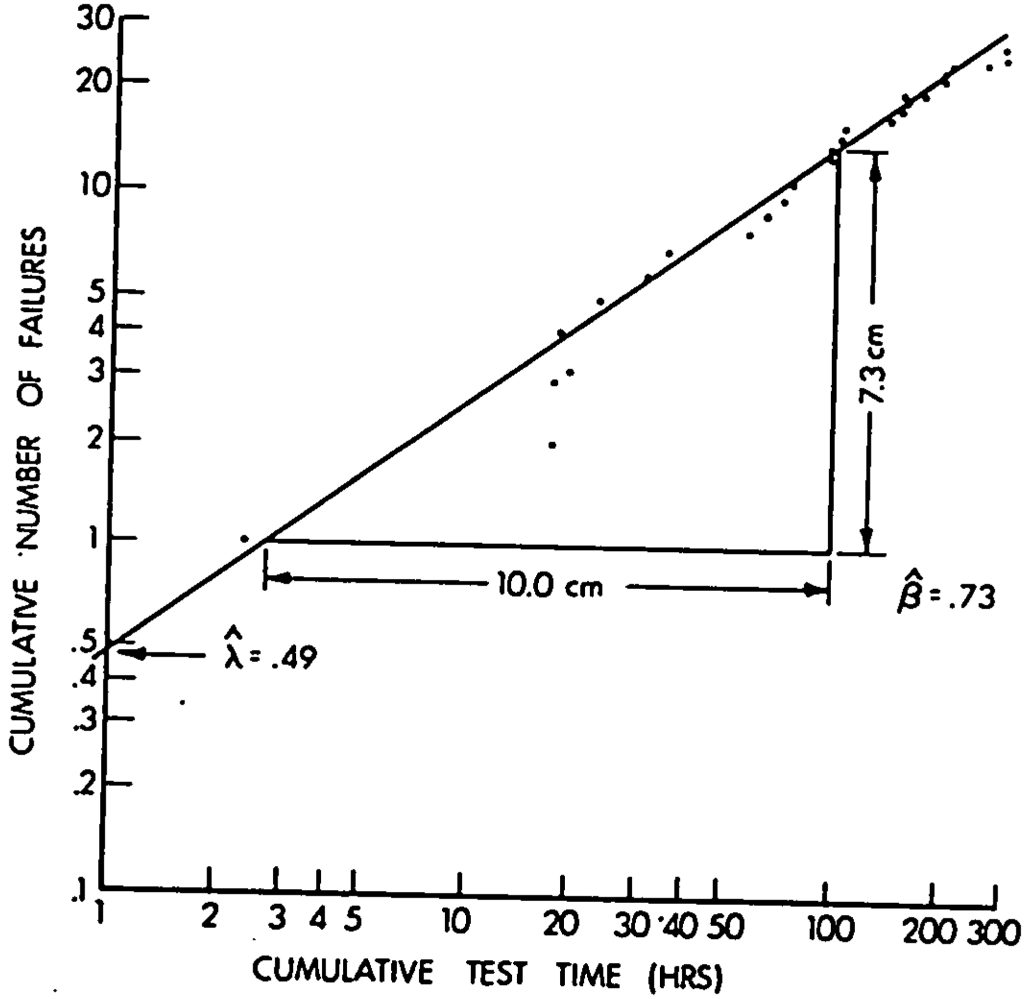


Figure C-8: Cumulative Failure Plot.

failures. The procedures vary slightly for these two types of tests. A goodness of fit test to determine whether the model is appropriate to describe the data is also described below. If the exact times of failure occurrence are unknown, it may still be possible to utilize the reliability growth model. This is the case when inspections are conducted to uncover hidden failures. Procedures to use in that instance are in Section 20.2.3.

20.2.1 Time terminated testing. The procedures described in this section are to be used to analyze data from tests which are terminated at a predetermined time or tests which are in progress with data available through some time. The required data consists of the cumulative test time on all systems at the occurrence of each failure as well as the accumulated test time. To calculate the cumulative test time of a failure occurrence it is necessary to sum the test time on every system at that instant. The data then consists of the N successive failure times $X_1 < X_2 < \ldots < X_N$ which occur prior to the accumulated test time T.

20.2.1.1 Point estimation. The method of maximum likelihood provides point estimates of the parameters of the reliability growth process. The estimate of the shape parameter β is

$$\hat{\beta} = \frac{N}{N \ln T - \sum_{i=1}^{N} \ln X_i}$$

Subsequently, the scale parameter λ is estimated by $\hat{\lambda} = N/T^{\hat{\beta}}$. It follows that for any time t the intensity function is estimated by $\hat{\rho}(t) = \hat{\lambda}\hat{\beta}t^{\hat{\beta}-1}$. In particular, this holds for T, the accumulated test time. For small sample sizes use of the unbiased estimator $\bar{\beta}$ given in 20.2.1.3 is advisable.

The reciprocal of $\hat{\rho}(T)$ provides an estimate of the mean time between failures which could be anticipated if the system configuration remains as it is at time T. If the reliability program is expected to continue without any shift in emphasis or environment, then the intensity function may be projected into the future to predict the benefit of continued attempts to improve reliability. Although the estimators use all failure occurrences, the model is effectively self purging. The estimator $\hat{\rho}(T)$ can be written as $\hat{\beta}(N/T)$. Note that N/T would be the estimate of the intensity function for a homogeneous Poisson process. Hence the fraction $(1-\hat{\beta})$ of the failures are effectively eliminated.

20.2.1.2 Interval estimation. Interval estimates provide a measure of the uncertainty regarding the demonstration of reliability by testing. For the reliability growth process the parameter of primary interest is the mean time between failures that the system would exhibit after the initiation of production. The probability distribution of the point estimate of the intensity function at the end of the test is the basis for the interval estimate of the true value of the intensity function at that time. The values in Table C-I facilitate computation of confidence interval estimates for the mean time between failures. The table provides two-sided interval estimates on the ratio of the true MTBF to the estimated MTBF for several values of the confidence coefficient. If the number of failures is N and γ is the selected confidence coefficient, then the appropriate tabular values are L_N,γ and U_N,γ. The interval estimate of MTBF is

$$\frac{L_N,\gamma}{\hat{\rho}(T)} \leq \text{MTBF} \leq \frac{U_N,\gamma}{\hat{\rho}(T)}$$

Because the number of failures has a discrete probability distribution, these interval estimates are conservative, that is, the actual confidence coefficient is slightly larger than the stated confidence coefficient.

20.2.1.3 Goodness of Fit. The null hypothesis that a nonhomogeneous Poisson process with an intensity function of the form $\lambda\beta t^{\beta-1}$ properly describes the reliability growth of a particular system is tested by the use of a Cramér-von Mises statistic. An unbiased estimate of the shape parameter is used to calculate that statistic. This estimate of β is

$$\bar{\beta} = \frac{N-1}{N}\hat{\beta}$$

for a time terminated test with N failure occurrences. The estimate $\hat{\beta}$ is described in Section 20.2.1.1. The goodness of fit statistic is

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

in which the failure times must be ordered so that $0 \leq X_1 \leq X_2 \leq \ldots \leq X_N$. The null hypothesis is rejected if the statistic C_M^2 exceeds the critical value for the level of significance selected by the analyst. Critical values of C_M^2 for the .20, .15, .10, .05 and .01 levels of significance (α) are in Table C-II. That table is indexed by a parameter labeled M. For time terminated testing M is equal to N, the number of failures. If the test rejects the reliability growth model, an examination of the data may reveal the reason for the lack of fit. Possible causes of rejection include the occurrence of more than one failure at the same time or the occurrence of a discontinuity in the intensity function. In the first case, an appropriate procedure may be to group the data as explained in 20.2.2. In the latter case the data should be treated as described in 20.2.4.

20.2.1.4 Example of time terminated test. The data used in 20.1.3 will be used to demonstrate the statistical estimation procedures.

The point estimate of β is

$$\hat{\beta} = \frac{27}{27 \ln 300 - (\ln 2.6 + \ln 16.5 + \ldots + \ln 286.1)} = .716$$

Thus, as the graphical techniques indicated, the AMSAA model estimates that reliability is improving. The estimate of the scale parameter is

$$\hat{\lambda} = \frac{27}{300^{.716}} = .454.$$

The estimated intensity function is

$$\hat{\rho}(t) = .325t^{-.284}$$

Figure C-9 shows the graph of this equation superimposed on the average failure frequencies. The intensity function at the end of the test is estimated as .0645. The point estimate of MTBF at the end of the 300 hour test is 15.5 hours. The interval estimate of MTBF with a 90 percent confidence coefficient is (.636/.0645, 1.682/.0645), that is 9.9 to 26.1 hours. These results together with the tracking curve and the planned growth curve are shown in Figure C-10. The level of significance chosen to test the goodness of fit is .05. The critical value is .218 as determined by interpolation in Table C-II for M equal to 27. The unbiased estimate of β is .690. This is used to calculate the Cramer-von Mises statistic. The value of that statistic is .091. Since this is below the critical value we accept the hypothesis that the AMSAA model is appropriate.

20.2.2 Failure terminated testing. The procedures described in this section are applicable to tests which are terminated upon the accumulation of a specified number of failures. The procedures are only slightly different from those used for time terminated testing. The data consist of the N successive failure times $X_1 < X_2 < \ldots < X_N$ expressed in terms of cumulative test time.

20.2.2.1 Point estimation. The method of maximum likelihood furnishes point estimates of the shape parameter β and the scale parameter λ. The estimate of β is

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

Note that this is equivalent to the estimate for time terminated testing with the test time equal to the time of occurrence of the last failure. The scale parameter λ is estimated by

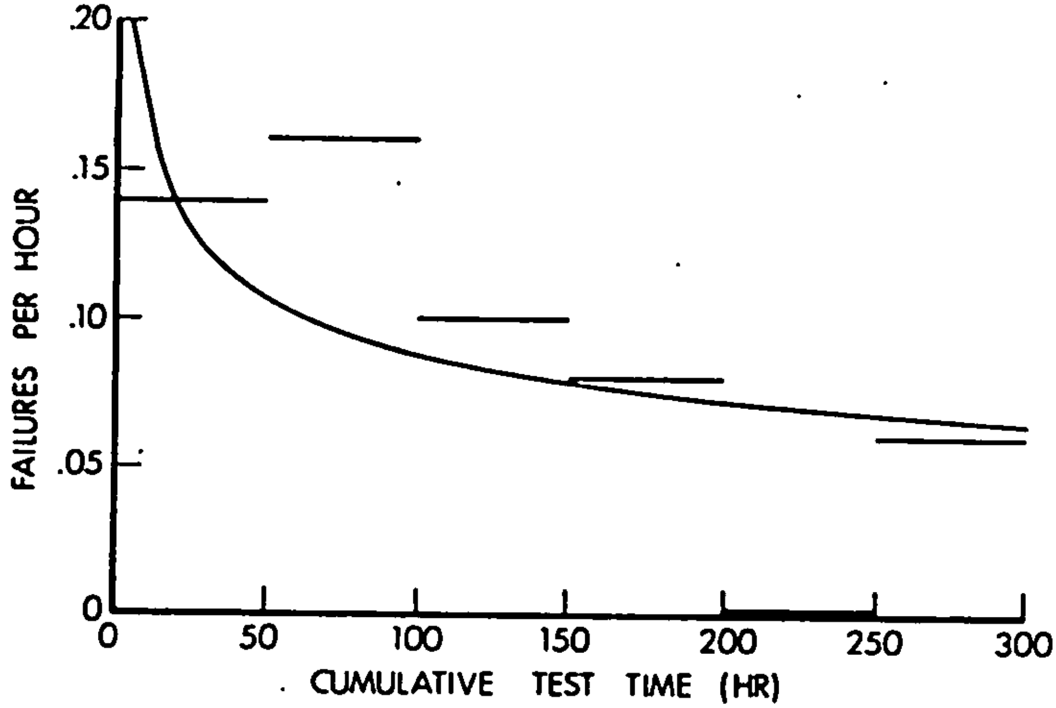


Figure C-9. Estimated Intensity Function.

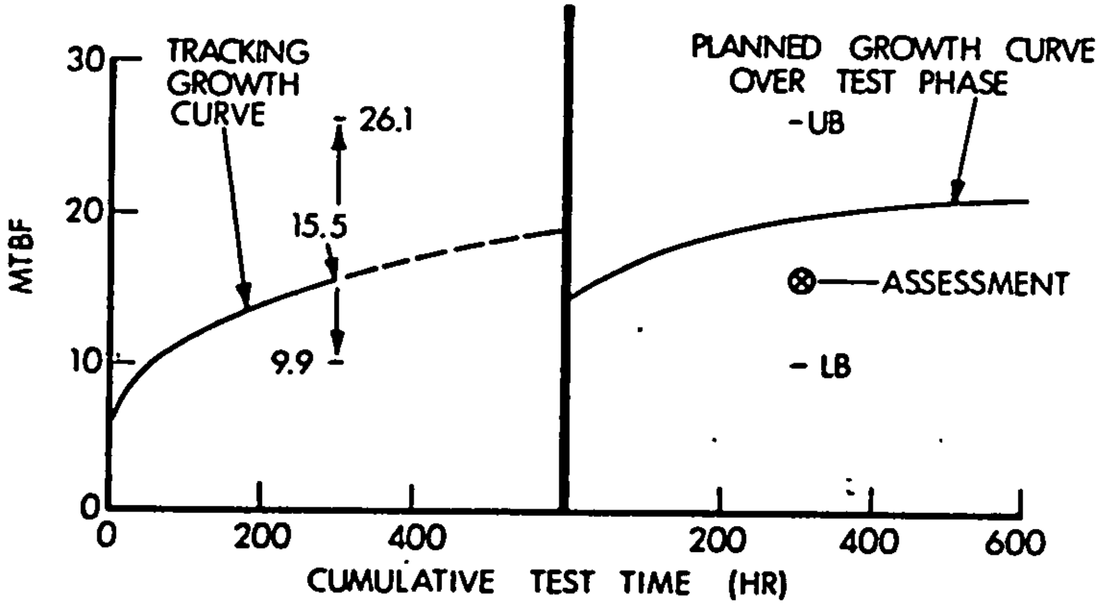


Figure C-10 Planned and Tracking Growth Curves.

$$\hat{\lambda} = \frac{N}{T^{\hat{\beta}}}$$

as before. The intensity function and mean time between failures are estimated as in 20.2.1.1. For small sample sizes use of the unbiased estimator $\bar{\beta}$ given in 20.2.2.3 is advisable.

20.2.2.2 Interval estimation. An interval estimate of the mean time between failures that the system would exhibit in the absence of further changes is also available for the case of failure terminated testing. Table C-III provides factors for the construction of two-sided interval estimates of the MTBF for several values of the confidence coefficient γ. The small value corresponding to the number of failures and desired confidence coefficient is divided by the point estimate of the intensity function at the end of the test to yield the lower limit of the interval. Division of the U value by the intensity function estimate provides the upper limit.

20.2.2.3 Goodness of Fit. The hypothesis that the AMSAA model is appropriate can be tested using a Cramér-von Mises statistic. It is important to note the difference in the calculations from those for time terminated testing. An unbiased estimate of the shape parameter given by

$$\bar{\beta} = \frac{N-2}{N}\hat{\beta}$$

is used in the calculation of the goodness of fit statistic. The parameter for indexing that statistic is M = N-1, where N is the number of failures. The Cramér-von Mises statistic is then

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

Table C-II provides critical values for use in the test. The model is deemed inappropriate if the statistic C_M^2 exceeds the critical value for some specified level of significance α.

20.2.2.4 Example of failure terminated test. In this example the data are from testing of a tank system. This illustrates that the basis for measurement of test length is not necessarily time. In this case, test duration is measured in miles accumulated. The test terminates at the occurrence of the twenty-sixth failure. Failures occur at accumulated mileages of 1, 57, 187, 252, 310, 485, 693, 720, 727, 779, 1028, 1561, 1766, 1793, 1938, 2030, 2065, 2289, 2423, 2560, 2879, 3086, 3458, 3626, 4252, and 4582 miles. The point estimate of the shape parameter is β = .626, which indicates substantial reliability growth. The scale parameter estimate is $\hat{\lambda} = .132$. At the end of the test the estimated intensity function is .00355 failures per mile. This corresponds to mean miles between failure equal to 281 miles. The interval estimate of MTBF with a 95 percent confidence coefficient is (.6333/ .00355, 1.919/.00355) or from 178 to 541 miles as determined by use of table C-III. This indicates uncertainty in the MMBF estimated from this amount of testing. The goodness of fit is tested at the .10 level of significance. The critical value for M equal to 25 is determined to be .172 by interpolation in table C-II. The Cramér-von Mises statistic is .058 which indicates that the model represents the data quite well.

20.2.3 Grouped data. It may happen that an event included within the scope of the definition of the term "failure" does not preclude the operation of the equipment. It is possible that such events are not uncovered until a thorough inspection is conducted. In this case the exact time of the failure is unknown; however, one can presume that it happened in the interval since the last inspection. The total number of failures in the interval between inspections is therefore the sum of the number of failures detected at the time of occurrence and the number of failures found in the inspection. Such totals for each interval can be used to estimate reliability growth in accordance with the AMSAA model if there are at least three intervals.

20.2.3.1 Point estimation from grouped data. The data consist of the total number of failures in each of K intervals of test time. The first interval starts at test time zero. The intervals do not have to be of equal length. Denote the number of failures in the interval from t_{i-1} to t_i by N_i. By convention t_0 is equal to zero. The maximum likelihood estimate of the shape parameter β is the value which satisfies

$$\sum_{i=1}^{K} N_i \left[\frac{t_i^{\hat{\beta}} \ln t_i - t_{i-1}^{\hat{\beta}} \ln t_{i-1}}{t_i^{\hat{\beta}} - t_{i-1}^{\hat{\beta}}} - \ln t_K \right] = 0,$$

in which $t_0 \log t_0$ is defined as zero. This nonlinear equation can be solved by any of several iterative procedures. The scale parameter estimate is

$$\hat{\lambda} = \frac{\sum_{i=1}^{K} N_i}{t_K^{\hat{\beta}}}$$

which corresponds to the result for testing when all failure times are known with the exception that the estimate of β is calculated differently. Point estimates of the intensity function and the mean time between failures are calculated as in 20.2.1.1.

20.2.3.2 Goodness of Fit. A chi-square goodness of fit test can be used to test the hypothesis that the AMSAA reliability growth model adequately represents a set of grouped data. The expected number of failures in the interval from t_{i-1} to t_i is approximated by

$$e_i = \hat{\lambda}(t_i^{\hat{\beta}} - t_{i-1}^{\hat{\beta}}).$$

Adjacent intervals may have to be combined so that the expected number of failures in any combined interval is at least five. Let the number of intervals after this combination be K and let the number of failures in the i-th interval be N_i. Furthermore, let e_i be the expected number of failures in the i-th new interval. Then the statistic

$$\chi^2 = \sum_{i=1}^{K} \frac{(N_i - e_i)^2}{e_i}$$

is approximately distributed as a χ^2 random variable with K-2 degrees of freedom. The critical values for this statistic can be found in tables of the chi-square distribution.

20.2.3.3 Example of estimation from grouped data. In this example an aircraft has scheduled inspections at intervals of twenty flight hours. All failures occurring between consecutive inspections are combined with those discovered during the inspection at the end of an interval to determine the total number of failures assigned to that interval. For the first 100 hours of flight testing the results are:

Start Time	End Time	No. of Failures
0	20	13
20	40	16
40	60	5
60	80	8
80	100	7

There are 49 failures accumulated. Solution of the equation for β yields an estimate of .753 for the shape parameter. The scale parameter estimate is 1.53. At the end of the test, the intensity function estimate is .369 failures per flight hour. If no further changes are made, the mean time between failures is estimated as 2.7 flight hours. The goodness of fit statistic is 5.4. The critical value for a χ^2 statistic with 3 degrees of freedom at the .05 level of significance is 7.8. Since the statistic is less than the critical value the applicability of the model is accepted.

20.2.4 Discontinuities in the intensity function. The simultaneous introduction of several design changes, a change in emphasis in the reliability program, or some other factor may cause an abrupt change in the intensity function. Such a jump should be detected by a departure from linearity in the full logarithmic plot of cumulative failures, a large change in the level of the average failure frequency, or rejection of the model by a goodness of fit test.

20.2.4.1 Location of discontinuity. The cumulative test time at which a discontinuity has occurred can be determined by inspection from graphs of cumulative failures or average failure frequency. The methods presented above can then be used to estimate the intensity function by use of different parameters for the period before the jump and for the period after the jump. That is, if the discontinuity occurs at time T_J, then the intensity function is estimated by

$$\hat{\rho}(t) = \hat{\lambda}_1 \hat{\beta}_1 t^{\hat{\beta}_1 - 1} \qquad 0 < t \leq T_J$$

$$= \hat{\lambda}_2 \hat{\beta}_2 (t - T_J)^{\hat{\beta}_2 - 1} \qquad t > T_J$$

in which λ_1 and β_1 are estimated only from failures on or before T_J and λ_2 and β_2 are estimated from those failures occurring after T_J. Only the second of these equations is needed to estimate the currently achieved value of the intensity function.

20.2.4.2 Example of discontinuity in intensity function. The following 56 failure times are recorded for a mechanical system: .3, .5, .6, 1.0, 2.1, 2.2, 3.5, 4.2, 4.2, 5.3, 8.1, 8.3, 9.7, 9.8, 10.3, 10.7, 12.2, 13.4, 13.9, 14.3, 14.4, 15.1, 18.2, 20.3, 21.2, 21.8, 22.4, 24.8, 26.6, 28.3, 29.0, 29.3, 29.5, 29.9, 30.6, 33.2, 33.3, 33.4, 34.4, 34.4, 34.6, 36.9, 37.5, 37.7, 38.3, 39.2, 40.3, 41.3, 43.1, 52.4, 81.0, 100.4, 101.0, 162.2, 165.2, and 188.1. The test is terminated at 200 test hours.

Calculation of the Cramér-von Mises statistic as described in 20.2.1.3 yields a value of .401. This exceeds the .05 level of significance critical value of .220; hence, the applicability of the model is rejected. The full logarithmic plot of cumulative failures is Figure C-11. The reliability growth trend changes abruptly at about 40 test hours. Therefore, the intensity function after 40 hours is of interest. The 10 failures subsequent to 40 test hours are used to estimate the parameters λ_2 and β_2. The estimate $\hat{\lambda}_2$ and $\hat{\beta}_2$ are .942 and .465, respectively.

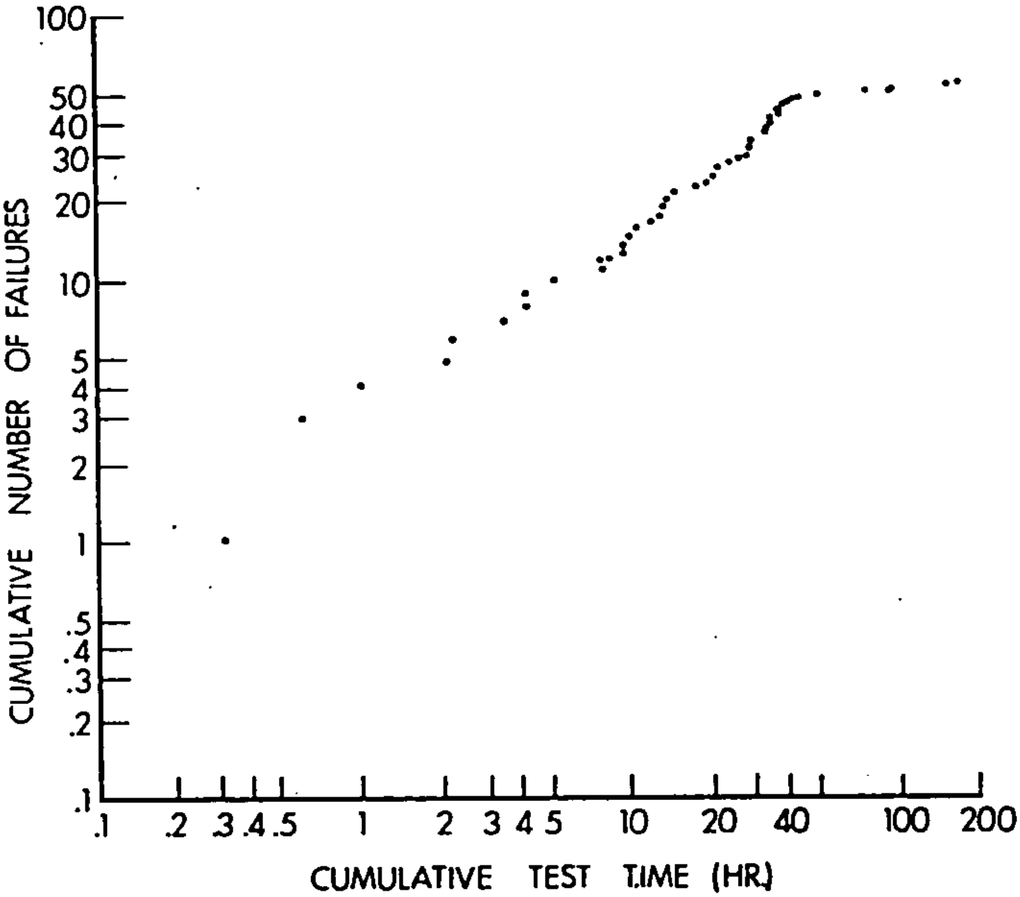


Figure C-11. Discontinuity in Intensity Function.

N \ γ	.80 L	.80 U	.90 L	.90 U	.95 L	.95 U	.98 L	.98 U
2	.261	18.66	.200	38.66	.159	78.66	.124	198.7
3	.333	6.326	.263	9.736	.217	14.55	.174	24.10
4	.385	4.243	.312	5.947	.262	8.093	.215	11.81
5	.426	3.386	.352	4.517	.300	5.862	.250	8.043
6	.459	2.915	.385	3.764	.331	4.738	.280	6.254
7	.487	2.616	.412	3.298	.358	4.061	.305	5.216
8	.511	2.407	.436	2.981	.382	3.609	.328	4.539
9	.531	2.254	.457	2.750	.403	3.285	.349	4.064
10	.549	2.136	.476	2.575	.421	3.042	.367	3.712
11	.565	2.041	.492	2.436	.438	2.852	.384	3.441
12	.579	1.965	.507	2.324	.453	2.699	.399	3.226
13	.592	1.901	.521	2.232	.467	2.574	.413	3.050
14	.604	1.846	.533	2.153	.480	2.469	.426	2.904
15	.614	1.800	.545	2.087	.492	2.379	.438	2.781
16	.624	1.759	.556	2.029	.503	2.302	.449	2.675
17	.633	1.723	.565	1.978	.513	2.235	.460	2.584
18	.642	1.692	.575	1.933	.523	2.176	.470	2.503
19	.650	1.663	.583	1.893	.532	2.123	.479	2.432
20	.657	1.638	.591	1.858	.540	2.076	.488	2.369
21	.664	1.615	.599	1.825	.548	2.034	.496	2.313
22	.670	1.594	.606	1.796	.556	1.996	.504	2.261
23	.676	1.574	.613	1.769	.563	1.961	.511	2.215
24	.682	1.557	.619	1.745	.570	1.929	.518	2.173
25	.687	1.540	.625	1.722	.576	1.900	.525	2.134
26	.692	1.525	.631	1.701	.582	1.873	.531	2.098
27	.697	1.511	.636	1.682	.588	1.848	537	2.068
28	.702	1.498	.641	1.664	.594	1.825	.543	2.035
29	.706	1.486	.646	1.647	.599	1.803	.549	2.006
30	.711	1.475	.651	1.631	.604	1.783	.554	1.980
35	.729	1.427	.672	1.565	.627	1.699	.579	1.870
40	.745	1.390	.690	1.515	.646	1.635	.599	1.788
45	.758	1.361	.705	1.476	.662	1.585	.617	1.723
50	.769	1.337	.718	1.443	.676	1.544	.632	1.671
60	.787	1.300	.739	1.393	.700	1.481	.657	1.591
70	.801	1.272	.756	1.356	.718	1.435	.678	1.533
80	.813	1.251	.769	1.328	.734	1.399	.695	1.488
100	.831	1.219	.791	1.286	.758	1.347	.722	1.423

TABLE C-I - CONFIDENCE INTERVALS FOR MTBF
FROM TIME TERMINATED TEST

For N > 100,

$$L \doteq \left(1 + Z_{.5+\frac{\gamma}{2}} / \sqrt{2N}\right)^{-2} \qquad U \doteq \left(1 - Z_{.5+\frac{\gamma}{2}} / \sqrt{2N}\right)^{-2}$$

in which $Z_{.5+\frac{\gamma}{2}}$ is the $(.5 + \frac{\gamma}{2})$-th percentile of the standard normal distribution.

M \ α	.20	.15	.10	.05	.01
2	.138	.149	.162	.175	.186
3	.121	.135	.154	.184	.23
4	.121	.134	.155	.191	.28
5	.121	.137	.160	.199	.30
6	.123	.139	.162	.204	.31
7	.124	.140	.165	.208	.32
8	.124	.141	.165	.210	.32
9	.125	.142	.167	.212	.32
10	.125	.142	.167	.212	.32
11	.126	.143	.169	.214	.32
12	.126	.144	.169	.214	.32
13	.126	.144	.169	.214	.33
14	.126	.144	.169	.214	.33
15	.126	.144	.169	.215	.33
16	.127	.145	.171	.216	.33
17	.127	.145	.171	.217	.33
18	.127	.146	.171	.217	.33
19	.127	.146	.171	.217	.33
20	.128	.146	.172	.217	.33
30	.128	.146	.172	.218	.33
60	.128	.147	.173	.220	.33
100	.129	.147	.173	.220	.34

For M > 100 use values for M = 100.

TABLE C-II - CRITICAL VALUES FOR CRAMER-VON MISES GOODNESS OF FIT TEST

N \ γ	.80 L	.80 U	.90 L	.90 U	.95 L	.95 U	.98 L	.98 U
2	.8065	33.76	.5552	72.67	.4099	151.5	.2944	389.9
3	.6840	8.927	.5137	14.24	.4054	21.96	.3119	37.60
4	.6601	5.328	.5174	7.651	.4225	10.65	.3368	15.96
5	.6568	4.000	.5290	5.424	.4415	7.147	.3603	9.995
6	.6600	3.321	.5421	4.339	.4595	5.521	.3815	7.388
7	.6656	2.910	.5548	3.702	.4760	4.595	.4003	5.963
8	.6720	2.634	.5668	3.284	.4910	4.002	.4173	5.074
9	.6787	2.436	.5780	2.989	.5046	3.589	.4327	4.469
10	.6852	2.287	.5883	2.770	.5171	3.286	.4467	4.032
11	.6915	2.170	.5979	2.600	.5285	3.054	.4595	3.702
12	.6975	2.076	.6067	2.464	.5391	2.870	.4712	3.443
13	.7033	1.998	.6150	2.353	.5488	2.721	.4821	3.235
14	.7087	1.933	.6227	2.260	.5579	2.597	.4923	3.064
15	.7139	1.877	.6299	2.182	.5664	2.493	.5017	2.921
16	.7188	1.829	.6367	2.144	.5743	2.404	.5106	2.800
17	.7234	1.788	.6431	2.056	.5818	2.327	.5189	2.695
18	.7278	1.751	.6491	2.004	.5888	2.259	.5267	2.604
19	.7320	1.718	.6547	1.959	.5954	2.200	.5341	2.524
20	.7360	1.688	.6601	1.918	.6016	2.147	.5411	2.453
21	.7398	1.662	.6652	1.881	.6076	2.099	.5478	2.390
22	.7434	1.638	.6701	1.848	.6132	2.056	.5541	2.333
23	.7469	1.616	.6747	1.818	.6186	2.017	.5601	2.281
24	.7502	1.596	.6791	1.790	.6237	1.982	.5659	2.235
25	.7534	1.578	.6833	1.765	.6286	1.949	.5714	2.192
26	.7565	1.561	.6873	1.742	.6333	1.919	.5766	2.153
27	.7594	1.545	.6912	1.720	.6378	1.892	.5817	2.116
28	.7622	1.530	.6949	1.700	.6421	1.866	.5865	2.083
29	.7649	1.516	.6985	1.682	.6462	1.842	.5912	2.052
30	.7676	1.504	.7019	1.664	.6502	1.820	.5957	2.023
35	.7794	1.450	.7173	1.592	.6681	1.729	.6158	1.905
40	.7894	1.410	.7303	1.538	.6832	1.660	.6328	1.816
45	.7981	1.378	.7415	1.495	.6962	1.606	.6476	1.747
50	.8057	1.352	.7513	1.460	.7076	1.562	.6605	1.692
60	.8184	1.312	.7678	1.407	.7267	1.496	.6823	1.607
70	.8288	1.282	.7811	1.367	.7423	1.447	.7000	1.546
80	.8375	1.259	.7922	1.337	.7553	1.409	.7148	1.499
100	.8514	1.225	.8100	1.293	.7759	1.355	.7384	1.431

TABLE C-III - CONFIDENCE INTERVALS FOR MTBF FROM FAILURE TERMINATED TEST

For N > 100,

$$L \doteq \left[1 + \sqrt{\frac{2}{N}}\; Z_{.5+\frac{\gamma}{2}}\right]^{-1} \qquad U \doteq \left[1 - \sqrt{\frac{2}{N}}\; Z_{.5+\frac{\gamma}{2}}\right]^{-1}$$

in which $Z_{.5+\frac{\gamma}{2}}$ is the $(.5 + \frac{\gamma}{2})$-th percentile of the standard normal distribution.

Next page is blank.

This page intentionally left blank.

APPENDIX D

BIBLIOGRAPHY

1. Aroef, M. 1957. Study of Learning Curves of Industrial Manual Operations. Unpublished Master's Thesis. Cornell University. Ithaca, NY.

2. Ascher, H. and Feingold, H. 1969. Bad-as-Old Analysis of System Failure Data. Proceedings Reliability and Maintainability Conference, pp. 49-62. New York, NY: ASME.

3. Barlow, R., Proschan. F., Scheuer, E. 1966. Maximum Likelihood Estimation and Conservative Confidence Interval Procedures in Reliability Growth and Debugging Problems. Report RM-4749-NASA. RAND Corporation. Santa Monica, CA.

4. Barlow, R. and Scheuer, E. 1966. Reliability Growth During A Development Testing Program. Technometrics. 8:53-60.

5. Bassin, W. M. 1969. Increasing Hazard Functions and Overhaul Policy. Proceedings Annual Symposium on Reliability, pp. 173-178. New York, NY: IEEE.

6. Box, G. E. P. and Jenkins, G. M. 1970. Time Series Analysis: Forecasting and Control. San Francisco, CA: Holden-Day.

7. Cox, D. R. and Lewis, P. A. W. 1966. The Statistical Analysis of Series of Events. New York, NY: John Wiley and Sons.

8. Crow, L. H. 1974. Reliability Analysis for Complex Repairable Systems, US Army Materiel Systems Analysis Activity, Technical Report 138, Aberdeen Proving Ground, MD. Also, Reliability and Biometry, ed. F. Proschan and R. J. Serfling, pp. 379-410. Philadelphia, PA: SIAM.

9. Crow, L. H. 1975. On Tracking Reliability Growth, Proceedings of the 1975 Annual Reliability and Maintainability Symposium, pp. 438-443. Washington, DC.

10. Crow, L. H. 1977. Confidence Interval Procedures for Reliability Growth Analysis. US Army Materiel Systems Analysis Activity, Technical Report 197. Aberdeen Proving Ground, MD: AD-A044788.

11. Duane, J. T. 1964. Learning Curve Approach to Reliability Monitoring. IEEE Transactions on Aerospace. 2: 563-566.

12. Englehardt, M. and Bain, L. J. 1978. Prediction Intervals for the Weibull Process. Technometrics. 20: 167-169.

BIBLIOGRAPHY (continued)

13. Finkelstein, J. M. 1976. Confidence Bounds on the Parameters of the Weibull Process. Technometrics. 18: 115-117.

14. Hollander, M. and Proschan, F. 1974. A Test for Superadditivity of the Mean Value Function of a Nonhomogeneous Poisson Process. Stochastic Processes and Their Applications. 2: 195-209.

15. Kempthorne, O. and Folks, L. 1971. Probability, Statistics, and Data Analysis. Ames, IA: The Iowa State University Press.

16. Lee, L. and Lee, S. K. 1978. Some Results on Inference for the Weibull Process. Technometrics. 20: 41-45.

17. Lewis, P. and Shedler, G. 1976. Statistical Analysis of Non-Stationary Series of Events. IBM Journal of Research and Development. 20: 465-482.

18. Lloyd, D. K. and Lipow, M. 1962. Reliability: Management, Methods and Mathematics. Englewood Cliffs, NJ: Prentiss-Hall.

19. Perkowski, N. and Hartvigsen, D. E. 1962. Derivations and Discussions of the Mathematical Properties of Various Candidate Growth Functions. Report CRA62-8. Aerojet-General Corporation, Azusa, CA. AD-349304.

20. Rome Air Development Center, 1975. Reliability Growth Study. RADC-TR-75-253, Air Force Systems Command, Griffiss Air Force Base, NY.

21. Rosner, N. 1961. System Analysis - Nonlinear Estimation Techniques. Proceedings National Symposium on Reliability and Quality Control, pp. 203-207. New York, NY: IRE.

22. Selby, J. D. and Miller, S. G. 1970. Reliability Planning and Management (RPM). Proceedings ASQC/SRE Seminar. Milwaukee, WS: ASQC.

23. Singpurwalla, N. 1978. Estimating Reliability Growth (or Deterioration) Using Time Series Analysis: 25: 1-14.

24. Virene, E. P. 1968. Reliability Growth and Its Upper Limit. Proceedings Annual Symposium on Reliability, pp. 265-270. New York, NY: IEEE.

25. Wolman, W. 1963. Problems in System Reliability Analysis. Statistical Theory of Reliability, ed. M. Zelen, pp. 149-160. Madison, WS: The University of Wisconsin Press.

(TO DETACH THIS FORM, CUT ALONG THIS LINE.)

STANDARDIZATION DOCUMENT IMPROVEMENT PROPOSAL

(See Instructions – Reverse Side)

1. DOCUMENT NUMBER	2. DOCUMENT TITLE

3a. NAME OF SUBMITTING ORGANIZATION

b. ADDRESS *(Street, City, State, ZIP Code)*

4. TYPE OF ORGANIZATION *(Mark one)*

- [ ] VENDOR
- [ ] USER
- [ ] MANUFACTURER
- [ ] OTHER *(Specify)*: ______

5. PROBLEM AREAS

a. Paragraph Number and Wording:

b. Recommended Wording:

c. Reason/Rationale for Recommendation:

6. REMARKS

7a. NAME OF SUBMITTER *(Last, First, MI)* – Optional	b. WORK TELEPHONE NUMBER *(Include Area Code)* – Optional
c. MAILING ADDRESS *(Street, City, State, ZIP Code)* – Optional	8. DATE OF SUBMISSION *(YYMMDD)*