AMC PAMPHLET AMCP 706-133

ENGINEERING DESIGN HANDBOOK

MAINTAINABILITY ENGINEERING THEORY AND PRACTICE

HEADQUARTERS, US ARMY MATERIEL COMMAND JANUARY 1976

DEPARTMENT OF THE ARMY
HEADQUARTERS UNITED STATES ARMY MATERIEL COMMAND
5001 Eisenhower Ave., Alexandria, VA 22333

AMC PAMPHLET
No. 706-133

10 January 1976

ENGINEERING DESIGN HANDBOOK
MAINTAINABILITY ENGINEERING THEORY AND PRACTICE

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd)

Paragraph *Page*

SECTION III
EQUIPMENT CONSIDERATIONS IN MAINTAINABILITY

CHAPTER 2
SYSTEM EFFECTIVENESS

SECTION 1
INTRODUCTION

SECTION II
SYSTEM EFFECTIVENESS METHODS

TABLE OF CONTENTS (Cont'd)

CHAPTER 3
MAINTAINABILITY ORGANIZATION AND MANAGEMENT

CHAPTER 4
MAINTAINABILITY ALLOCATION AND PREDICTION

SECTION I
INTRODUCTION

TABLE OF CONTENTS (Cont'd)

TABLE OF CONTENTS (Cont'd)

TABLE OF CONTENTS (Cont'd)

TABLE OF CONTENTS (Cont'd)

TABLE OF CONTENTS (Cont'd)

CHAPTER 7
ECONOMETRICS

SECTION I
COST FACTORS & ANALYSES

TABLE OF CONTENTS (Cont'd)

TABLE OF CONTENTS (Cont'd)

CHAPTER 8
STATISTICAL MAINTAINABILITY

SECTION I
INTRODUCTION

TABLE OF CONTENTS (Cont'd)

SECTION II
PRACTICAL APPLICATIONS

TABLE OF CONTENTS (Cont'd)

Paragraph *Page*

CHAPTER 9
MAINTAINABILITY DATA

SECTION I
LIFE CYCLE DATA

SECTION II
DATA ACQUISITION

CHAPTER 10
CONTRACTING RESPONSIBILITIES

SECTION I
INTRODUCTION

TABLE OF CONTENTS (Cont'd)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont'd)

LIST OF ILLUSTRATIONS (Cont'd)

LIST OF TABLES

LIST OF TABLES (Cont'd)

PREFACE

The Engineering Design Handbooks of the US Army Materiel Command have evolved over a number of years for the purpose of making readily available basic information, technical data, and practical guides for the development of military equipment.

This handbook was prepared by Igor Bazovsky and Associates, Inc., of Sherman Oaks, California, for the Engineering Handbook Office of Duke University, prime contractor to the US Army Materiel Command. It was completed through the coordinated efforts of Mr. Bazovsky, Sr., and the Engineering Handbook Office of the Research Triangle Institute, prime contractor to the US Army Materiel Command. Technical guidance was provided by an Ad Hoc Working Group under the chairmanship of Mr. H. J. Bukowski, Headquarters, US Army Materiel Command.

Igor Bazovsky, Sr., Igor Bazovsky, Jr., George W. Dauncey, Dr. Melvin B. Kline, Dr. Ernest M. Scheuer, and Dr. David Sternlight participated as co-authors in the writing of the handbook; each contributed his particular expertise and practical experiences.

The individual chapters were written to stand on their own, with a minimum of cross-referencing between the chapters, so that the reader can concentrate on the chapters which are of specific interest to him or to his activity. The interrelations of maintainability with design engineering and other disciplines (reliability, system effectiveness, logistic support, and life cycle costing) are highlighted through the whole text. Notation and symbols differ in some instances because of the variety of subjects covered, and in an attempt to be consistent with notation used in the referenced standard texts, documents, and papers pertaining to the various subjects. A standardization of notation is long overdue, as evidenced throughout the maintainability and reliability literature and also in statistics and probability theory.

The Engineering Design Handbooks fall into two basic categories–those approved for release and sale, and those classified for security reasons. The US Army Materiel Command policy is to release these Engineering Design Handbooks in accordance with current DOD Directive 7230.7, dated 18 September 1973. All unclassified Handbooks can be obtained from the National Technical Information Service (NTIS). Procedures for acquiring these Handbooks follow:

a. All Department of Army activities having need for the Handbooks must submit their request on an official requisition form (DA Form 17, dated Jan 70) directly to:

Commander
Letterkenny Army Depot
ATTN: AMXLE-ATD
Chambersburg, PA 17201

(Requests for classified documents must be submitted, with appropriate "Need to Know" justification, to Letterkenny Army Depot.) DA activities will not requisition Handbooks for further free distribution.

b. All other requestors—DOD, Navy, Air Force, Marine Corps, nonmilitary Government agencies, contractors, private industry, individuals, universities, and others—must purchase these Handbooks from:

National Technical Information Service
Department of Commerce
Springfield, VA 22151

Classified documents may be released on a "Need to Know" basis verified by an official Department of Army representative and processed from Defense Documentation Center (DDC), ATTN: DDC-TSR, Cameron Station, Alexandria, VA 22314.

Comments and suggestions on this Handbook are welcome and should be addressed to:

Commander
US Army Materiel Development
and Readiness Command
ATTN: DRCRD-TT
Alexandria, VA 22333

(DA Forms 2028, Recommended Changes to Publications, which are available through normal publications supply channels, may be used for comments/suggestions.)

CHAPTER 1

THE MAINTAINABILITY CONCEPT
SECTION I

INTRODUCTION

1-1 GENERAL

The rapid technological advances which have occurred in the past 25 years have made operating realities today of complex and costly systems. With the advent of jet aircraft, large helicopters, nuclear submarines, digital computers, automated combat vehicles and guns, satellites, manned spacecraft, worldwide command and communication systems, and other sophisticated systems, greater emphasis has been placed on the need for efficient and effective design in terms of system performance, support, cost, and life.

In the design of a system, many different requirements must be taken into consideration. Some of these are shown in Fig. 1-1. In addition to the more familiar requirements of performance, packaging, and environment, there are requirements for supportability, human factors, safety, reliability, maintainability, and producibility—all of which contribute to the measure of system worth and utilization. These requirements exist within the constraints of time and cost which also must be satisfied by the system, during its acquisition period as well as its use period.

In order to achieve the effective design desired, we must be able to handle qualitatively and quantitatively all of these parameters in our system models. Optimization of the system design will then consist of cost-effective trade-offs among pertinent parameters. The methodology for combining each of these parameters into the optimized system, as well as for handling each one separately within its own discipline, is called the System Engineering Process.

Maintainability is one of the system design parameters which must be given careful consideration, along with the other parameters of design, as part of system engineering. The ability of a system to be maintained—i.e., retained in or restored to effective usable condition—is often as important to system usefulness as is its ability to perform its intended function reliably. In spite of this, system designers are often more concerned with system performance features than with reliability and maintainability.

Reliability, as an engineering discipline, experienced rapid development shortly after World War II as an outgrowth of the requirements of missile and space technology. Within recent years, the realization that, in many cases, a more cost-effective system can be obtained by trading off some reliability for the ability to maintain a system easily has led to a considerable research and development effort to describe a new engineering discipline–maintainability. This discipline is new not in basic concept, but rather in the concentration given to its attributes, its relationship to other system parameters, the quantitative prediction and evaluation of maintainability during design, and its management.

Maintainability is a characteristic of system and, equipment design. It is concerned with such system attributes as accessibility,. test points, controls, displays, test equipment, tools, connectors, maintenance manuals, checklists, test and checkout, and safety. Maintainability engineering is the discipline which is concerned with the design and development of weapon systems and equipment to ensure effective and economical maintenance within prescribed readiness requirements.

Maintainability may be defined as a characteristic of design and installation which imparts to a system or end item a greater inherent ability to be maintained, so as to lower the required maintenance manhours, skill levels, tools, facilities, and logistic costs, and to achieve greater mission availability.

This engineering handbook is concerned with the theory and practice of maintainability as an engineering discipline which influences design.

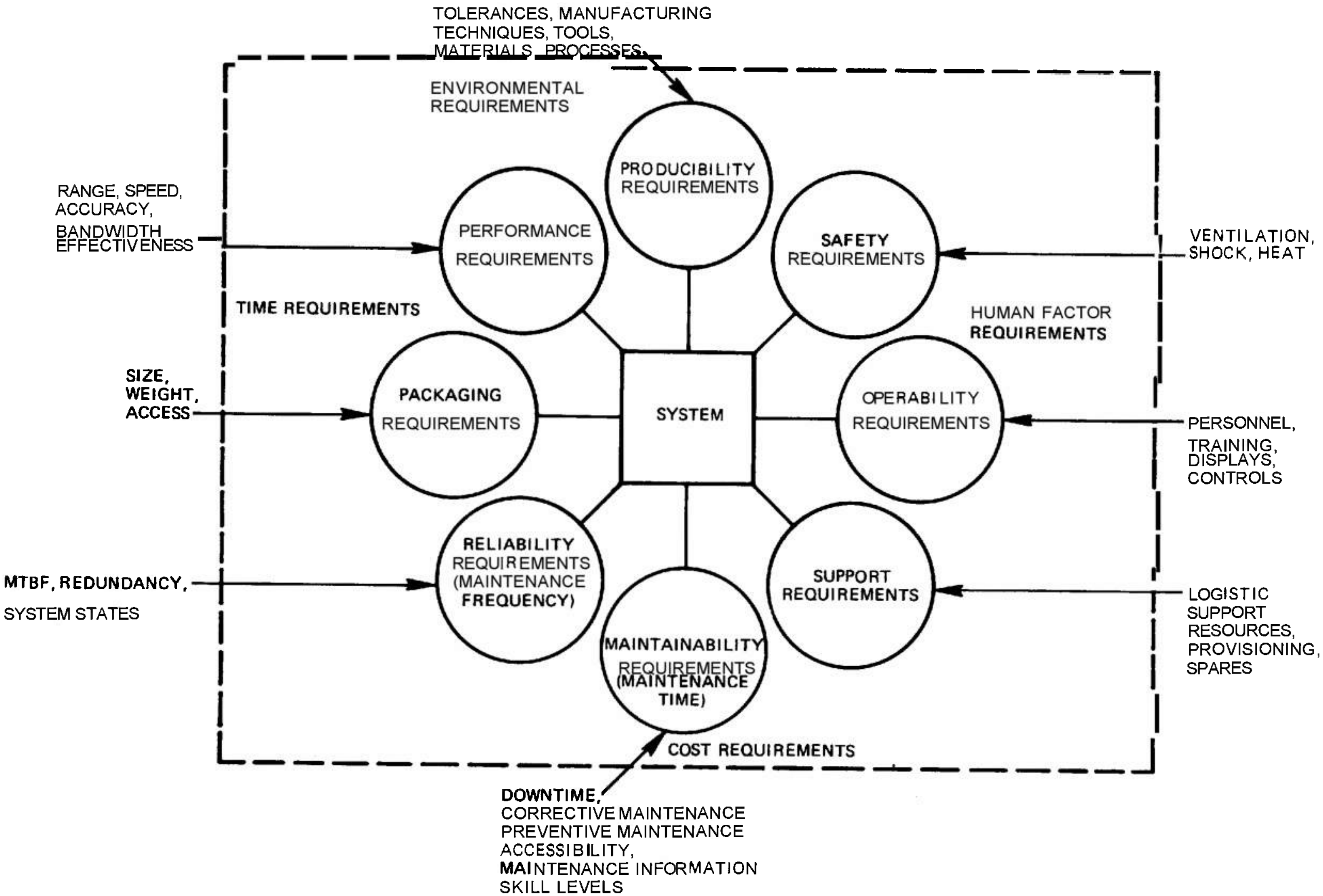


Figure 1-1. The Framework of System Design

Maintainability, as an engineering discipline, is not quite 20 years old. However, the ability to maintain equipment has been of concern for a much longer time. For example, in 1901 the Army Signal Corps contract for the development of the Wright Brothers' famous airplane contained a requirement that the airplane be "simple to operate and maintain". However, in its modem context, maintainability dates back to the early 1950's as an outgrowth of the intensive development of reliability after World War 11. At that time, concern with regard to maintainability was centered on the ability of systems to be serviced and repaired, without a formal approach.

By the late 1950's, concern with maintainability was focused on specific maintainability features in equipment design. Human factors engineers and psychologists, rather than equipment designers, took the lead in the development of maintainability. Numerous conferences, seminars, and informal group and panel meetings resulted in the development of a number of good design guides to an extent not ,continued in the 1960's. These design guides contain many worthwhile considerations still applicable to design for maintainability.

The growing concern for maintainability resulted in the development of military specifications as part of system requirements, the first of which, MIL-M-26512(USAF), appeared in June 1959. Subsequently, in the early 1960's general specifications for maintainability were issued by various Army and Navy Materiel Command organizations, in addition to the Air Force. As a result of the rapid proliferation of reliability and maintainability specifications—along with the development of the concept of system effectiveness as a combination of performance, reliability, and maintainability—the Department of Defense in the mid-1960's launched a standardization effort to reduce the number of specifications and to replace them with DoD-wide standards and a common language applicable to all the military services. One of the first of these was MIL-STD-778 on definition of maintainability terms. Subsequently, DoD issued in 1966 MIL-STD-470 on maintainability program requirements (Ref. 1), MIL-STD-471 on maintainability demonstration (Ref. 2), MIL-HDBK-472 on maintainability prediction (Ref. 3), and MIL-STD-721B on definition of effectiveness terms for reliability, maintainability, human factors, and safety (Ref. 4). The latter standard replaced MIL-STD-778, and the others replaced the individual service maintainability specifications. In addition, continued efforts in the maintainability engineering discipline resulted in refined techniques and additional maintainability design guides, such as AMCP 706-134 (Ref. 5).

Parallel with the development of the Military Standards and Specifications of the 1960's, the trend in maintainability turned away from guides for maintainability design and human factors to the quantification of maintainability, with time generally adopted as the common measure. Significant effort has been given to the development of techniques for prediction, demonstration, and evaluation of maintainability using statistical measures, such as mean time to repair *(MTTR)* and median repair time, as the quantification parameters. Other measures frequently used are maintenance manhours per unit of use (e.g., flying hours, miles, rounds), minimum time to failure, maximum time to repair, minimum time between overhaul. In addition, considerable attention has been given to maintainability program management throughout system development and design, as part of system engineering, including the interface relationship of maintainability with reliability, integrated logistic support, and cost-effectiveness.

The rapid development of maintainability as a discipline in the 1960's, along with other system engineering disciplines, has resulted in some instances in specification of maintainability program requirements that have become too costly when applied. Recently, it has been recognized that maintainability, as well as other system disciplines, must be selectively tailored to the needs of each particular program or specific categories of equipment.

Experience has shown that specifications often have expressed optimistic desires rather than operational needs. Maintainability demonstrations and predictions have not agreed with subsequent field use of systems, with actual repair times proving to be several times longer than predictions and demonstrations had indicated (Refs. 6-8).

It is already apparent that the 1970's will see the continued development and accelerated maturation of maintainability as one of the system engineering disciplines. Current specifications and standards will undoubtedly be modified as experience dictates and as new technology requires. For example, the advent of microelectronics and new methods of constructing and packaging electronic systems requires that data formerly applicable for vacuum tube, discrete component, and conventional wiring and construction contained in current maintainability prediction and demonstration specifications be revised. New maintenance concepts and maintainability design techniques must also be devised to keep up with such change. The long neglected and more difficult need to develop maintainability design and quantification techniques for nonelectronic systems and equipment, particularly mechanical and hydraulic, has been recognized and will become one of the primary areas to receive considerable attention.

1-2 THE IMPORTANCE OF MAINTAINABILITY

If a system is to be cost-effective over its designated operational life, its ability to meet performance requirements is only one of many considerations. Also of concern is system ability to perform when needed and for the duration of its assigned mission. This latter concern deals with system operational readiness and mission reliability; for this, a proper balance between system reliability and maintainability is required. Not only is such a balance necessary, but in order to be achieved, reliability and maintainability considerations must begin *early* in the conceptual and definition phases of system acquisition, as part of the overall system engineering effort.

The need for maintainability is emphasized by the alarmingly high operating and support costs which exist due to failures and the necessary subsequent maintenance. Lack of reliability and poor maintainability carry the major responsibility for this situation.

One study, made in the 1950's, showed that one-third of all Air Force operating cost was for maintenance, and one-third of all Air Force personnel was engaged in maintenance, even though a large portion of the maintenance was done by contract (Ref. 9). Army studies indicate that the orginal purchase price of electronic equipment represented only 25 to 40 percent of the total life-cycle cost, with the remainder resulting from operation and maintenance (Ref. 10, Chapter 1; Ref. 11).

No exact or up-to-date data on the cost of maintenance of military equipment exist at present. Service and General Accounting Office studies indicate that, when averaging maintenance costs over all systems deployed, these costs exceed three to ten times the procurement costs during the life cycle of equipment.

The system resources associated with maintainability, and their attendant costs, include test and support equipment, repair parts, maintenance personnel and their training, training equipment, maintenance facilities, maintenance instructions and data, and other logistic costs. The extent of the resources depends upon the specific reliability and maintainability features designed into the equipment and specified in contract work statements. Because they represent such a significant part of total system resources and costs, the need for a logical, cost-effective approach to maintainability is emphasized.

There is a multiplier or leverage effect involved in system design, particularly with respect to maintainability and logistic support. In effect this means, as illustrated in Fig. 1-2, that maintenance and support considerations have a strong leverage effect on system cost and effectiveness when taken into account early in the system life cycle and have much less effect later on. One can consider the system life cycle to be a long lever with its fulcrum placed at the life-cycle phase where maintainability and logistic support are considered. Thus, in the conceptual development phases, a relatively moderate investment in reliability, maintainability, and support design requirements can produce very substantial savings in the operation phase. On the other hand, waiting until late validation or production phases to consider maintainability and support features may tip the balance in the other direction and result in excessive maintenance and support costs. No other factor affects the life-cycle logistic cost with the preponderance of inclusion of proper implementations of its maintainability and reliability.

In personnel costs alone, the savings realized from using just one less maintenance technician has been estimated to be approximately $15,000 per year in pay and allowances, administrative support, and training costs. Couple with this the savings in repair parts, maintenance information, and support equipment costs, and a significant impact on life-cycle cost can be achieved.

It is readily seen, therefore, that an original investment in maintainability made during system acquisition may produce a manifold saving in operating costs and a substantial improvement in system effectiveness. The Weapons Systems Effectiveness Industry Advisory Committee (WSEIAC) study on system effectiveness (Ref. 12) states:

"The high cost and complexity of modern military systems require the most efficient management possible to avoid wasting significant resources on inadequate equipment.

"Efficient systems management depends on the successful evaluation and integration of numerous different but interrelated system characteristics such as reliability, maintainability, performance and cost. If such evaluation and integration is to be accomplished in a scientific rather than intuitive manner, a method must be formulated to assess quantitatively the effects of each system characteristic on overall system effectiveness."

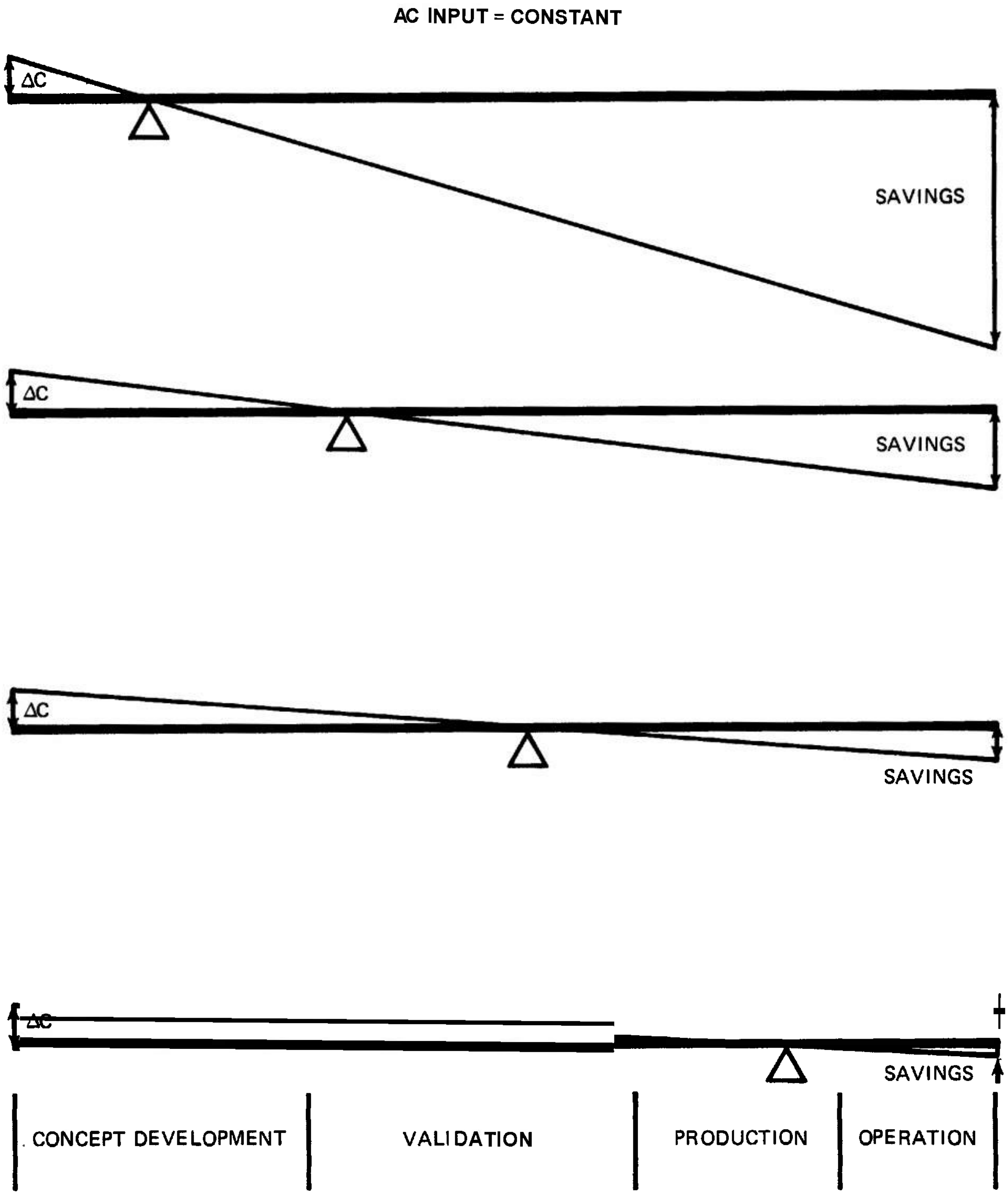


Figure 1-2. The Lever Effect

However, although extremely important, cost is not the only consideration with regard to the need for maintainability engineering. The ability of a system to operate when needed and to do so for the duration of the specified mission is often as important, and sometimes even more important, than cost savings. This suggests then that time is an important parameter in maintainability. Time is used as a common measure in system effectiveness. A system to which maintainability engineering has been properly applied can be expected to have:

1. Lower downtime, and therefore a higher operational readiness (availability)

2. The capability of being restored quickly to operating status when downtime is due to random failures (corrective maintenance)

3. The capability of being retained in an operationally ready state by inhibiting those types of failures which result from age or wearout (preventive maintenance).

In some Army systems, the failure of one critical item of equipment due to lack of maintenance or provision of adequate maintainability features may cause an important mission or battle to be lost, with a resultant loss of life and equipment. This could be vital to our national security.

The need, therefore, is to provide a maintainability program which,will assure that maintainability features reflecting operational maintenance requirements are included in system design throughout system acquisition from the early conceptual phase through at least system development, test, and evaluation.

1-3 PURPOSE OF MAINTAINABILITY

Maintainability engineering is concerned with the operational readiness of a system or equipment. Operational readiness (sometimes called materiel readiness in the Army) is the term used to indicate the ability of a system to be utilized upon demand. It consists of a number of factors—primary ones being the inherent reliability of the system/equipment, its ability to be maintained, and its mission or operational demand requirement in its operational environment. AR 702-3 states "The primary objectives of the reliability and maintainability program are to assure that during the life cycle, items of materiel provided to Army forces will be ready for use when needed, will be able to successfully perform their assigned functions, and will fulfill all required maintenance characteristics" (Ref. 13).

It is possible to achieve operational readiness by making the system so reliable that failures are rare. However, such a system, if feasible within the state-of-the-art, could require components that might be so costly that the system would not be economical or cost-effective. On the other hand, it is possible to design a system in such a manner that any failure could occur frequently but the failure could be corrected in a short time. Such a system might also be very expensive in terms of its design characteristics (number of test points, accessibility, skill levels required, displays, troubleshooting logic, repair levels), or in terms of maintenance resources required (skilled technicians, maintenance float, repair cycle float, repair parts, tools and test equipment, manuals), so that it also would not be cost-effective. In addition, when considering system or equipment utilization in terms of mission times, a system that might fail frequently, even though it could be repaired quickly, might be intolerable to a field commander and might well result in loss of confidence by the user or in mission failure, with consequent disastrous results. Operational readiness, therefore, requires a suitable balance between reliability and maintainability. Maintainability, then, is used to obtain maximum operational readiness in such a way that an end item can be maintained in the least time consistent with other system requirements, and with a minimum expenditure of support resources.

In order to achieve such a proper balance, maintainability considerations, like reliability, must start with the original materiel requirement in the concept development phase of the system life cycle. Maintenance and maintainability considerations must be part of the original system/equipment planning effort. Integrated logistic support concepts must be developed during these early phases and must be approved before subsequent phases can be entered by the developer. Further, there must be a proper balance of logistic support resource needs versus cost, schedule, and performance in order to achieve maximum system effectiveness and operational readiness.

1-4 MAINTENANCE ENGINEERING AND MAINTAINABILITY

Maintenance and maintainability have different meanings. ***Maintenance*** is concerned with those actions taken by a system user to retain an existing system/equipment in, or restore it to, an operable condition. ***Maintainability*** is concerned with those actions taken by a system/equipment ***designer***, during development, to incorporate those design features which will enhance ease of maintenance. Its function is to ensure that—when produced, installed, and operated—the fielded system/equipment can be maintained at minimum life-cycle support cost and with minimum downtime.

The life-cycle support (user) aspects are the responsibility of ***maintenance engineering***, and they influence the design aspects which are the responsibility of ***maintainability engineering***. This difference in perspective and responsibility is recognized in AR **750-1** (Ref. **14**) and TM **38-703** (Ref. **15**).

1-4.1 THE USER-PRODUCER DIALOGUE

Every system has a user and a producer. The system user is the one whose needs for the system must be met by the system producer. Thus, a dialogue is necessary between system users and producers, as, for example, between someone who wants a house built and the architect and builder who design and produce the house to satisfy the user's needs.

The system user is concerned with formulating and developing the needs and concepts for the system and for its operation and support. He provides the requirements to which the producer designs. The producer is concerned with translating the user's formulated needs into the design, production, and installation of the system which meets these needs and which can be operated and supported in a cost-effective manner. The system life cycle is the logical framework for carrying out the user/producer dialogue. (See par. **3-2**.)

There is a user-producer relationship within the Army. The ultimate users in the Army are the various combat Field Army Commanders and other operating forces. The Army Materiel Command (AMC) is responsible for system and equipment research and development, acquisition, and support; and the Training and Doctrine Command is responsible for training. These are the internal producers in the Army. AMC represents the Army as user and developer to the industry which is the external producer.

The user-producer dialogue allows maintenance engineering and maintainability engineering to be put into proper perspective. Maintenance engineering represents the user's needs; maintainability engineering represents the producer's response to these needs. The responsibility for the conduct of both maintenance and maintainability engineering rests with the AMC commodity commands.

1-4.2 MAINTENANCE ENGINEERING

Maintenance engineering is defined in AMCR **750-42** as "that activity of equipment maintenance which develops and maintains concepts, criteria, and technical requirements from concept through obsolescence of materiel to assure timely, adequate, and economic maintenance support of AMC materiel" (Ref. **16**). It is defined in AMCP **706-134** as "the application of techniques, engineering skills, and effort organized to ensure that the design and development of weapons, systems, and equipment provide adequately for effective and economical maintenance" (Ref. **5**). Of particular note in these definitions is the important role assigned to maintenance engineering in the concept, validation, and design phases of system and equipment development.

This is further emphasized in AMCR **750-42** as follows:

"During the concept formulation, validation and production phases, the maintenance engineering activity provides necessary maintenance support concepts, plans, and maintenance experience data to be used in developing technical requirements for new weapons and equipments. Maintenance engineers participate in the design reviews and evaluation of test results to reduce the need for maintenance support. Thus, effective maintenance engineering participation significantly influences technical requirements in design which, in general, dictate initial and future support investments and operating costs associated with new military hardware."

The maintenance engineer is concerned with how the fielded system will be operated and maintained. Since he represents the user needs, he is concerned with system mission/operational and support profiles, the environment in which the system will be operated and maintained, the levels of maintenance, maintenance and other support resources, and maintenance actions. It is his responsibility to see that user needs with regard to maintenance are reflected in system development and design requirements.

Within the defined operational use concepts, the maintenance engineer must help develop the overall system integrated logistic support (ILS) concept and the maintenance concepts and constraints which will guide the system designer with respect to maintainability design. Maintainability design requirements for maintainability engineers are provided through the process of maintenance engineering analysis, the development of maintenance concepts, the analysis of maintenance tasks and requirements, and the determination of maintenance resource requirements. The development of a maintenance concept must precede maintainability design, not result from it. Maintenance and maintainability engineering must influence system design to be effective. The output of maintenance engineering analysis should be a "Plan for Maintenance" which is consistent with the maintenance concept and which serves as the basis for maintenance planning for the system during its use period as well as a basis for maintainability design.

1-4.3 MAINTAINABILITY ENGINEERING

Since maintainability is defined as "the inherent ability of a design to be maintained" (Refs. 1 and 2), maintainability engineering is concerned with incorporating required maintainability features in system/equipment design. Maintainability design requirements are an output of the maintenance engineering analysis which reflects user needs. It is the task of the maintainability engineer to see that maintainability features required to meet these needs are incorporated in the system/equipment design contracts. Maintainability engineering must be integrated with the other elements of system engineering so as to provide the necessary effectiveness, considering all costs over the entire life cycle of the system equipment (Ref. 13).

Maintainability engineering is concerned with specific features of system/equipment design and with other physical characteristics of the system pertinent to its rapid maintenance with the least logistic resources. Examples of such design features are accessibility, human factors considerations, test, checkout, calibration, and replace/repair/discard features resulting from the selected maintenance concept and from maintenance engineering analysis.

Maintainability engineering is also concerned with specific features for fault detection—Built-in Test Equipment (BITE), fault isolation, correction, and verification—at each maintenance level. It is concerned with contributions of various parts of the system to the allocation, prediction, and demonstration of quantitative measures of maintainability. It is concerned with incorporating preventive and corrective maintenance requirements in such a way that the system will meet stated operational readiness and system effectivenes. goals within specified mission and logistic time profiles. Maintainability engineering is concerned with designing for specified manpower skills and with the development of maintenance instructions, aids, and training for maintenance personnel.

AMCP 706-134, ***Maintainability Guide for Design*** (Ref. *5)*, is an engineering design handbook which contains many of the design requirements, features, and concepts that maintainability engineers will apply to Army systems and equipment.

1-4.4 EXAMPLES OF MAINTENANCE POLICY INTERRELATIONSHIPS

The following examples illustrate the interrelationships between maintenance engineering and maintainability engineering. In each example, a maintenance concept is stated, followed by the resulting maintainability design implications.

Example 1. Maintenance Concept. Organizational maintenance shall be performed by equipment operators, organization repairmen, and direct support technicians as needed. Organizational maintenance activities shall be limited to inspection, preventive maintenance, servicing, and minor adjustment. Only minor repairs and replacements shall be made by direct support technicians. No special tools or limited general-purpose test equipment shall be required for this maintenance level.

Maintainability Design Implication. Organizational repairmen shall not require high skill levels. BITE features shall be incorporated into equipment so that the operator need only turn a function test switch and note an indicator reading, preferably by a go/no-go or lo-go-hi type of indication. Repairs shall be made primarily by replacing faulty items without the need for special tools and test equipment, utilizing built-in signal sources and indicators, and with minimum dependence on repair parts.

Example 2. Maintenance Concept. *MTTR* at the organizational level shall not exceed 10 min.

Maintainability Design Implication. No time for detailed troubleshooting and repair is allowed at organizational level. Fault localization and isolation and verification features must be incorporated directly in the equipment, using a test function switch. Repairs shall be made by replacement, using plug-in units and standard tools. Quick-access fasteners shall be used to gain access to units.

Example 3. Maintenance Concept. Organizational level maintenance shall make maximum feasible use of plug-in modules which can be discarded at failure. No module repair shall be performed at the organizational level. A repair/discard criterion of $100 might be used.

Maintainability Design Implication. Module design shall be such that, insofar as possible, those modules requiring replacement at the organizational level should cost less than $100. Where modules costing more than $100 must be removed, they should be replaced and the failed unit sent back to general support or depot for repair.

Example **4.** Maintenance Concept. At the direct support level, replacement of one module shall not require removal or adjustment of other modules or important units, except for those adjustments normally provided by BITE for operator use in order to align unit performance to peak efficiency.

Maintainability Design Implication. Replaceable modules must be designed so that they contain all necessary performance functions, components, and adjustments within the module, except for interface adjustments.

1-5 PRIMARY CONSIDERATIONS

Reliability and maintainability are elements of system engineering and are viewed as interrelated characteristics (Ref. 13). They are different but complementary engineering disciplines.

Reliability engineering provides the methodologies for increasing the ability of a system to operate without failure or serious degradation for prolonged periods of time in its operational environment (Ref. 17). It is thus concerned with extending system "up" time. Maintainability engineering, on the other hand, provides the methodologies for reducing the "down" time of systems when maintenance becomes necessary because of failures or in order to reduce the need for preventive maintenance actions when system performance is drifting out of, the specified performance limits.

Reliability and maintainability of a system are related to each other in terms of operational readiness, mission success, and system availability which measure system uptime with respect to the total time the system is required to operate.

Although reliability and maintainability are closely allied disciplines, one significant difference between them is the extent to which they are dependent upon the use of manpower, and, therefore, human factors. Inherent (equipment) reliability is primarily dependent upon the physical characteristics of the equipment and its components—such as stress-strain relationships, failure modes and effects, and environmental factors. Mission (operational) reliability is dependent, in addition to the stated physical characteristics, on the number and skill level of the equipment operators and, therefore, of the specific human engineering features which have been incorporated in the equipment to assist the operator in performing his task reliably.

Inherent maintainability cannot be divorced from human factors considerations, except in the improbable event of completely self-healing systems.

By self-healing is meant the ability of a system to correct its own defect or failure, such as removing a short or restoring an imbalance. The automatic switching in a standby redundant item to replace a failed item does not constitute self-healing. From the outset, therefore, the maintainability engineer must be concerned with human factors, maintenance technician skill levels and capabilities, and safety. Thus, maintainability engineering requires a multi-disciplined approach utilizing personnel with backgrounds in such areas as equipment design, statistical techniques, safety, and human factors. Maintainability is a joint effort of these types of personnel with the reliability and system effectiveness engineers, maintenance and logistic engineers, and system engineers (see Fig. 1-1).

The actual preventive and corrective maintenance tasks which can be performed on a system are a direct consequence of the maintainability characteristics which have been designed into the system. To design for these features is the responsibility of the maintainability engineers and equipment designers. The maintainability design requirements are derived from maintenance and logistic support concepts and operational requirements. Maintainability design considerations are discussed in Chapter **5.**

Maintainability as an element of system effectiveness is predicated on the fact that system maintainability requirements can be specified quantitatively and, therefore, can be predicted, measured, demonstrated, and evaluated. Maintainability quantification, as part of system effectiveness, is discussed in Chapters 2, **4,** 6, and 8.

Maintainability is part of integrated logistic support, system engineering and program management, and, therefore, must be considered in terms of the system life cycle with respect to program and system planning, system trade-offs, and life-cycle costs. These aspects of maintainability are discussed in Chapters 3, **7,** 9, and 10.

SECTION II

QUANTIFICATION OF MAINTAINABILITY

1-6 MAINTAINABILITY MEASURES

In MIL-STD-721B (Ref. 4) maintainability is defined as "a characteristic of design and installation expressed as a probability that an item will be retained in or restored to specified conditions within a given period of time, when maintenance action is performed in accordance with prescribed procedures and resources." Expressed somewhat differently, maintainability is the Probability that an item in need of maintenance will be retained in/or restored to a specified operational condition within a given period of time. The variable in this probabilistic definition of maintainability is the maintenance time.

Obviously, maintenance time will differ from case to case according to the nature of the failure or malfunction which requires maintenance. Therefore, maintenance time is not a constant but is in some way statistically distributed. This is in a sense similar to the distribution of time-to-failure in reliability. The difference is that in maintainability the variable is always time, while in reliability the variable may be the time to failure, or miles to failure, or rounds fired to failure, or cycles to failure, or number of successful trials to failure, etc. This difference, as will be seen later in the text, shows up in evaluating the availability of systems, where uptime may be measured in miles traveled or rounds fired without failure and downtime is measured in hours or minutes; it is thus not always easy to combine the two into meaningful and realistic measures of availability. Another difference between reliability and maintainability is the fact that while reliability is the probability that an event, i.e., failure, will *not* occur in a specific time, maintainability is the probability that the event, i.e., successful completion of maintenance, *will* occur in a specific time.

1-6.1 THE EXPONENTIAL CASE

The simplest and mathematically easiest way to handle a case is with exponential distribution. It applies in maintainability to corrective maintenance when the duration of repair times is exponentially distributed, according to the equation

$$M(t) = 1 - \exp(-t/MTTR) \qquad (1\text{-}1)$$

where

$M(t)$ = probability that repair will be successfully completed in time t when it starts at $t = 0$

t = variable repair time

MTTR = mean time to repair

exp = base of the natural logarithm ($e = 2.71828....$).

Looking at this equation, we see that it has only a single parameter, namely the *MTTR*. Once the MTTR is given, $M(t)$ can be calculated for any specific value of t Thus for each value of t, the probability $M(t)$ of completing repair in t is fully defined by the MTTR. Fig. 1-3 illustrates two such maintainability functions $M(t)$—one for an equipment with an *MTTR* of 0.5 hr and the other for an equipment with an *MTTR* of 1 hr.

1-6.2 THE CONCEPTS OF MEDIAN REPAIR TIME AND M_{MAX}

From Fig. 1-3 we can make some interesting observations and draw definite conclusions. Looking at the maintainability function $M(t) = 1 - \exp(-2t)$ of the equipment which has an *MTTR* of 0.5 hr, we see that the probability of accomplishing repair $M(t)$ in a time $t = 0.5$ hr (30 min) is approximately 0.63 or 63 percent, while the probability of accomplishing repair in $t = 0.25$ hr (15 min) is only about 0.40 or 40 percent. On the other hand, the probability of accomplishing repairs in 1 hr becomes approximately 0.865 or 86.5 percent, and we find that for a repair time of 2.3 X MTTR, or for t = 1.15 hr (about 69 min) there is a probability of $M(t) = 0.9$ or 90% of accomplishing repair.

To generalize, an exponentially repaired equipment has a probability of about 63% of accomplishing repair in a time t which equals its *MTTR* (i.e., t = MTTR), a probability of about 40% for $t = 0.5$ *MTTR*, a probability of about 22% for t = 0.25 *MTTR*, a probability of about 90% for $t = 2.3$ X

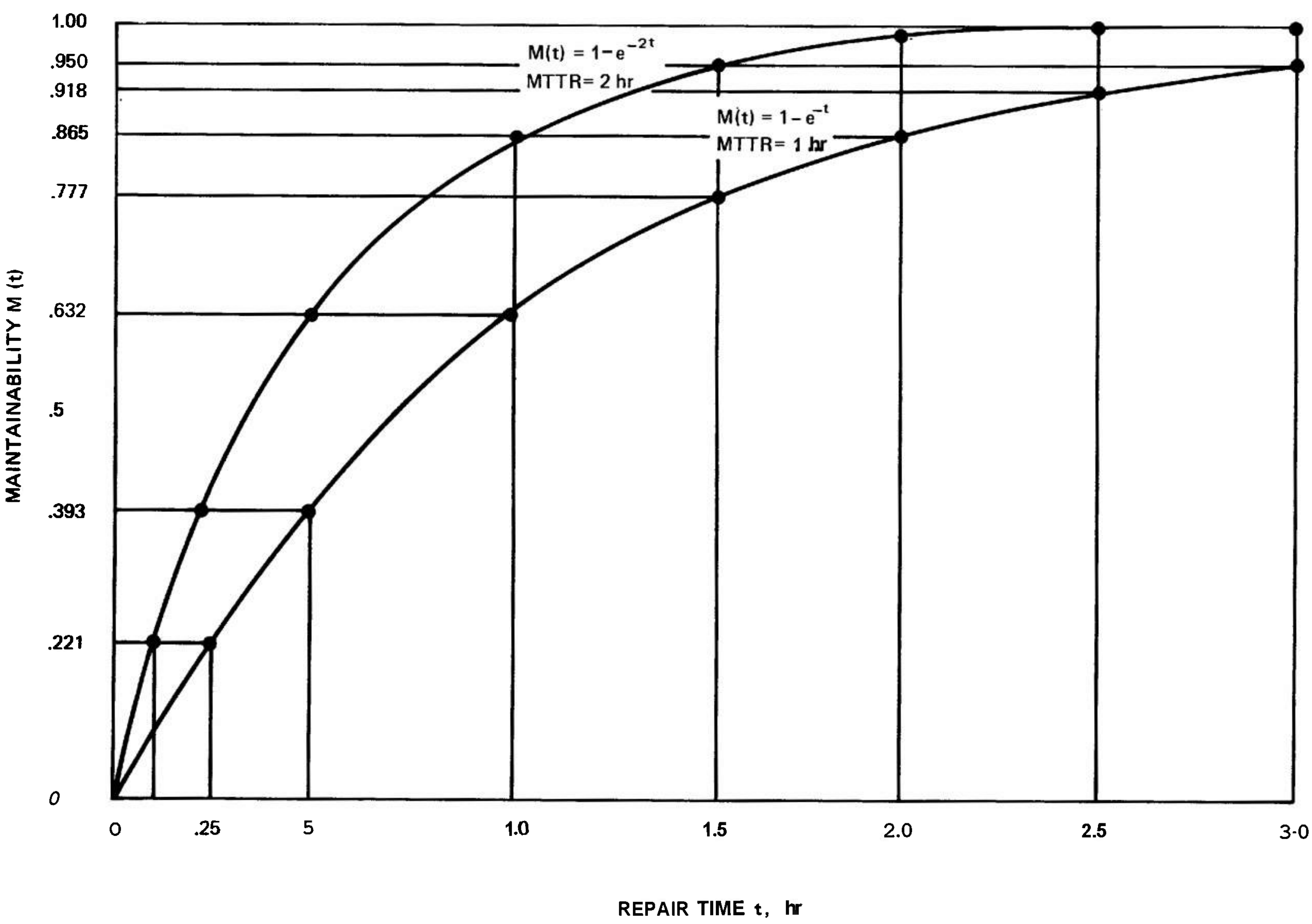


Figure 1-3. Two Exponential Maintainability Functions

MTTR and a probability of about 95% for $t = 3 \times$ MTTR. Finally, there is a 50% probability of accomplishing repair in approximately $t = 0.7$ MTTR which is called the median time to repair.

Of specific interest in maintainability specifications are the last two numbers, i.e., 50% and 90% probabilities. It is often desirable to specify a maximum repair or maintenance time M_{MAX} which should possibly not be exceeded or, exceeded only with a small probability. Such constraints on maximum maintenance time are usually associated with the 90th or 95th percentile, i.e., the probability of accomplishing maintenance in a specified time $t = M_{MAX}$ should be 0.9 or 0.95, according to what the specification demands. In the case of an exponential distribution of repair times, $M(t) = 0.9$ for approximately $t = M_{MAX} = 2.3 \times$ MTTR and $M(t) = 0.95$ for $t = M_{MAX} = 3 \times$ MTTR. The explanation of such a requirement is that 90% or 95% of all repair actions shall require less than $t = 2.3 \times$ MTTR or $t = 3 \times MTTR$, respectively, according to which percentage is associated with the M_{MAX} requirement. For example, if the MTTR is 1 hr (refer to Fig. **1-3**), 90% of all repair actions should take less than **2.3** hr and 95% should take less than **3** hr.

In the exponential case it makes no difference whether the MTTR or the M_{MAX} are specified along with the associated probability or percentile. If M_{MAX} (maximum maintenance time) is specified with probability $M(t) = 1 - \alpha$, i.e.,

$$M(t) = 1 - \alpha = 1 - \exp(-M_{MAX}/MTTR) \quad (1\text{-}2)$$

which may also be written as

$$\ln \alpha = -M_{MAX}/MTTR \quad (1\text{-}3)$$

we obtain from such requirement the MTTR as a design goal by taking the natural logarithm of the above equation, i.e., $\ln \alpha = -M_{MAX}/MATTR$, and solving for *MTTR* we get

$$MTTR = -M_{MAX}/\ln \alpha \quad (1\text{-}4)$$

For example, if the assumption of an exponential distribution of maintenance time is valid and a customer specifies that with probability $M(t) = 1 - \alpha = 0.9$, the maintenance time must not exceed 1 hr, i.e., $M_{MAX} = 1$ hr, the *MTTR* to design for is obtained from Eq. 1-4 by finding $\alpha = 0.1$, $\ln 0.1 = -2.30259$, and $MTTR = -M_{MAX}/\ln \alpha = -1/(-2.30259) = 0.434$ hr or about 26 min.

So far we have determined that one of the maintainability measures is the length of time it takes to perform maintenance actions and that this time may be distributed according to a maintainability function $M(t)$, such as the exponential function in **Eq. 1-**1. When the exponential distribution is applicable, a specific, unique, and sufficient measure of maintainability is the *MTTR*. When this is specified, all percentile points are also automatically defined, such as M_{MAX}, and associated with this, the median time to repair. The mathematical formulas by which these measures are interrelated have been shown, and the relationships are illustrated in Fig. **1-3.** It must be emphasized that all the equations presented so far apply only to the case of the exponential distribution of repair or maintenance time. However, the maintainability measures developed—i.e., the concept of maintainability function $M(t)$, mean time to repair MTTR, maximum repair time M_{MAX}, and median time to repair—apply also to other statistical maintenance time distributions, such as the lognormal, normal, gamma, and others; only the mathematical formulas by which these measures are interrelated become different.

1-6.3 THE REPAIR RATE μ

In the maintainability literature one often finds the concept of maintenance rate or repair rate μ, especially when dealing with the exponential distribution. For the exponential case, the repair rate is given as the reciprocal of the MTTR, i.e.,

$$\mu = 1/MTTR \quad (1\text{-}5)$$

Since the *MTTR* is a fixed number, the repair rate μ is a constant for the exponential distribution. For all other distributions, the repair rate is nonconstant. It usually increases as a function of the progressing maintenance time t. When this is the case, the probability of completing or finishing a repair in a short period dt when repair started t time units ago, i.e., $\mu(t)dt$, increases the longer repair has been in progress. On the other hand, in the exponential case μdt is always constant, regardless of how long a repair action has been in progress.

1-6.4 THE MEAN TIME TO REPAIR (*MTTR*)

To return to the concept of MTTR, this is an important parameter, easy to quantify, and easy to measure

(Ref. 18). Unfortunately, by itself, except for the exponential distribution, ***MTTR*** does not tell us enough about the tails of the distribution, such as the frequency and duration of the very long maintenance actions. Still, ***MTTR*** is an important design requirement especially for complex. pieces of equipment and systems, and it can be measured when the hardware is tested.

By its nature, ***MTTR*** depends on the frequencies at which various replaceable or repairable components in the equipment fail (i.e., on the failure rates or replacement rates), and on the times it takes to repair the equipment as the different kinds of failure occur. There is a predicted ***MTTR*** for which we need to know the predicted failure rates and estimated repair times down to the lowest repair level at a given repair level, and there is the measured ***MTTR*** observed on actual hardware. Ideally, the two ***MTTRs*** will be close to each other. But if the predicted failure rates are not correct, the measured ***MTTR*** may deviate significantly from the predicted value, even though the individual repair times initially were well estimated. When designing an equipment for maintainability, prediction techniques such as are in **MIL-HDBK-472** are used. An ***MTTR*** estimate of an exponentially failing equipment is obtained from the formula

$$MTTR = \sum_{i=1}^{N} \lambda_i t_i / \lambda \qquad (1\text{-}6)$$

where

N = total number of replaceable or repairable components

λ_i = failure rate of the ith component

t_i = equipment repair time when the ith component fails

λ = failure rate of the whole equipment, usually taken as the sum of the failure rates of all components in the equipment

Eq. 1-6 is a very practical design tool for maintainability. When the predicted failure rates are available, the maintainability engineer evaluates the expected repair times t. They are estimated by maintenance time analysis methods based on previous field data or expert engineering judgment which consider fault verification, fault localization, fault isolation, disassembly, replacement, reassembly, adjustment, servicing, and checkout. Each of these actions takes a certain time to perform, but these times can well be estimated from the design, testability, and packaging concept for the equipment. Trade-off techniques are used to change design and packaging characteristics, as well as test capabilities, to achieve the desired repair times t_i for the various types of failures and thus to comply with the *MTTR* requirement. As to the measured ***MTTR***, this is determined from hardware test, simulated maintainability demonstrations, or field data by computing the total observed repair downtime over an extended period of time (the sum of all individual downtimes), and dividing this by the number of repair actions N_r which occurred in the period of observation, i.e.,

$$MTTR = \sum_{i=1}^{N_r} t_i / N_r \qquad (1\text{-}7)$$

Observing Eqs. 1-6 and 1-7, one can see that the ***MTTR*** computations are very simple, requiring only simple summations, multiplications, and divisions/ easily done by the help of an inexpensive **desk** calculator or slide rule. **As** to the preceding Eqs. 1-1 through **1-5,** these are also easily handled by exponential tables (Ref. 18) and slide rules. Some more complex mathematics, however, will be involved when discussing the specifics of the more complicated distributions.

1-7 SPECIFIC MEASURES IN MAINTAINABILITY

In par. 1-6, certain measures in maintainability have been identified and some equations for these measures developed, with an emphasis on the simple exponential distribution of repair time. However, in many instances maintenance is performed not only when a system or equipment develops a failure or malfunction but also preventively to forestall the possible occurrence of such an undersirable event. Maintenance actions can thus be divided into two major categories.

1. Corrective maintenance, performed when the equipment fails to perform to required performance specifications.

2. Preventive maintenance, performed to avoid the equipment getting into a condition requiring corrective maintenance.

Whether maintenance is corrective **or** preventive, it usually causes a definite amount of downtime for the equipment so it cannot be used while the maintenance actions are performed. But there is a distinct difference between downtime due to Corrective maintenance actions and downtime due to preventive maintenance actions. While the need for corrective maintenance is usually due to equipment breakdowns and malfunctions which occur at random times when the equipment is operating and therefore interfere with equipment operational schedules, preventive maintenance can be scheduled so that it is performed at predetermined times when the equipment is not required to operate or when substitute equipment can be used, so that either no undesired reduction of output or use is encountered, or effects of such are minimized.

Still, it is obvious that the need for preventive or scheduled maintenance imposes additional burdens on an undisturbed equipment operation in terms of the costs associated with it, the possible need for substitute equipment, or the loss of the function for some periods of time. In the context of maintainability, it **is** therefore necessary to consider preventive maintenance as well as corrective maintenance when evaluating the usefulness, maintenance costs, and availability of an equipment.

Though the penalties due to scheduled preventive maintenance may be smaller than those resulting from corrective maintenance, they are still real losses and subtract from the value of the equipment to the user. Even though such loss is usually not of the same magnitude as the loss suffered due to failures during operation, to assess it and to include it in the evaluation of overall worth of the equipment to the user in terms of maintainability, availability, and pay-off capability becomes a necessity.

1-7.1 MEASURES OF MAINTENANCE DOWNTIME

Although maintainability has been defined as a probability (Ref. **4),** there are a number of useful time measures by which quantitative maintainability requirements can be specified and trade-offs performed with reliability, availability, and other system engineering disciplines. Quantitative requirements for maintainability may be expressed in different ways according to the type of equipment/system, their usage, and the maintenance concept. There may be a quantitative availability requirement specified which, in conjunction with the reliability requirement, yields a quantitative maintainability requirement in terms of the mean time to repair ***(MTTR)*** or mean downtime. In other instances, the maintenance manhours per system operating hour (*MMH/OH*) may be specified and maintainability design goals then derived from such specification. Other useful measures applicable to specific systems are time between overhauls, turnaround time, and a number of maintenance downtime measures currently used by maintainability engineers, such as mean time to repair ***(MTTR)***, mean active corrective maintenance time ($\overline{M}_c$), mean active preventive maintenance time ($\overline{M}_p$), mean active corrective and preventive maintenance time (@, median equipment repair time (*ER1*), maximum equipment repair time (ERT_{MAX}), geometric mean time to repair ($MTTR_G$), and maximum maintenance time (M_{MAX}). Ref. 19, Chapter **4,** and Ref. **3**, pages 2-3 through 2-6, define these various terms somewhat differently. In the paragraphs that follow definitions are used which give more consistent results.

1. ***Mean Time to Repair (MTTR)*** is defined as the mean of the distribution of equipment or system repair time. In its simplest form, the ***MTTR*** is given by the equation

$$MTTR = \sum_{i=1}^{N} \lambda_i t_i \Big/ \sum_{i=1}^{N} \lambda_i \qquad (1\text{-}8)$$

where

A, = failure rate of the ith repairable or replaceable component in the equipment/system

t_i = time required to repair the system when the ith component fails

The ***MTTR*** is sometimes given in hours and at other times in minutes. It is important to use the same time units for the A's and for the *t*'s. Failure rates are usually (but not always) given in units of "failures per hour". Then the repair times should also be given in hours. This becomes obvious in availability calculations.

As an example of ***MTTR*** computation, assume a system consisting of three replaceable subassemblies (components) which have the following *MTBF*'s and replacement times:

Subassembly 1: $MTBF_1$ = 1000 hr, t_1 = 1 hr
Subassembly 2: $MTBF_2$ = 500 hr, t_2 = 0.5 hr
Subassembly 3: $MTBF_3$ = 500 hr, t_3 = 1 hr

To compute the *MTTR* of the system, we first convert the *MTBF*s into failure rates, i.e., A, = 1/1000 = 0.001; A, = 1/500 = 0.002; and A, = 1/500 failures per hour. Then, using Eq. 1-8 we calculate

$$MTTR = \frac{\lambda_1 t_1 + \lambda_2 t_2 + \lambda_3 t_3}{\lambda_1 + \lambda_2 + \lambda_3}$$

$$= \frac{(0.001)(1) + (0.002)(0.5) + (0.002)(1)}{0.001 + 0.002 + 0.002}$$

$$= \frac{0.004}{0.005} = 0.8 \text{ hr} \qquad (1\text{-}8a)$$

When the time to failure is exponentially distributed according to the reliability equation

$$R(\tau) = \exp(-\lambda\tau) \qquad (1\text{-}9)$$

where

λ = failure rate
τ = operating time

the reciprocal of λ is the mean time between failures, i.e., $MTBF = 1/\lambda$ (Ref. 20, Chapter 3). The *MTBF* is often used as a measure of reliability, just as the *MTTR* is often used as a measure of maintainability.

2. ***Mean Active Corrective Maintenance Time*** ($\bar{M}_c$), is defined the same way as the *MTTR*, except that emphasis is on active maintenance time, which means that no idle time must be included when measuring the duration of maintenance tasks. However, this applies to the *MTTR* measure, also.

Denoting the active maintenance time of a system by M_{ci} when the ith component with failure rate λ_i fails, the mean active maintenance time of the system is given by

$$\bar{M}_c = \sum \lambda_i M_{ci} \Big/ \sum \lambda_i \qquad (1\text{-}10)$$

3. ***Mean Active Preventive Maintenance Time*** ($\bar{M}_p$), is defined as the arithmetic mean of the active preventive maintenance times of an equipment or system and is given by

$$\bar{M}_p = \sum f_i M_{pi} \Big/ \sum f_i \qquad (1\text{-}11)$$

where

f_i = frequency at which the ith preventive maintenance task is performed
M_{pi} = system active maintenance time when the ith preventive maintenance task is performed.

If the frequencies f_i are given in maintenance tasks per hour, the downtimes M_{pi} should also be given in hours.

4. ***Mean Active Corrective and Preventive Maintenance Time*** ($\bar{M}$) is defined as the mean of the distribution of time of all maintenance actions, both corrective and preventive, of an equipment or system. It is given by the equation

$$\bar{M} = \frac{\sum \lambda_i M_{ci} + \sum f_i M_{pi}}{\sum \lambda_i + \sum f_i} \qquad (1\text{-}12)$$

where the terms λ_i, f_i, M_{ci}, and M_{pi} are as defined in the preceding paragraphs. In this equation the same units ***must*** be used for the λ_i's and f_i's, and the same time units for M_{ci}'s and M_{pi}'s.

5. ***Equipment Repair Time*** (*ERT*) is defined as the ***median*** of the distribution of repair times of an equipment/system. It was discussed in par. 1-6.2 in connection with the exponential distribution. Fig. 1-4 is presented here to indicate more generalization. As seen in Fig. 1-4, the *ERT* corresponds to that repair time within which **50%** of all repair actions can be accomplished.

The numerical relationships between *ERT* and *MTTR* are different for different distributions. For the normal distribution, because of its symmetry, the median and the mean coincide

$$ERT = MTTR \qquad (1\text{-}13)$$

For the exponential distribution, we have approximately

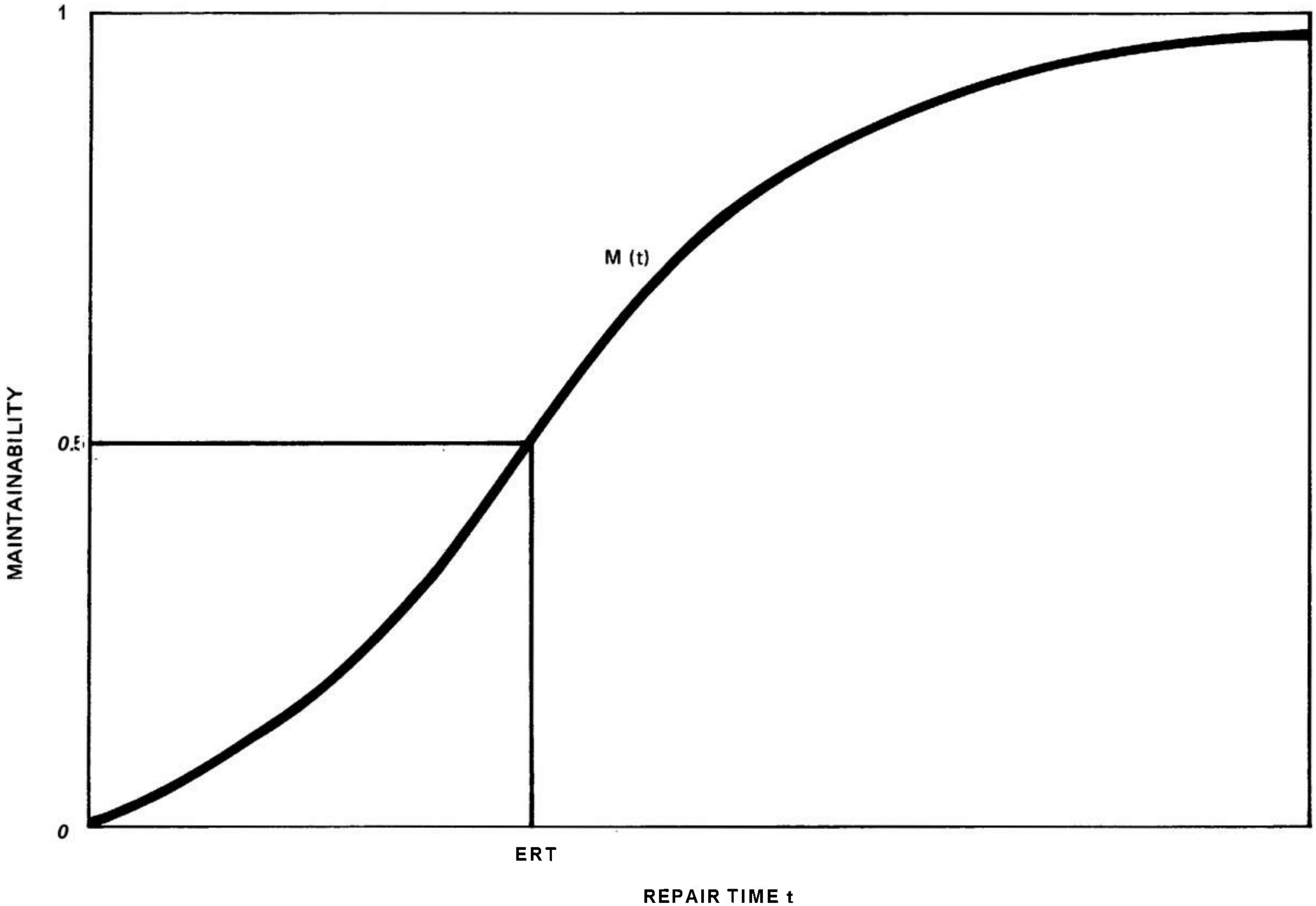


Figure 1-4. Equipment Repair Time (ERT)

$$ERT = 0.7\ MTTR \tag{1-14}$$

For the lognormal distribution the relationship holds

$$MTTR = ERT \exp(\sigma^2/2) \tag{1-15}$$

which yields

$$ERT = MTTR/\exp(\sigma^2/2) \tag{1-16}$$

where σ^2 is the variance around the mean of the ***natural logarithm*** of repair times.

6. ***Geometric Mean Time to Repair*** ($MTTR_G$) is used in the lognormal distribution, where it happens to be identical with ***ERT***. It is given by Eq. 1-17 which is identical with Eq. 1-16, i.e.,

$$MTTR_G = MTTR/\exp(\sigma^2/2) \tag{1-17}$$

It can be directly obtained from the mean m of the natural logarithms of the repair times t_i which is given by

$$m = \sum \lambda_i \ln t_i \Big/ \sum \lambda_i \tag{1-18}$$

and the $MTTR_G$ is then given by

$$MTTR_G = e^m \tag{1-19}$$

7. ***Maximum Maintenance Time*** (M_{MAX}) is defined as the 95th percentile of the maintainability function $M(t)$, as shown in Fig. 1-5. M_{MAX} is that maintenance time within which 95% of all maintenance action can be accomplished, i.e., not more than 5% of the maintenance may exceed M_{MAX}. For the normal distribution M_{MAX} occurs at approximately

$$M_{MAX} = MTTR + 1.65\,\sigma \tag{1-20}$$

where σ is the standard deviation of the normally distributed maintenance time. For the exponential distribution M_{MAX} is approximately

$$M_{MAX} = 3\ MTTR \tag{1-21}$$

and for the lognormal distribution the relationship holds

$$\ln M_{MAX} = m + 1.65\,\sigma \tag{1-22}$$

where m is given by Eq. 1-18, and σ is the standard deviation of the natural logarithm of the repair times.

In some instances the concept of Maximum Equipment Repair Time (ERT_{MAX}) also has been introduced into maintainability. It is defined as the "maximum allowable value" of ***ERT***, and is quoted to be $ERT_{MAX} = 0.45\ M_{MAX}$ for the lognormal distribution (Ref. 19, page 87).

1-7.2 TIME FACTORS IN MAINTENANCE

In the preceding paragraph specific maintenance downtime measures were defined, mostly pertaining to the active maintenance time as it occurs in repairs and preventive maintenance tasks. The active maintenance time can be corrective or preventive.

The active corrective maintenance time consists of the sum of certain elemental times it takes to perform the various activities which jointly result in the completed repair. These are failure verification time, fault location time, fault isolation time, access time, fault correction time, reassembly time, adjustment-calibration time, checkout time, and cleanup-servicing time. Fault correction time may involve repair in place; or remove, repair and replace; or remove and replace with a like item. The active preventive maintenance time involves inspection time and servicing time or turnaround time in the case of scheduled maintenance actions (Ref. 21).

However, when considering the total downtime, almost invariably delays occur, such as supply delay time, administrative time, and work breaks, which can be summarized under the concept of delay time. Fig. 1-6 presents a useful block diagram of time relation-

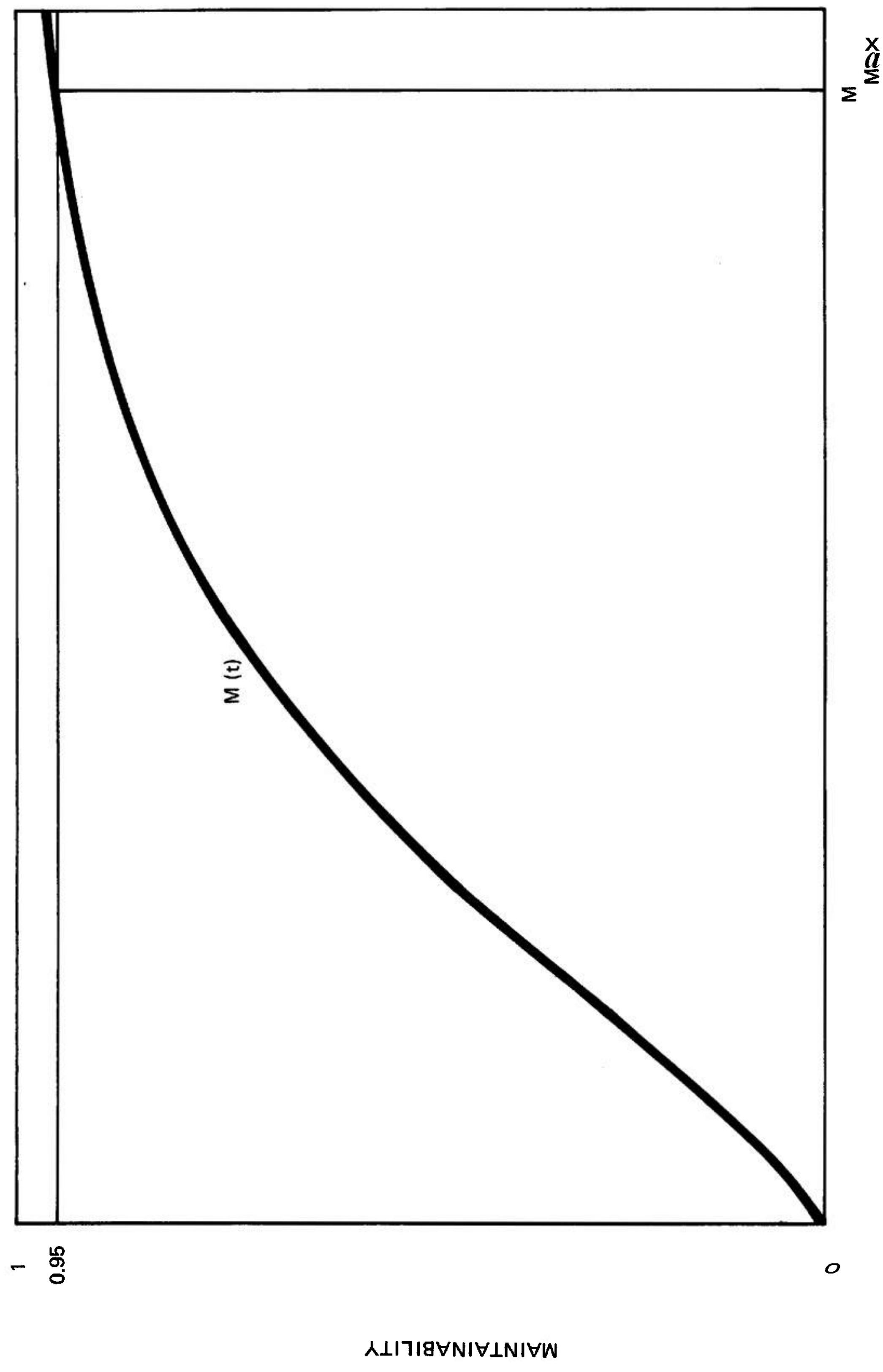


Figure 1-5. Maximum Maintenance Time M_{MAX}

ships which considers system uptime as well as downtime and thus establishes a good basis for the discussion of availability and related factors (Ref. 4).

1-7.3 AVAILABILITY FACTORS

The concept of availability is best explained in terms of a continuously operating system which is either operating and thus "up", or is in maintenance and thus "down". Availability is then defined as the probability that at an arbitrary point in time the system is operable, i.e., is "up".

Of specific interest to maintainability engineers, who look at the long-term or steady-state operation of systems, are the concepts of Inherent Availability A_i, Achieved Availability A_a, and Operational Availability A_0 (Ref. 19, pages 6, 7, 82-84).

Inherent Availability A_i considers the mean time between failures ($MTBF$) and the ***MTTR*** of a system and is by definition given by the formula

$$A_i = MTBF/(MTBF + MTTR) \tag{1-23}$$

It excludes idle time, logistic time, waiting time, and preventive maintenance time and is therefore a useful parameter for equipment/system design. Fig. 1-7 is a nomograph for fast determination of A_i, $MTBF$, or $MTTR$ if two of these parameters are known.

Achieved Availability A_a includes preventive maintenance and is given by the formula

$$A_a = MTBM/(MTBM + \bar{M}) \tag{1-24}$$

where $\bar{M}$ is the mean active corrective and preventive maintenance time as given by Eq. 1-12, and $MTBM$ is the mean interval between corrective and preventive maintenance actions equal to the reciprocal of the frequency at which these actions occur, which is the sum of the frequency or rate λ at which corrective actions occur, and the frequency or rate f at which preventive maintenance actions occur.

Therefore

$$MTBM = 1/(\lambda + f) \tag{1-25}$$

Operational Availability includes in addition to A_a logistic time, waiting time, and administrative time, so that the total mean downtime MDT becomes

$$MDT = \bar{M} + \text{Mean Waiting Time} + \text{Mean Logistic Time} + \text{Mean Administrative Time}$$

and adds to the uptime the ready time RT, i.e.,

$$A_0 = (MTBM + RT)/(MTBM + RT + MDT) \tag{1-26}$$

It is important to realize that RT is the system average ready time in a complete operational cycle, the cycle being MTBM + MDT + RT.

1-7.4 MAINTENANCE MANHOURS

The maintenance manhours expended in equipment maintenance are not identical with active maintenance downtime. This would be so only in a case where a single maintenance man would perform the maintenance actions. Quite frequently two or more men, or a whole maintenance crew, work on a system. In addition, maintenance manhours are expended at various maintenance levels—such as at the organizational level, direct support level, general support level, and depot level.

For instance, a system may have only a short maintenance downtime to replace a failed "black box". But the failed black box may require many maintenance manhours at some rear maintenance level to be repaired and made available again as a spare part.

Since maintenance manhours are expensive, it became necessary to specify certain constraint for these support labor costs in terms of an index called maintenance manhours per system operating hour (MMH/OH). This is a necessity especially for larger systems where several maintenance levels are usually involved. The MMH/OH index, when specified, must be and can be considered in maintainability design and becomes a design parameter not only for the maintainability of the system, but also for maintainability of the "black boxes" at rear levels and for appropriate planning of the maintenance concept.

1-8 STATISTICAL ASPECTS AND STATISTICAL DISTRIBUTIONS

Statistics play an important role in the estimation of the various measures in maintainability. Maintenance downtime is always in some way statistically distributed, and when maintenance time data are collected they must first be ordered in some way. The kind of statistical distribution they most likely belong to must be determined, and then the parameters of the distribu-

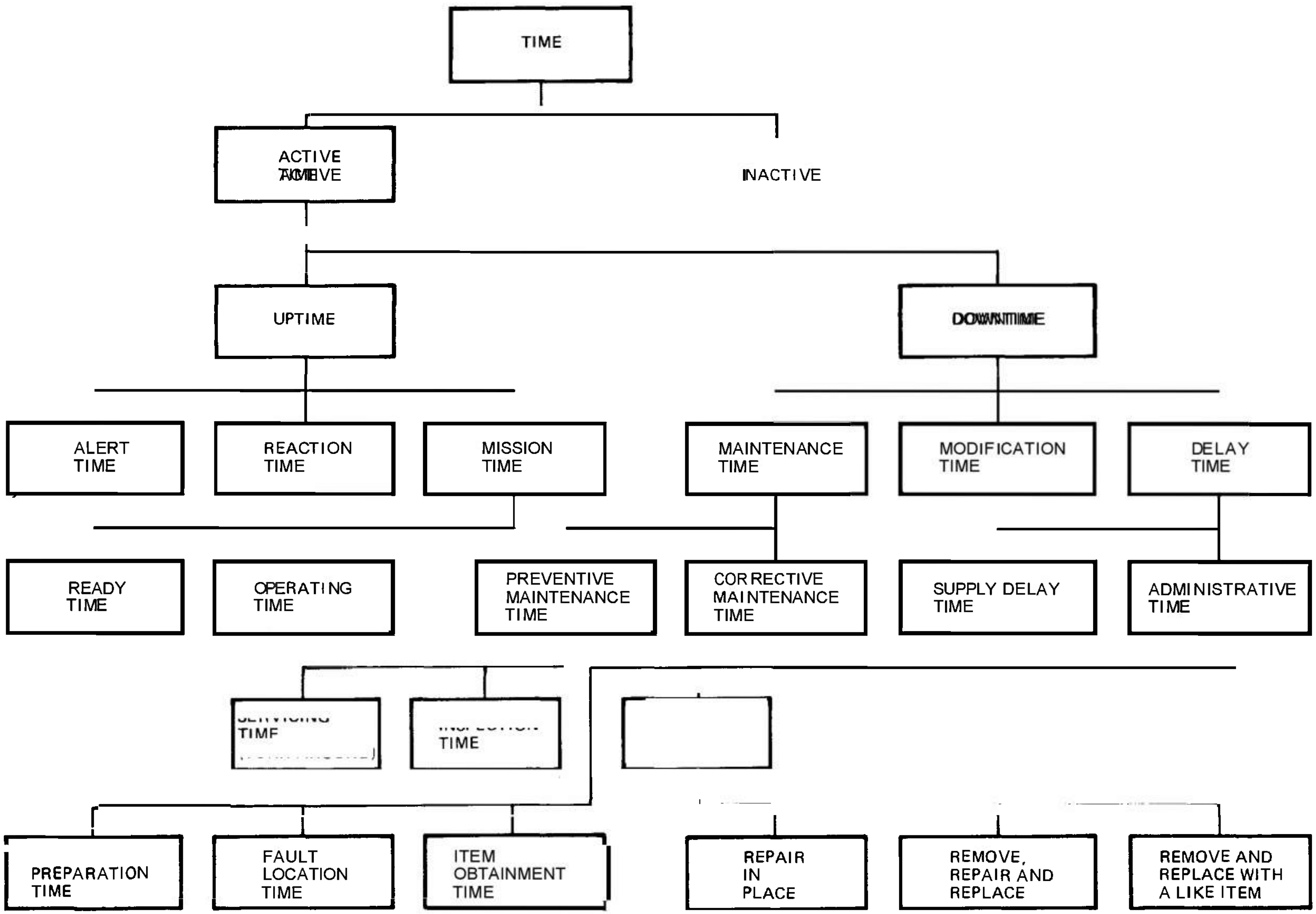


Figure 1-6. Time Relationship

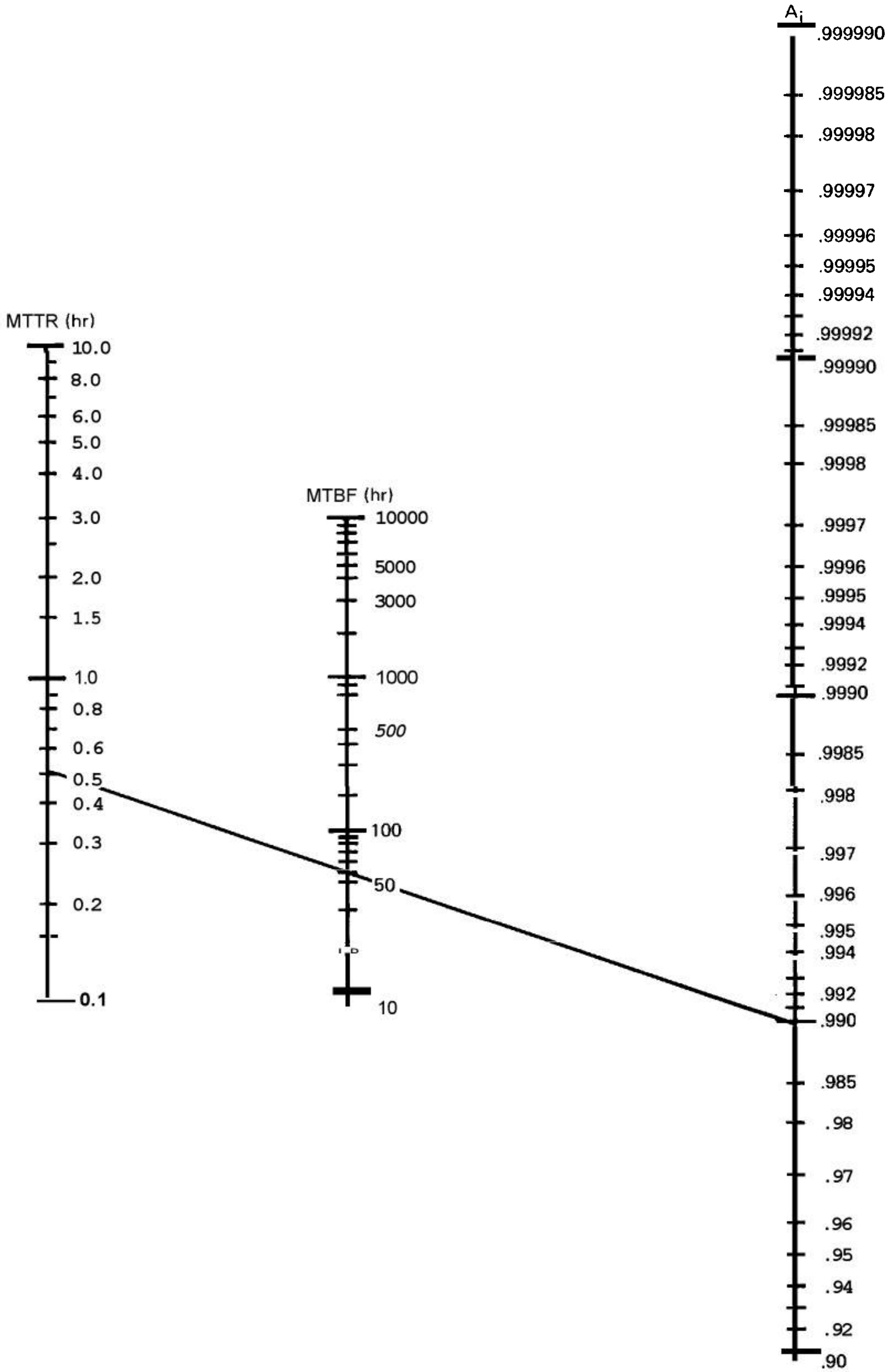


Figure 1-7. Availability Nomograph

.ion are predicted using estimation techniques. Estimates thus obtained also serve to verify whether the predicted parameters, i.e., the maintenance downtime measures, were predicted closely enough during the design phase.

Probability is an important aspect of maintainability, in view of the fact that maintenance times are statistically distributed. There are several statistical distributions which can be well applied in maintainability and are used commonly in solving maintainability problems. Some of these distributions are now discussed.

In this paragraph the major statistical distributions are introduced in a form usually given in texts on statistics and probability (Refs. **22, 23, 24**), and for simplicity of presentation, will use the notation *t* for the variable maintenance time and M for the mean of the distribution of maintenance time. The exponential distribution has already been introduced in par. **1-6.** All distributions introduced in this paragraph, including the exponential distribution, are discussed in great detail with numerical examples in Chapter 8.

1. *The Normal Distribution*

The probability density function (*pdf*) of the normal distribution (Ref. **22**, Chapter **10**, and Ref. **23**, Chapter 3) has the equation

$$f_t = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{t-M}{\sigma}\right)^2\right] \qquad (1\text{-}27)$$

where σ is the standard deviation of the variable maintenance time t around the mean *M.* Fig. **1-8** shows a typical normal density function, which is always symmetrical about the mean M

The area under this curve, taken from the left to any point t is the cumulative distribution $M(t)$ which is the maintainability function (see Fig. **1-9**).

Therefore, the maintainability function $M(t)$ is given **by**

$$M(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x=-\infty}^{t} \exp\left[-\frac{1}{2}\left(\frac{t-M}{\sigma}\right)^2\right] dt \qquad (1\text{-}28)$$

The mean *M*, which corresponds to the *MTTR,* is estimated from observed and measured maintenance times t_i

$$M = \frac{t_1 + t_2 + t_3 \cdots + t_n}{n} = \sum_{i=1}^{n} t_i/n \qquad (1\text{-}29)$$

and the standard deviation σ is estimated by the equation

$$\sigma = \sqrt{\sum_{i=1}^{n}(t_i - M)^2/(n-1)} \qquad (1\text{-}30)$$

We call the normal distribution a two-parameter distribution, since when the mean M and the standard deviation σ are known, the shape of the curves $f(t)$ and $M(t)$ is fully defined.

2. The Lognormal Distribution

The lognormal distribution is a skewed two-parameter distribution, widely used in maintainability. In its most general form the probability density function $f(t)$ of the lognormal distribution is given by:

$$f(t) = \frac{1}{(t-c)\sigma\sqrt{2\pi}} \cdot \exp\left\{-\frac{1}{2}\left[\frac{\ln(t-c) - m}{\sigma}\right]^2\right\} \qquad (1\text{-}31)$$

where

- t = maintenance time
- m = mean of the natural logarithms of the maintenance times
- σ = standard deviation with which the natural logarithm of the maintenance times are spread around the mean m
- c = a constant, the shortest time below which no maintenance action can be performed.

The effect of c is to shift the origin of $f(t)$ from $t = 0$ to $t = c$. In subsequent discussions, we assume c to be zero so that $f(t)$ starts at $t = 0$. Fig. **1-10** shows a typical density $f(t)$ and maintainability $M(t)$ function of the lognormal distribution.

Like all skewed distributions, the lognormal density function has three characteristic points (Ref. **24**), which are shown in Fig. **1-10:** the mode M_M at which $f(t)$ has its maximum; the median M_G which bisects the area under $f(t)$ into two equal parts of 50 percent; and the mean M which is the expected or average value of maintenance time t and is defined as the first moment of the distribution.

$$M(t) = \int_0^{\infty} t f(t) dt$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_0^{\infty} \exp\left[-\frac{1}{2}\left(\frac{\ln t - m}{\sigma}\right)^2\right] dt \qquad (1\text{-}32)$$

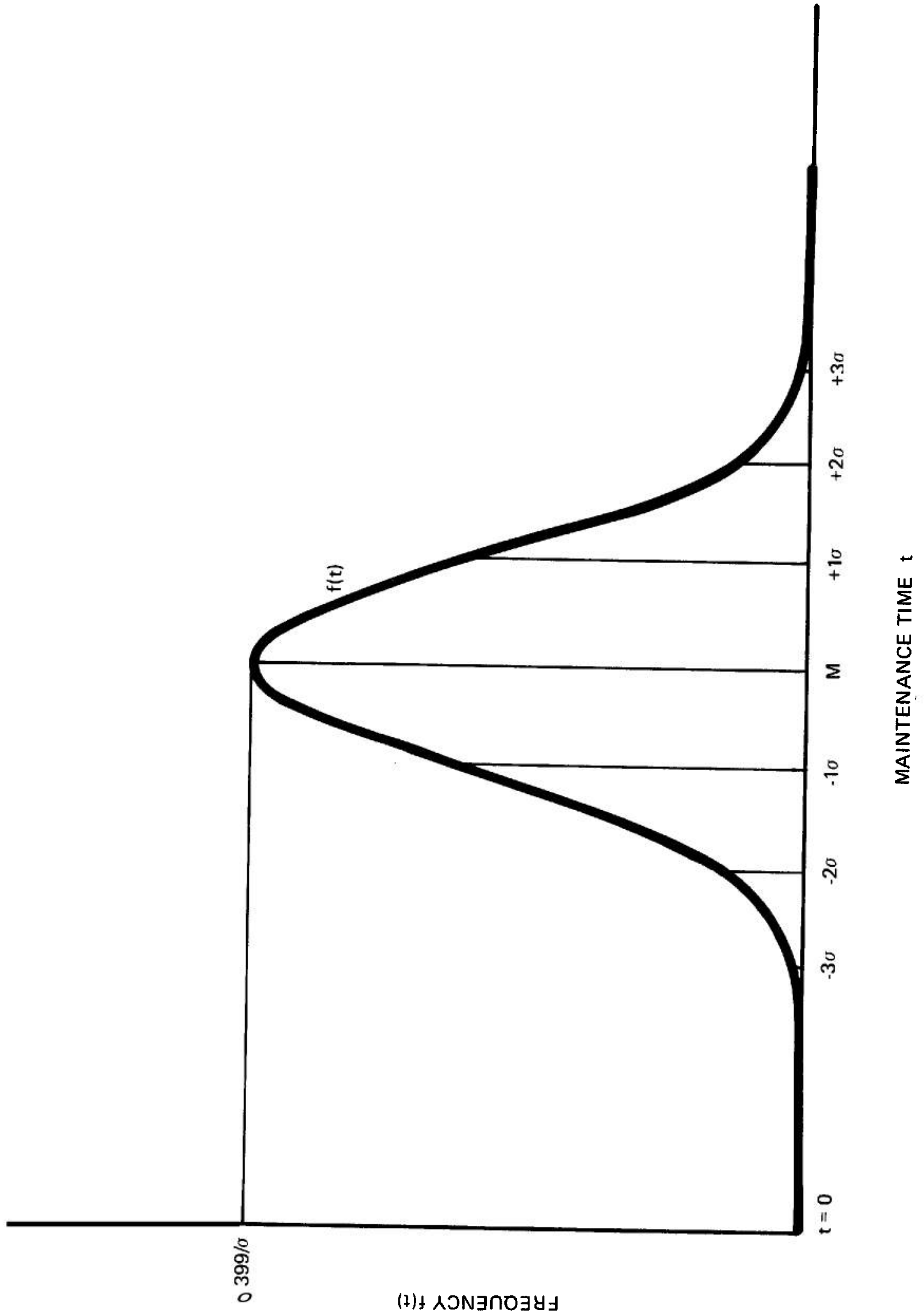


Figure 1-8. Normal Density Function

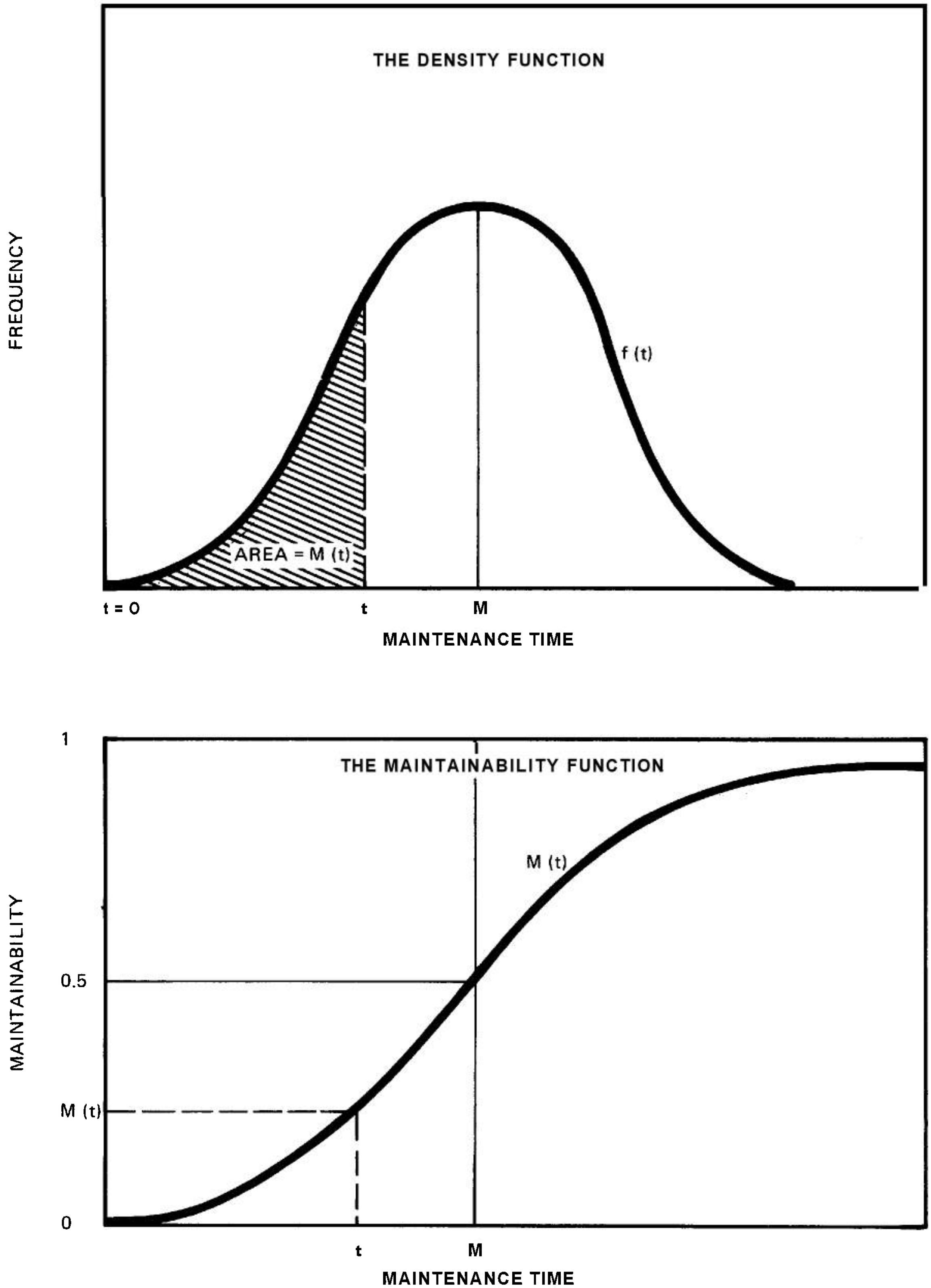


Figure 1-9. Normal Cumulative Distribution

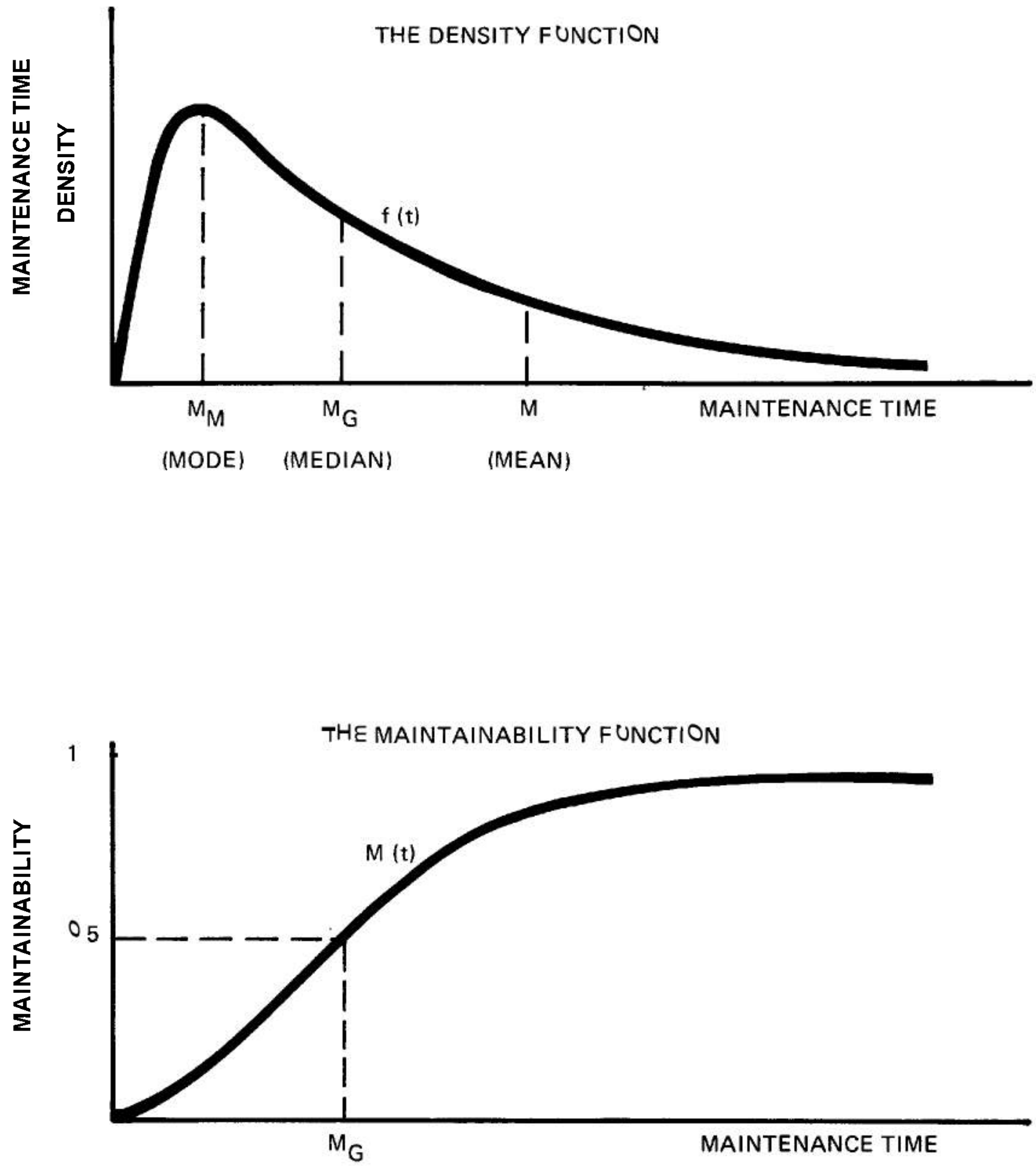


Figure 1-10. A Typical Lognormal Distribution

To find explicit formulas for the mode, median, and mean, we make use of a convenient feature of the lognormal distribution, namely, that the natural logarithm of the variable maintenance time t is normally distributed. This fact, on which the derivation of the lognormal distribution is based, makes it also easier to obtain numerical values of the maintainability function $M(t)$, for given arguments t, by looking up in normal tables the cumulative probability values (areas) corresponding to $x = \ln t$. Fig. 1-11 shows the transform property of the lognormal distribution graphically.

The transformed density function *Ax)* $= f(\ln t)$, which is of the normal form, has an $x = \ln t$scale on the abscissa. For $t = 0$, $x = -\infty$. The mean of the $\ln t$'s is m. It bisects the area of the normal density curve. Since every point ton the abscissa of the lognormal curve corresponds to a point $x = \ln t$on the abscissa of its normal transform curve and vice versa, the point m on the $\ln t$ scale will correspond to a point M_G on the t scale such that M_G bisects the area under the lognormal curve, and is thus its median, and in this case also its geometric mean. Realizing that $\ln M_G = m$, we also have $M_G = e^m$, as the antilog. Now, if we want to know $M(T)$, i.e., the area from $t = 0$ to T under the lognormal curve, we form $X = \ln T$and look up in standardized normal tables the corresponding normal tail area after determining how many standard deviations σ is Xaway from m to the left or to the right. Of course σ and $m = \ln M_G$ must be given to be able to plot the density curve of Eq. 1-31. The magnitudes of σ and m determine the shape of the lognormal distribution. Thus its shape changes as σ changes and also as the location of M_G changes.

The estimators of m and σ, from measured maintenance times t_L are

$$m = \sum_{i=1}^{n} \ln t_i / n = \frac{\ln t_1 + \ln t_2 + \ln t_3 + \cdots + \ln t_n}{n} \qquad (1\text{-}33)$$

and

$$\sigma = \sqrt{\sum_{i=1}^{n} (\ln t_i - m)^2/(n-1)} \qquad (1\text{-}34)$$

3. *The Gamma Distribution*

The gamma distribution is one of the most flexible distributions and can, probably better than any other, approximate any set of maintenance time data drawn from a population which is assumed to be continuously distributed and positively skewed. It has two parameters, exists only for positive values of t, includes the exponential distribution, and, in the limit, approaches the normal distribution. Certainly, in maintainability work it deserves as much attention as the lognormal distribution (Ref. 25). Besides, the gamma distribution has the advantage of mathematical tractability.

In its most general form, the gamma probability density function *At)* is of the form

$$f(t) = \frac{k^n}{\Gamma(n)} \cdot t^{n-1} e^{-kt} \qquad (1\text{-}35)$$

where $\Gamma(n)$ is called the gamma function given by

$$\Gamma(n) = \int_0^{\infty} x^{n-1} e^{-x}\, dx \qquad (1\text{-}36)$$

and k and n are positive constants (Ref. 26, Chapter 9). We call n the shape parameter and k the scale parameter. For $n = 1$, $\Gamma(n) = \Gamma(1) = 1$, and the gamma distribution becomes the exponential distribution

$$f(t) = ke^{-kt} \qquad (1\text{-}37)$$

with k representing the repair rate μ.

If $n \neq 1$, the gamma distribution will not have an exponential shape.

The cumulative probability, or the maintainability function $M(t)$ of the gamma distribution, is given by:

$$M(t) = \int_0^t f(x)\, dx = \frac{k^n}{\Gamma(n)} \int_0^t x^{n-1} e^{-kx}\, dx \qquad (1\text{-}38)$$

where $\Gamma(n)$ is defined by Eq. 1-36. For known values of k and n, $M(t)$ can be found by the use of tables of the Incomplete Gamma Function (Ref. 27) which tabulate the values of the following integral $I(t)$:

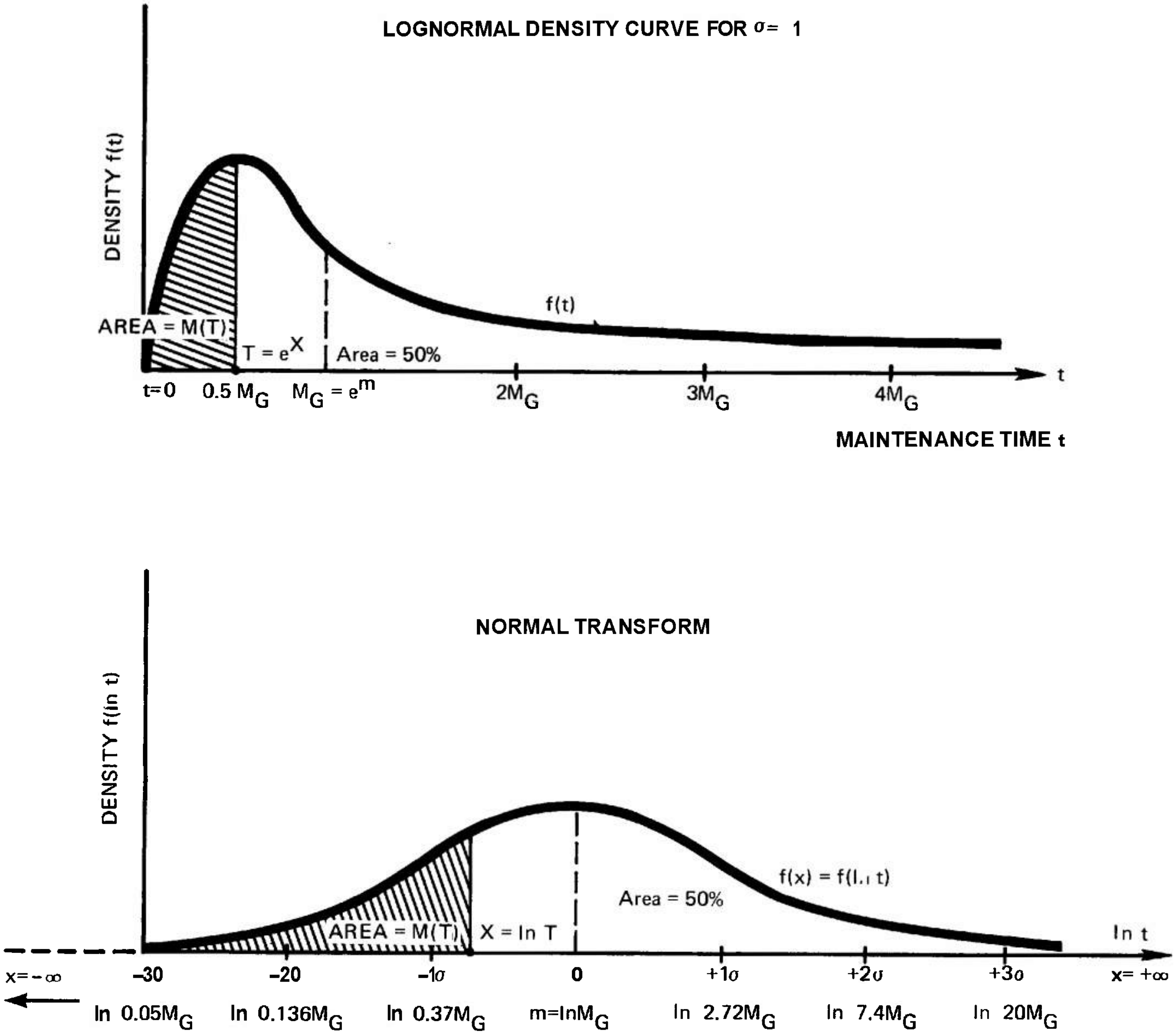


Figure 1-11. Normal Transformation of the Lognormal Distribution

$$I(t) = \frac{1}{\Gamma(n)} \int_0^t x^{n-1} e^{-kx} dx \qquad (1\text{-}39)$$

This integral $I(t)$ has the same shape as $M(t)$ of Eq. 1-38, except for the missing multiplication factor k^n. Using $I(t)$, we may write for $M(t)$,

$$M(t) = k^n I(t) \qquad (1\text{-}40)$$

which gives us direct numerical answers when k and n are known and reading $I(t)$ from tables.

The mean M of the gamma distribution has the simple form of

$$M = n/k \qquad (1\text{-}41)$$

which is the ratio of the shape parameter to the scale parameter, and the variance Var(t) is

$$\text{Var}(t) = n/k^2 = M/k \qquad (1\text{-}42)$$

so that standard deviation σ of the gamma distribution is

$$\sigma = \sqrt{n}/k = \sqrt{M/k} \qquad (1\text{-}43)$$

For positive integer values of the shape parameter n, the gamma density function $f(t)$ assumes a simple form because $\Gamma(n) = (n-1)!$, so that we get

$$f(t) = [k/(n-1)!]\, t^{n-1} e^{-kt} \qquad (1\text{-}44)$$

This is often referred to as the Special Erlangian distribution. It has the physical interpretation of a "stage-by-stage" repair. The corresponding maintainability function $M(t)$ is then given by

$$M(t) = 1 - \sum_{i=0}^{n-1} [e^{-kt}(kt)^i/i!] = \sum_{i=n}^{\infty} [e^{-kt}(kt)^i/i] \qquad (1\text{-}45)$$

and can be read directly from Poisson tables (Ref. 28) as $M(t) = D(X)$ for $X = n$, or as $M(t) = 1 - C(X)$ for $X = n - 1$ with the argument $U = kt$, since the summations are the cumulative terms of the Poisson distribution. Also, the density function $f(t)$ of *Eq.* 1-44 can be written in individual Poisson terms multiplied by the scale factor k, i.e.,

$$f(t) = ke^{-kt}(kt)^{n-1}/(n-1)! \qquad (1\text{-}46)$$

4. *The Weibull Distribution*

At times it is assumed that the field maintenance time of complex electronic equipment is Weibull distributed. In fact, it was found in some specific cases that the distribution of administrative times which delay field maintenance can be closely approximated by the Weibull distribution (Ref. 29, page 366). Of course, a gamma distribution also can be fitted as closely to such data. In general, the Weibull distribution in maintainability work has not become popular or useful.

The Weibull density function $f(t)$ (Ref. 30) is given by

$$f(t) = (n/k^n)t^{n-1} \exp[(t/k)^n] \qquad (1\text{-}47)$$

where n is the shape parameter and k is the scale parameter. The maintainability function $M(t)$ is then

$$M(t) = 1 - \exp[-(t/k)^n] \qquad (1\text{-}48)$$

and the mean maintenance time M is

$$M = k\Gamma(1 + 1/n) \qquad (1\text{-}49)$$

5. *The Poisson Distribution*

The Poisson distribution (Ref. 22, Chapter 8) s a discrete distribution with the density function p

$$P(N_S = n) = p(n, t) = e^{-kt}(kt)^n/n! \qquad (1\text{-}50)$$

which in maintainability work is interpreted as the probability that in time t a single repair channel, man, or crew will successfully complete exactly n maintenance actions in sequence, when the maintenance time of the actions is exponentially distributed with a mean maintenance time of $M = 1/k$. The variable is here $N_s(t)$, i.e., the number of successfully completed maintenance actions in time t, where t is fixed and the maintenance rate of $k = \mu$ is known. N_s theoretically can assume any integer value of n from zero to infinity. The values of $p(n,t)$ are found as individual terms $P(X)$ in Poisson tables, where the cumulative terms $C(X)$ also are tabulated (Ref. 28).

The mean E of the Poisson distribution is

$$E(N_s) = kt = t/M \qquad (1\text{-}51)$$

which is the expected number of successfully completed maintenance actions in time t, when the actions are performed in sequence.

Observing Eq. 1-50, we may write

$$P(N_s = 0) = p_0 = e^{-kt} \qquad (1\text{-}50a)$$

which is the probability no maintenance action will be completed in t,

$$P(N_s = 1) = p_1 = kte^{-kt} \qquad (1\text{-}50b)$$

which is the probability that exactly one and only one maintenance action will be completed in t

$$P(N_s = 2) = p_2 = [(kt)^2/2!]e^{-kt} \qquad (1\text{-}50c)$$

which is the probability that exactly two maintenance actions will be completed in t, etc.

As to the cumulative probability, we get

$$P(N_s \leq X) = \sum_{n=0}^{X} p_n = e^{-kt}[1 + kt + (kt)^2/2! + \cdots + (kt)^X/X!] \qquad (1\text{-}52)$$

which is the probability that X or less maintenance actions will be completed in time t, or, we may say, the probability that at the most X maintenance actions will be completed in time t. We may also write Eq. 1-52 in the form

$$P(N_s \leq X - 1) = \sum_{n=0}^{X-1} p_n = e^{-kt}[1 + kt + \cdots + (kt)^{X-1}/(X - 1)!] \qquad (1\text{-}53)$$

which is the probability that at the most $X - 1$ maintenance actions will be completed in t. Consequently, we get

$$P(N_s \geq X) = \sum_{n=X}^{\infty} p_n = e^{-kt}[(kt)^X/X! + (kt)^{X+1}/(X+1)! + \cdots] \qquad (1\text{-}54)$$

as the probability that X or more (or at least X) maintenance actions will be completed in t, so that by adding Eqs. 1-53 and 1-54, we get

$$P(N_s \leq X - 1) + P(N_s \geq X) = \sum_{n=0}^{\infty} p_n = 1 \qquad (1\text{-}55)$$

Fig. 1-12 shows the probability density and the cumulative probability of a Poisson distribution with a repair rate of $k = 0.5$ per hr and an observation time of $t = 10$ hr, so that $kt = 5$ is the mean or the expected number of completed maintenance actions in 10 hr, when equipments are repaired in sequence (i.e., no parallel simultaneous repairs take place in this repair channel).

The bars in the upper graph of Fig. 1-12 represent the probabilities of completing exactly $n = 0, 1, 2, 3,$ ... maintenance actions in 10 hr, while the lower graph of Fig. 1-12 represents the cumulative probability of completing at least n maintenance actions in t hours (i.e., n or more).

To conclude this discussion let us mention the very interesting relationship between the discrete Poisson distribution and the time-continuous gamma distribution. When we observe a Poisson maintenance process, we may ask what is the expected or mean time $E(t_n)$ to the occurrence of the nth successfully completed repair. This is given by

$$E(t_n) = n/k \qquad (1\text{-}56)$$

since the time t_n to the nth completed repair when the Poisson maintenance process starts at $t = 0$, is gamma distributed with the density $f(t_n)$

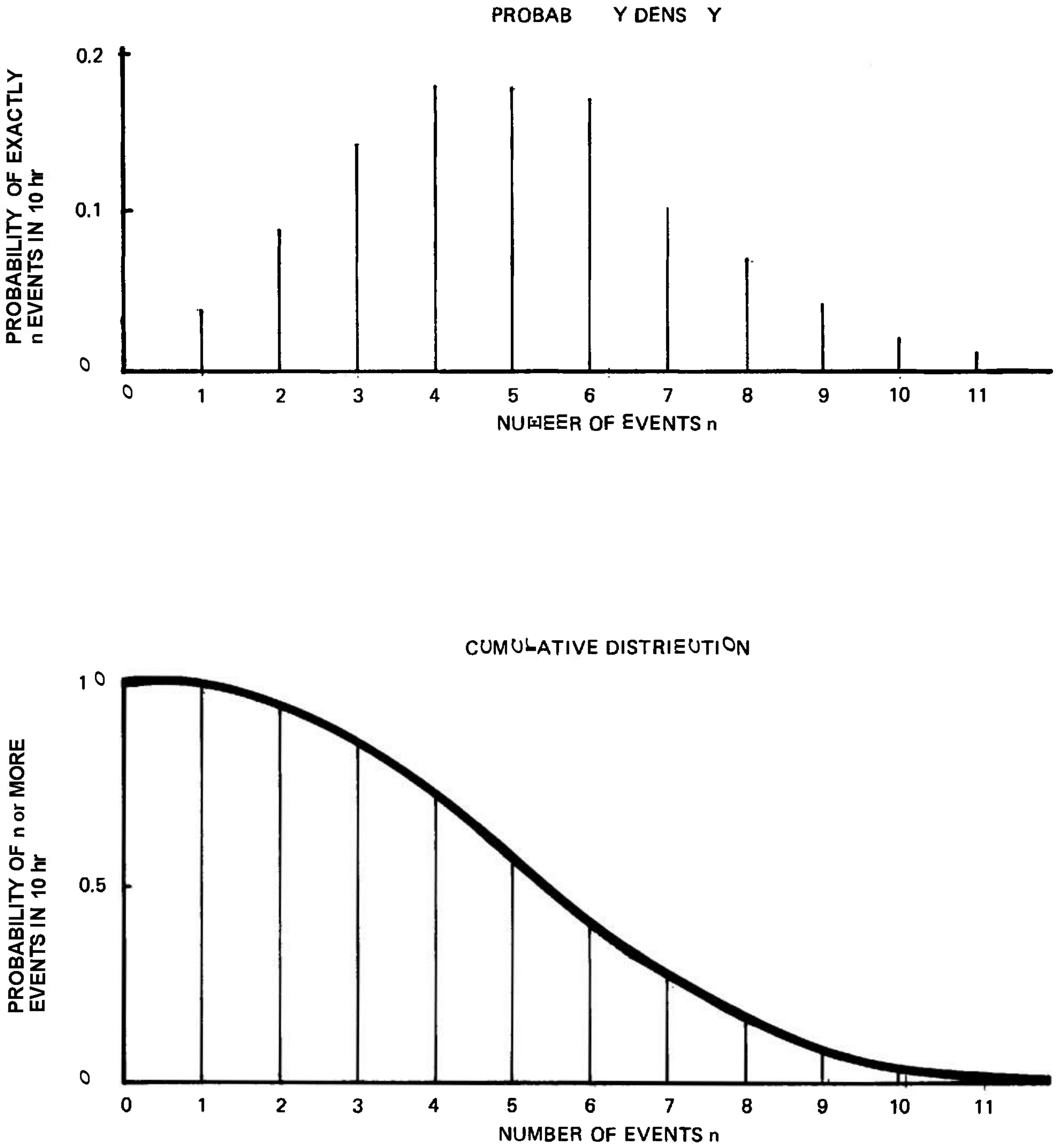


Figure 1-12. The Poisson Distribution ($a = kt = 5$)

$$f(t_n) = [k^n/(n-1)!]\, t_n^{n-1} \exp(-kt_n) \quad (1\text{-}57)$$

which, as we know, has the mean n/k.

6. The Binomial Distribution

Another discrete distribution frequently used in statistical work is the binomial distribution. Its application in maintainability appears to be rather limited.

The binomial distribution applies to so-called Bernoulli trials where each trial has the same probability of success P_s.

Assume that one has to perform a fixed number N of trials of the same kind where each trial can end with a success or with a failure and where S_N successes are counted in the N trials, so that there are N $-$ S_N failures. If one would observe the number of successes S_N in a repeated *series* of N trials, the number S_N would very likely change in each N trials. In fact, S_N is a random variable which may assume all integer values from zero to N, i.e., $S_N = K$, where K = 0, 1, 2, 3, . . ., N. The probability that S_N assumes a definite value of K is then given by the binomial probability density function p as

$$P_K = p(S_N = K) = \binom{N}{K} P_S^K (1 - P_S)^{N-K} \quad (1\text{-}58)$$

where, by definition

$$\binom{N}{K} = N!/[K!(N-K)!] \quad (1\text{-}59)$$

The mean value of this distribution is the expected or average number of successes $E(S_N)$ in N trials given by

$$E(S_N) = NP_S \quad (1\text{-}60)$$

That is, if one would run a large series of experiments, with N trials performed in each experiment, the averaged number of successes observed per N trials should approach the value of Eq. 1-60.

Observing the binomial probability density function, one can write the equations for S_N assuming any of the values K = 0, 1, 2, 3, . . ., N. For example, the probability that not a single success will occur in N trials, i.e., $p(S_N = 0)$, is given by

$$p_0 = p(S_N = 0) = (1 - P_S)^N \quad (1\text{-}61)$$

which is obtained by setting K = 0 in the Eq. 1-58. The probability that exactly one success will be observed in N trials is

$$p_1 = p(S_N = 1) = NP_S(1 - P_S)^{N-1} \quad (1\text{-}62)$$

The probability that exactly two successes will be observed in N trials is

$$p_2 = p(S_N = 2) = [N(N-1)/2!]\, P_S^2 (1 - P_S)^{n-2} \quad (1\text{-}63)$$

etc., until one gets the probability that all trials will be successful, i.e., $S_N = N$, is

$$p(S_N = N) = P_S^N \quad (1\text{-}64)$$

The cumulative binomial distribution $P(S_N \geq X)$ is then given by the partial sum of the probability densities p_K summing from K = X to K = N, i.e.,

$$P_K = (PS_N \geq X) = \sum_{K=X}^{N} p_K = \sum_{K=X}^{N} \binom{N}{K} P_S^K (1 - P_S)^{N-K} \quad (1\text{-}65)$$

which is the probability that in N trials X or more successes will be observed.

To perform these calculations one must know the probability of success P_S in any one Bernoulli trial. In real life one obtains only an estimate of P_S, because it is not possible to run an infinite series of N trials each to get the true value of P_S. Running just one set of N trials one obtains only an estimate of P_S, denoted by P_S, as

$$P, = S_N/N \quad (1\text{-}66)$$

How good this estimate of P_S is depends on the number N of trials performed. If one wants to determine the goodness of this estimate, he would be interested in the lower confidence limit of this estimated probability of success P_S, denoted by P_{SL}, such that with a confidence (or probability) of $1 - \alpha$ one could confidently make the statement that the true P_S exceeds P_{SL}, which is given by

$$P(P_S > P_{SL}) = 1 - \alpha \quad (1\text{-}67)$$

If the value of P_S was obtained from N trials in which S_N successes were observed, the lower confidence limit P_{SL} is given by

$$P_{SL} = \{1 + [(N - S_N + 1)/S_N][F_{(\alpha)(f_1)(f_2)}]\}^{-1} \quad (1\text{-}68)$$

where F is the α percentage point of Fisher's F distribution for $f_1 = 2(N - P_S + 1)$ and $f_2 = 2S_N$ degrees of freedom.

Fig. 1-13 shows a typical binomial distribution (density and cumulative) for $N = 100$ trials, and $P_S = 0.9$ and $1 - P_S = 0.1$ per trial.

In maintainability work the application of the binomial distribution could occur in cases where the duration of many maintenance actions of the same kind is observed, and one would be interested in obtaining an estimate of the probability (and confidence limit) that such specific action will be completed in a specified time t. Each action completed by the specified time t would be designated as a success and when it exceeds t it would be designated as a failure.

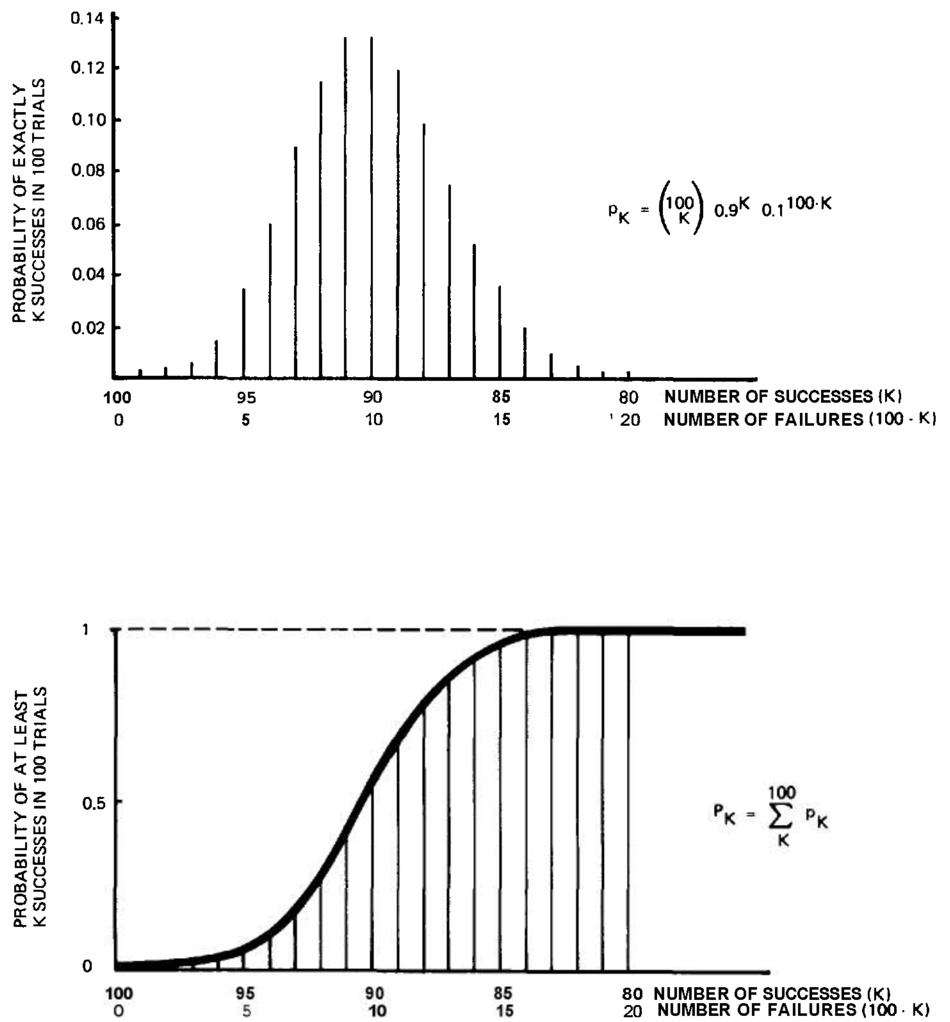


Figure 1-13. The Binomial Distribution ($N = 100$, $P_S = 0.9$)

SECTION III

EQUIPMENT CONSIDERATIONS IN MAINTAINABILITY

1-9 CATEGORIES OF EQUIPMENT

It would be convenient and would simplify the task of both the reliability and the maintainability engineer if different categories of systems and equipment could be treated in the same manner with respect to their reliability and maintainability characteristics. Unfortunately, this is not the case. Each category of equipment may require specific considerations which are peculiar to it. For example, reliability and maintainability considerations for mechanical systems (Ref. 20, Chapter 6) – systems in which there are moving parts subject to wear – have different maintenance requirements and implications for maintainability design than do electronic systems.

This paragraph contains a discussion of salient points of maintainability applicable to different categories of equipment, including electrical-electronic systems, electromechanical systems, hydraulic and pneumatic systems, optical systems, chemical systems, and systems containing nonreversible devices.

There are a number of considerations which affect maintainability design regardless of the category of equipment. These include:

1. The operational level at which maintenance is to be performed (organizational, direct support, general support, depot levels)

2. The system maintenance level (system, subsystem, equipment, group, unit, assembly, subassembly, stage, piece part)

3. The maintenance task to be performed (detection, diagnosis, correction, replacement verification).

In addition to these common considerations, there are those which are peculiar to the specific category of equipment. Among these are:

1. Equipment attributes such as accessibility, test points, connectors, controls, displays, inspection points, fittings, lubrication points, and packaging.

2. Maintenance methods such as module replacement, repair in place, periodic maintenance, adjustment, alignment, inspection, overhaul, remove, repair in shop, and reinstall.

3. Test methods such as built-in automatic checkout, monitoring, marginal testing, periodic check, and calibration.

1-9.1 ELECTRICAL-ELECTRONIC SYSTEMS

Electrical-electronic systems are in many ways the easiest to handle from a reliability and maintainability standpoint. More is known about their behavior, more reliability and maintainability data have been collected for such systems and prediction and demonstration techniques have been developed for these systems. Electrical systems generally are associated with the generation and distribution of electrical energy and may contain continuously rotating components, such as motors and generators. Electronic systems contain active as well as passive devices used for amplification, transformation, and shaping of electrical signals. They generally do not contain continuous rotating devices, but may contain intermittently operated electromechanical items, such as switches, relays, variable resistors, capacitors, and inductors.

Experience with reliability and maintainability of electronic systems has shown that where a constant hazard rate is experienced, (the flat bottom of the well-known bathtub curve in reliability), chance (random) failure is the predominant reliability phenomenon. Maintainability, in this case, primarily is concerned with corrective maintenance upon the occurrence of a failure. Indeed, it has been shown in such instances that the best maintenance policy may be to do no maintenance until failure occurs, the so-called hands-off or "leave well enough alone" policy. Studies have shown that where preventive maintenance, other than periodic test or performance monitoring, is performed, maintenance-induced failures often result. In these cases, and where the wearout portion of the failure rate curve is sufficiently far away in time, the assumption of the exponential failure distribution and the lognormal corrective maintenance distribution frequently have been shown to be valid for electronic systems.

A similar situation is true for electrical systems. In these cases—where rotating components such as motors, generators, and servos are used—wearout life characteristics can be expected to be approached at earlier points in time than for purely electronic systems in which no moving parts are involved. Preventive maintenance tasks—such as brush and contact inspection and replacement, lubrication and other servicing, or inspection of shafts and bearings for alignment and frictional wear—may be necessary in order to retain the system in its serviceable condition, effectively preventing the rising portion of the wearout curve from occurring too soon.

The inclusion of maintainability features and maintenance tasks in equipment design is usually simpler for electrical-electronic systems than for other types. Electrical-electronic systems lend themselves readily by their very nature to the use of automation with regard to monitoring, fault diagnosis, and verification. It is also simpler to achieve low corrective maintenance downtimes. Many of the studies and data collected as to the actual percentage of corrective maintenance times in the principal areas of detection, diagnosis, correction, and verification have been on electronic systems and equipment. Since corrective maintenance and the associated corrective maintenance tasks are generally of greater importance in electrical-electronic systems than preventive maintenance, maintainability characteristics which should be considered include:

a. built-in test points

b. built-in test equipment

c. automatic monitoring

d. automatic test and checkout

e. functional packaging into unit replaceable modules with provision for test points and failure indicators

f. controls

g. displays

h. connectors

i. parallel or standby redundancy to increase system availability

j. throwaway modules

k. the possibility of accomplishing a significant amount of corrective maintenance by replacement at the organizational and direct support level.

1-9.2 ELECTROMECHANICAL SYSTEMS

The primary difference between electromechanical systems and electrical-electronic systems is that mechanical actuating elements are utilized in electromechanical systems to perform some of the system prime functions in addition to electrical or electronic elements. Electromechanical systems may include such items as servo systems, actuators for moving missile control surfaces, autopilots, radar gun laying devices, tracking radars, and the like.

Electromechanical systems combine components in equipments which fail in different modes, and, therefore, have different failure distribution statistics. Some of the items may have constant hazard rates and thus obey an exponential failure distribution. Other parts may exhibit a hazard rate which increases with time and, therefore, may be described by one of a number of other distributions such as the Weibull distribution. For those parts which do have a constant hazard rate, corrective maintenance features are predominant; for those which have an increasing hazard rate, preventive maintenance features are more significant. Thus, one thing which distinguishes electromechanical systems from electronic systems is the necessity for concern with preventive maintenance features—such as periodic servicing, lubrication, and inspection—in addition to the corrective maintainability features provided for electronic systems.

1-9.3 MECHANICAL SYSTEMS

For purely mechanical systems, or those systems which are essentially mechanical, the situation with regard to maintainability considerations becomes quite different. Mechanical systems, in general, do not have constant hazard rates. They begin to wear out as soon as they are put to use. This does not mean that they necessarily have short wearout lives; it just means that friction and aging characteristics resulting from mechanical motion begin to exhibit themselves rather early. In order to obtain reasonable life expectancies or reasonable *MTBF*s, therefore, the maintainability designer's attention must be focused on those equipment considerations which will inhibit failures and will prolong component and equipment life.

One approach to this is to design long-life, low-friction elements, such as air bearings, or to use hard surface finishes. In many instances this may be costly and unrealistic, particularly when one considers the various environments in which the equipment will be expected to operate. This approach puts the emphasis on design for high reliability.

Another approach, which is often more cost-effective, is to recognize the essential nature of mechanical systems with regard to the physics of failure and to incorporate maintainability features during system design which will inhibit the rapidly rising wearout characteristic. Attention, therefore, must be on preventive maintenance features such as periodic inspection and replacement, lubrication, calibration and alignment, and overhaul (Ref. 20, Chapter 20). Indeed, for mechanical systems, this might be the most realistic means for achieving high operational readiness.

Cost-effective trade-offs between item life and maintenance intervals, maintenance personnel, and other maintenance resource requirements are of concern in mechanical systems. The ability to remove assemblies and components with a minimum of teardown emphasizes the need for modularization, interchangeability, and standardization. These are also important, of course, in electrical-electronic systems, but more difficult to accomplish in mechanical systems.

With regard to maintenance levels for mechanical systems, the simplest preventive maintenance functions—such as inspection, lubrication, removal and replacement, and adjustment and alignment—should be performed at organizational levels, assisted by Direct Support technicians and tools. Additional detailed maintenance tasks must be performed at the General Support level. For complex mechanical items, most corrective maintenance, repairs, and overhaul can be expected to be accomplished at the General Support and Depot levels. The concept of rotatable pools of mechanical components, assemblies, and equipments, such as the Army Direct Exchange Program (DX), is a feasible one. This concept, discussed in more detail in Chapter 5, is one in which forward level repairs are primarily accomplished by replacement of assemblies, components, and equipments, with detail repair in field operations performed at rear levels, and the repaired items returned to a repaired rotatable pool. As a matter of fact, when it is desirable to overhaul certain items after so many hours of use, the rotatable pool concept can be very cost-effective.

1-9.4 HYDRAULIC AND PNEUMATIC SYSTEMS

Hydraulic and pneumatic systems are examples of systems in which fluid flow is the primary energy transfer means. While there are instances of purely hydraulic and pneumatic systems, generally these types of equipment are combined with electrical or mechanical equipments to form electrohydraulic and other combination systems. Reliability and maintainability problems with respect to hydraulic and pneumatic systems are primarily concerned with pressure strengths, erosion, contamination and leakage of the fluids used (liquid or gas), and the reliability of seals, gaskets, and other sealing devices. Of concern to the designer then are the material and life characteristics of components, such as pressure vessels, piping, O-rings, gaskets, pumps, filters, and ports. Contamination from internal as well as external sources are of great importance. Maintainability design considerations are concerned with preventive maintenance as the principal means of obtaining long-lived hydraulic and pneumatic systems. Alignment, lubrication, visual indicators (such as sight gages, pressure and temperature indicators), oil and air spectral and chemical analysis, filter characteristics, and inspection and replacement are some of the primary maintainability considerations.

1-9.5 OTHER SYSTEMS

Among other categories of systems to which maintainability consideration may have to be given are optical and chemical systems. For fixed optical systems (no moving parts), reliability is generally high, and primary maintainability requirements are those of keeping the system clean, aligned, and calibrated. For electro-optical systems without moving parts, maintainability considerations for electronic systems apply. Similarly, when there are moving parts so that the systems are mechanico-optical or electromechanico-optical, then maintainability considerations for these types of equipments and systems, as discussed earlier in this paragraph, will also apply.

For chemical systems, maintainability considerations have to do with contamination, cleanliness, safety, visual inspection, chemical analysis, and with the specific nature of the chemical apparatus involved in the chemical reaction. The chemical system may contain features of several of the previously discussed categories of systems, and thus maintainability considerations of these will also apply where appropriate. Propulsion systems are examples of chemical systems.

1-10 NONREVERSIBLE DEVICES

Nonreversible devices are items which depend upon some physical, chemical, or biological reaction or effect which, once started, cannot be reversed or changed back to its original form or state. Ammunition, radioactive substances, and chemical processes are nonreversible devices. Bullets, bombs, and missiles are examples of the first; atomic bombs and nuclear power or propulsion of the second; napalm and rocket propellants of the third. Reliability and maintainability considerations for such devices are different from the categories of equipments discussed in par. 1-9. Because their reactions cannot be reversed and are thus not repairable once the action is initiated, it is essential that the mission reliability of such devices be high. Emphasis on these devices therefore has been and will continue to be high inherent reliability and safety. This does not mean that there are no maintainability considerations involved in nonreversible devices. It only means that once the mission has been started, maintainability can no longer be effected, such as is possible with a communication or radar system, aircraft, or tank.

Maintainability considerations for these types of devices primarily reside in the maintainability of the remaining parts of the system of which the nonreversible device is a part. These remaining parts include such items as launching and aiming devices, fuzing, and initiating devices. In the early guided missile days, many valuable lessons were learned with regard to reliability and maintainability. One of the principal lessons learned was that, although missiles are designed to operate for a short duration, measured in minutes, reliability design was originally performed so that the parts of the missile would operate only for the mission time. The necessity for test and checkout was not considered by the designers, and this resulted in many of the early missiles being worn out because of the need for frequent test and checkout and the accumulation of significantly more operating time than the missile components were designed for. In order to assure high mission reliability, however, it was necessary to exercise and test all those parts of the system except the nonreversible devices up to and including its fuzing circuitry.

With the emphasis almost completely on reliability in the early missile developments, there was a lack of maintainability considerations, and the consequent drastic effect on operational availability of the missile systems due to repeated testing helped spur the development of maintainability as a system design discipline.

Maintainability considerations, therefore, of nonreversible devices have to do with the state of readiness prior to the mission start. They have to do with designing features into the equipment which emphasize periodic test and checkout, and the prediction of the overall device effectiveness from test results. Of prime concern are:

1. The ability to simulate the operation of the nonreversible device where necessary in order to properly exercise and test the total system
2. The ability to safely test the system under various operational situations and environments without initiating the nonversible reaction
3. The ability to obtain high confidence levels of successful operation once committed to the mission.

This places great emphasis on the areas of safety, test, and checkout as prime equipment design considerations for the maintainability of nonreversible devices. Such test and checkout ranges all the way from relatively simple manual tests to highly sophisticated and complex automatic checkout equipment and procedures.

1-11 DESIGN GUIDES

A number of equipment design guides for maintainability have been written. These guides, in general, discuss the maintainability design features and problems in terms of maintenance methods, maintenance tasks or actions, maintenance time distributions, maintenance levels, and equipment attributes. No attempts are made to relate these to the maintainability quantitative requirements, except by implication in generic terms. In addition, many of these design guides, including AMCP 706-134 (Ref. 5), contain specific anthropometric and other human factors Considerations, specific equipment design features, and designer's checklists which can be applied to a wide variety of equipment and maintenance concepts (Refs. 31-39). Chapter 5 treats equipment design for maintainability in greater detail.

REFERENCES

1. MIL-STD-470, ***Maintainability Program Requirements,*** 21 March 1966.
2. MIL-STD-471, ***Maintainability Demonstration,*** 15 February 1966.
3. MIL-HDBK-472, ***Maintainability Prediction,*** 24 May 1966.
4. MIL-STD-721B, ***Definition of Effectiveness Terms for Reliability, Maintainability, Human Factors, and Safety,*** 25 August 1966.
5. AMCP 706-134, Engineering Design Handbook, ***Maintainability Guide for Design.***
6. A. Coppola and A. D. Pettinato, "RADC Case Histories in R & M Demonstration", ***Proceedings of the 1966 Annual Symposium on Reliability,*** IEEE, 1966, pp. 395-408.
7. J. E. Daveau and A. Coppola, "Reliability and Maintainability Case Histories", ***Proceedings of the Sixth Annual R/M Conference,*** July 1967, Society of Automotive Engineers, pp. 582-586.

8. D. W. Lowry, "Maintainability Demonstration Test Performed on a Computer System", *Proceedings of the 1970 Annual Symposium on Reliability,* IEEE, **1970,** pp. **308-318.**
9. A. C. Hall, *Relationship Between Operational Effectiveness of Electronic Systems and Maintenance Minimization and Training,* in *Proceedings, Symposium on Electronics Maintenance,* OASD (R&D), Washington, D.C., **1955.**
10. F. L. Ankenbrandt, Ed., *Electronic Maintainability,* Vol. **3,** Engineering Publishers, Elizabeth, N.J., **1960.**
11. Headquarters AMC Memo AMCRD-ES, **21** Jan. **1971.** Subject: "Economy vs. Life Cycle Cost", signed by **Maj.-Gen.** John R. Guthrie, Director, Research, Development and Engineering.
12. Weapons System Effectiveness Industry Advisory Committee (WSEIAC), "Requirements Methodology, Final Report of Task Group I", Air Force Systems Command, AFSC-TR-65-1, Jan. **1965** (AD-458 **453).**
13. **AR 702-3** *Army Materiel Reliability, ,Availability, and Maintainability.*
14 AR **750-1,** *Maintenance Concepts.*
15 TM **38-703,** *Integrated Logistic Support, 5* Volumes, Dept. of the Army, **1969.**
16 AMCR **750-42,** *Maintenance of Supplies and Equipment; AMC Maintenance Engineering Activity.*
17. **MIL-STD-785A,** *Reliability Program for Systems and Equipment Development and Production,* **28** March **1969.**
18. *Tables of the Exponential Function,* National Bureau of Standards, Applied Mathematics Series No. **14,** June **1951.**
19. B. **S.** Blanchard, Jr. and E. E. Lowery, *Maintainability Principles and Practices,* McGraw-Hill Book Company, Inc., N.Y., **1969.**
20. Igor Bazovsky, *Reliability Theory and Practice,* Prentice-Hall, Inc., Englewood Cliffs, N.J., **1961.**
21. Orest A. Meykar, "Maintainability Terminology Supports the Effectiveness Concept", IEEE Transactions on Reliability **R-16, No. 1, 10-15 (1967).**
22. M. J. Moroney, *Facts from Figures,* Penguin Books, Inc., Baltimore, Md., **1951.**
23. A. H. Bowker and G. J. Lieberman, *Engineering Statistics,* Prentice-Hall, Inc., Englewood Cliffs, N.J., **1959.**
24. Melvin B. Kline, "Application of the Lognormal Distribution to Maintainability Design", The Logistics Spectrum 1, No. **2, 13-20 (1967-68).**
25. *An Application of the Gamma Distribution to Downtime Distribution Functions,* Maintainability Bulletin No. **13,** Electronic Industries Association, Washington, D.C., September **1968.**
26. Paul L. Meyer, *Introductory Probability and Statistical Applications,* Addison-Wesley Publishing Company, Inc., Reading, Massachusetts, **1965.**
27. Karl Pearson, *Tables of the Incomplete Gamma Function,* Cambridge University Press, Cambridge, England, **1934.**
28. *Tables of the Individual and Cumulative Terms of Poisson Distribution,* Defense Systems Dept, General Electric Company, D. Van Nostrand Company, Inc., Princeton, N.J., **1962.**
29. William H. von Alven, Ed., *Reliability Engineering,* Prentice-Hall, Inc., Englewood Cliffs, N.J., **1964.**
30. Waloddi Weibull, "A Statistical Distribution Function of Wide Applicability", Journal of Applied Mechanics 18, No. **3, 293-297 (1951).**
31. L. V. Rigby, J. I. Cooper, and W. A. Spickard, *Guide to Integrated System Design for Maintainability,* ASD Technical Report **61-424,** U S Air Force, Wright-Patterson Air Force Base, Ohio, October **1961,** AD-271 **477.**
32. Melvin B. Kline, *Maintainability Considerations in System Design,* Doctoral Dissertation, Dept. of Engineering, University of California, Los Angeles, **1966.** University Microfilms, Ann Arbor, Michigan.
33. J. M. McKendry, et al., *Maintainability Handbook for Electronic Equipment Design,* Supplement IV to NAVTRADEVCEN **330-1,** US Naval Training Device Center, Port Washington, N.Y., April **1960, AD-241 284.**
34. J. G. Wohl, "Why Design for Maintainability?", *IRE* Transactions on Human Factors in Electronics HFE-2, **87-92 (1961).**
35. **J.** D. Folley, Jr. and J. W. Altman, *Guide to Design of Electronic Equipment for Maintainability,* WADC Technical Report **56-218,** U S Air Force, Wright-Patterson Air Force Base, Ohio, April **1956,** AD-101 **729.**
36. J. W. Altman, et al., *Guide to Design of Mechanical Equipment for Maintainability,* ASD

Technical Report 61-381, US Air Force, Wright-Patterson Air Force Base, Ohio, August 1961, AD-269 332.

37. *Maintainability Design Criteria for Designers of Shipboard Electronic Equipment,* NAVSHIPS 94324, Change No. 2, March 1965.

38. *Human Factors Design Standards for the Fleet Ballistic Missile Weapon System, Vol. II – Design of Equipment,* NAVWEPS OD 18413A, August 1962.

39. J. Cooper, "An Integrated Approach to System Maintainability", Chapter 14 in *Electronic Maintainability,* Vol. 3, ed. by F. L. Ankenbrandt, Engineering Publishers, Elizabeth, N.J., 1960.

CHAPTER 2

SYSTEM EFFECTIVENESS

SECTION I

INTRODUCTION

2-1 GENERAL

As pointed out in par. 1-2, the realization that in many cases a more cost-effective system can be obtained by trading off some reliability for the ability to maintain a system has led to a considerable research and development effort into describing maintainability as an engineering discipline. This, in turn, has led to the concepts of operational readiness, availability, and system effectiveness as elements of system worth—the ultimate measure of a system. Army Technical Manual TM 38-703-1 (Ref. 1) states:

"The worth of a system is determined primarily by the effectiveness with which it does its job. Subsequent to World War 11, system reliability came to the forefront as a measure of system performance. More recently the systems approach requiring such consideration as system maintainability and availability has received increasing attention. All of these factors are highly interdependent and tend to make the measurement of system performance very complex. A measure of system performance may be generally defined as a quantified assessment of the ability of a system to fulfill a specified function, when both the system and function are thoroughly defined. The parameter to be defined by such a measure is called system effectiveness." Specification of the support environment is also essential in system effectiveness assessment.

It is recognized by system designers today, particularly for systems that are not of the "one-shot" type but which are required to have a long operational life with repeated usage, that system effectiveness considerations, in which maintainability is indeed as significant a parameter as reliability, consist of more than just system performance and mission reliability considerations. Obviously, with repeated usage ease of maintenance assumes a very significant role.

An Air Force study on system effectiveness states:

"The high cost and complexity of modem military systems require the most efficient management possible to avoid wasting significant resources on inadequate equipment.

"Efficient systems management depends on the successful evaluation and integration of numerous different but interrelated system characteristics such as reliability, maintainability, performance, and costs. If such evaluation and integration are to be accomplished in a scientific rather than intuitive manner, a method must be formulated to assess quantitatively the effects of each system characteristic on overall system effectiveness." (Ref. 2).

How do availability, readiness, and maintainability relate to other system parameters? Considerable attention has been paid to this question in recent years, and many concepts have been proposed. Of these concepts, system effectiveness has been elevated to the position of highest rank.

The notions of effectiveness and measures of effectiveness are not new. Such measures have been used for many years for determining how well a device performs or for comparing one device with another. The use of figure-of-merit comparison is well known, e.g., the gain-bandwidth product for electronic amplifiers.

The extension to measuring system performance on some overall mission basis is, however, relatively recent. Many of the operations research and system analysis efforts, which became prominent starting in World War 11, were initiated in order to find quantitative methods for assessing and optimizing system effectiveness. Cost-effectiveness considerations have become a major item of system design in defense and space systems, due largely to the emphasis given by former Assistant Secretary of Defense Hitch.

A system is designed to perform a function or set of functions (meet a need). System effectiveness is a measure of how well the system performs its intended function in its operating environment. In order to be a useful measure, it is necessary to express system effectiveness in quantitative terms. A number of such measures have been derived, most in a probability sense:

The effectiveness of a system, in the final analysis, can only be really measured when the system is performing its mission in the environment for which it was designed or other accurately simulated environment. Of great concern, however, is how system effectiveness can be predicted while the system design concepts are being formulated and again later when the system is being designed and evaluated. Thus, most system effectiveness methodologies deal more with the predictive design and test aspects of effectiveness' of the system than with the later use of the system.

The effectiveness of a system, then, is concerned with

1. The ability of the system to perform satisfactorily for the duration of an assigned mission, often stated as mission reliability

2. The ability of the system to begin performing its mission when called upon to do so, often stated as operational readiness or availability; and

3. The actual performance measures of the system in terms of its performance functions and environment in which it performs, often stated as design adequacy or capability.

These may be related, as in AMCP 706-134 (Ref. 3), as
System Effectiveness = Reliability X Availability X Performance (How Long?) (How *Often?)* (How Well?)

Just about all system effectiveness methodologies which have been developed in the past 10 to 15 yr are concerned with these fundamental questions in one way or another. They include such system attributes as performance parameters, reliability, maintainability, and logistic supportability, as well as such other attributes as human factors, safety, and standardization, all of which condition the ability of a system to perform its assigned missions. (*See* Fig. 1-1.)

It is instructive, therefore, to discuss and compare the various concepts and methodologies that have been put forth and are being used today, the semantic barriers (sometimes very great) that have arisen, their points of similarity and difference, and the ease or difficulty of their application.

2-2 SYSTEM EFFECTIVENESS CONCEPTS

The three generally recognized components of system effectiveness described in the previous paragraph (reliability, availability, performance) will be used as the basis for description and comparison of the concepts and formulations of system effectiveness which are currently in use. It should be recognized that all of these effectiveness components must be derived from an analysis of the operational needs and mission requirements of the system, since it is only in relation to needs and missions that these three basic components can be meaningfully established.

Many semantic difficulties arise when discussing system effectiveness and its components. These difficulties result from the fact that some people use the same words to mean different things or different words to mean the same thing.

2-2.1 THE ARINC CONCEPT OF SYSTEM EFFECTIVENESS

One of the early attempts to develop concepts of system effectiveness was delineated by the ARINC Research Corporation in Chapter I of their book, Reliability *Engineering* (Ref. 4). It contains some of the earliest published concepts of system effectiveness and represents one of the clearest presentations of these concepts, from which many of the subsequent descriptions have been derived. The definition of system effectiveness in this early work is as follows: "System effectiveness is the probability that the system can successfully meet an operational demand within a given time when operated under specified conditions".

This definition includes the following concepts:

1. That system effectiveness can be measured as a probability

2. That system effectiveness is related to operational performance

3. That system effectiveness is a function of ***time***

4. That system effectiveness is a function of the ***environment*** or conditions under which the system is used

5. That system effectiveness may vary with the ***mission*** to be performed.

What is not obvious in this definition, with regard to system effectiveness as a function of time, is that there are two kinds of time to be considered. One is the ***point*** in time in which we wish to make use of the system and whether or not the system is usable at that time. The other is the ***continued period*** of time, starting with this point in time, for which we want the system to continue to operate (mission time). The three components of system effectiveness, according to the ARINC model Fig. 2-2(C), are mission reliability, operational readiness, and design adequacy, as shown in Fig. 2-1. Definitions of the words used in this figure are given in Table 2-1. These are essentially the three factors which contribute to system effectiveness as indicated at the beginning of this paragraph. A study of these definitions and their meaning is of particular significance. While most of these definitions are left to the reader to study, certain definitions and their meanings or implications will be discussed in more detail. This will be particularly helpful when other concepts of system effectiveness which have been developed are discussed.

Although it is not essential to describe system effectiveness and its component parts in terms of probabilities as opposed to other quantitative measures, it has often been found to be convenient to do so. The ARINC model may be expressed such that system effectiveness probability P_{SE} is the product of three probabilities as follows:

$$P_{SE} = P_{OR} \times P_{MR} \times P_{DA} \qquad (2\text{-}1)$$

where

P_{OR} = operational readiness probability

P_{MR} = mission reliability probability

P_{DA} = design adequacy probability

This equation states that the effectiveness of the system is the product of three probabilities: (1) the probability that the system is operating satisfactorily or is ready to be placed in operation when needed, (2) the probability that the system will continue to operate satisfactorily for the period of time required for the mission, (3) the probability that the system will successfully accomplish its mission given that it is operating within design limits (Fig. 2-2(A)).

Each of these terms may then be developed in terms of the specific problem. (See, for example, Chapter II of Ref. 4.)

2-2.2 THE AIR FORCE (WSEIAC) CONCEPT

A more recent definition of system effectiveness results from the work of the Weapon System Effectiveness Industry Advisory Committee (WSEIAC) established in late 1963 by the Air Force Systems Command "to provide technical guidance and assistance to Air Force Systems Command in the development of a technique to apprise management of current and predicted weapon system effectiveness at all phases of weapon system life". Five task groups worked for one year on various aspects of this problem. The result of these efforts has been published as Air Force Systems Command Technical Reports TR-65-1, TR-65-2, TR-65-3, TR-65-4, TR-65-5, and TR-65-6 (Ref. 2). The WSEIAC definition of system effectiveness, Fig. 2-2(B), is

"***System effectiveness*** is a ***measure*** of the extent to which a system may be expected to achieve a set of specific mission requirements and is a function of availability, dependability, and capability" (Ref. 2).

This definition may be expressed as

$$E = ADC \qquad (2\text{-}2)$$

where

A = ***availability,*** a measure of the system condition at the start of a mission, when the mission is called for at an unknown (random) point in time

D = ***dependability,*** a measure of the system condition at one or more points during the performance of the mission, given the system condition (availability) at the start of the mission

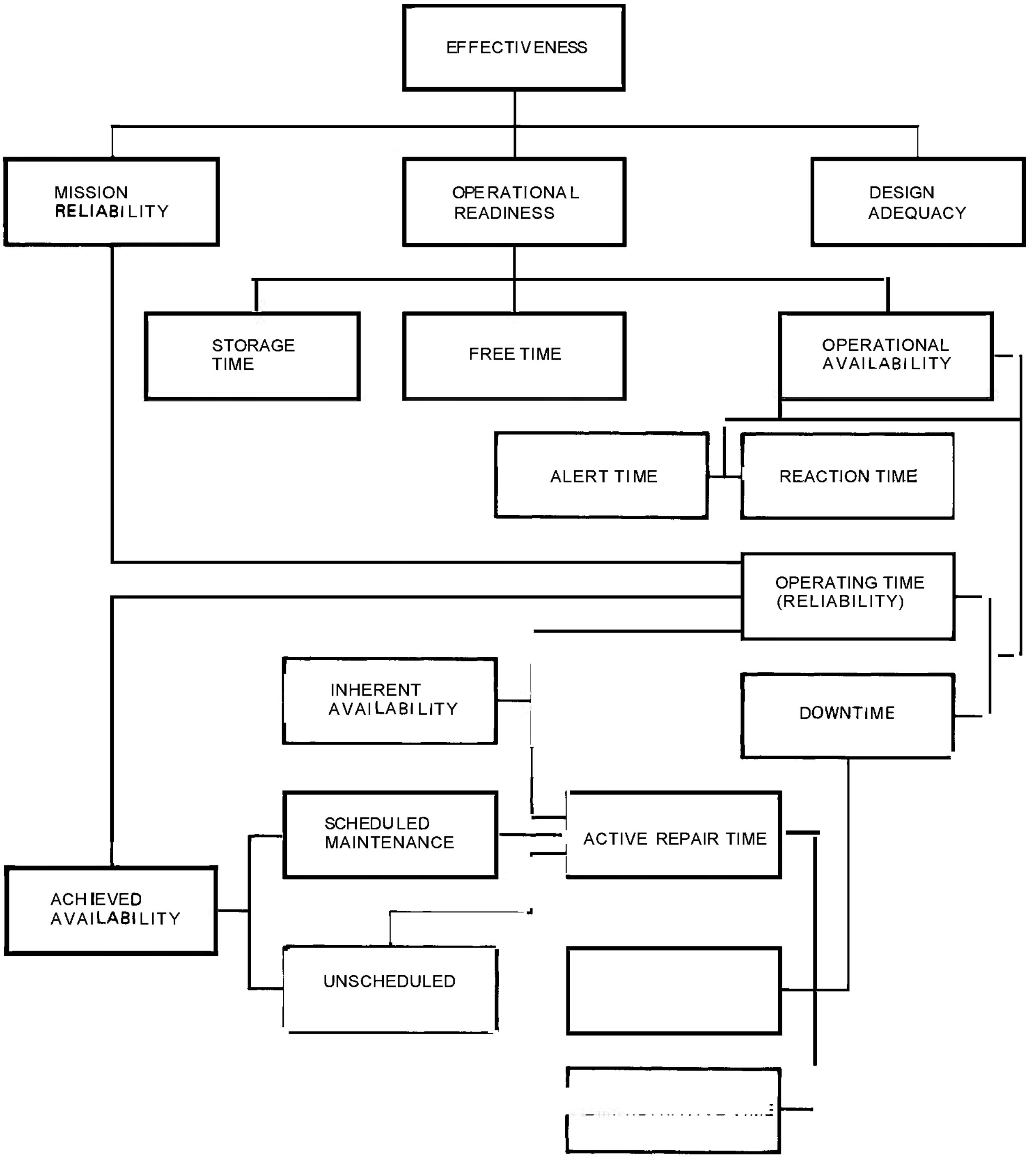


Figure 2-1. Concepts Associated With System Effectiveness
Adapted from: William H. vanAlven, Ed., *Reliability Engineering*, © 1964 by ARINC Research Corporation. Used with permission of Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

TABLE 2-1.
DEFINITIONS

Definitions of Concepts:

System Effectiveness is the probability that the system can successfully meet an operational demand within a given time when operated under specified conditions.

System Effectiveness (for a one-shot device such as a missile) is the probability that the system (missile) will operate sucoessfully (kill the target) when called upon to do so under specified conditions.

Reliability is the probability that the system will perform satisfactorily for at least a given period of time when used under stated conditions.

Mission Reliability is the probability that, under stated conditions, the system will operate in the mode for which it was designed (i.e., with no malfunctions) for the duration of a mission, given that it was operating in this mode at the beginning of the mission.

Operational Readiness is the probability that, at any point in time, the system is either operating satisfactorily or ready to be placed in operation on demand when used under stated conditions, including stated allowable warning time. Thus, total calendar time is the basis for computation of operational readiness.

Availability is the probability that the system is operating satisfactorily at any point in time when used under stated conditions, where the total time considered includes operating time, active repair time, administrative time, and logistic time.

Intrinsic Availability is the probability that the system is operating satisfactorily at any point in time when used under stated conditions, where the time considered is operating time and active repair time.

Design Adequacy is the probability that the system will accomplish its mission successfully, given that the system is operating within design specifications.

Maintainability is the probability that, when maintenance action is initiated under stated conditions, a failed system will be restored to operable condition within a specified total downtime.

Repairability is the probability that a failed system will be restored to operable condition within a specified active repair time.

Serviceability is the degree of ease or difficulty with which a system can be repaired.

TABLE 2-1.
DEFINITIONS (Cont.)

Definitions of Time Categories:

Operating time is the time during which the system is operating in a manner acceptable to the operator, although unsatisfactory operation (or failure) is sometimes the result of the judgment of the maintenance man.

Downtime is the total time during which the system is not in acceptable operating condition. Downtime can, in turn, be subdivided into a number of categories such as active repair time, logistic time, and administrative time.

Active repair time is that portion of downtime during which one or more technicians are working on the system to effect a repair. This time includes preparation time, fault-location time, fault-correction time, and final checkout time for the system, and perhaps other subdivisions as required in special cases.

Logistic time is that portion of downtime during which repair is delayed solely because of the necessity for waiting for a replacement part or other subdivision of the system.

Administrative time is that portion of downtime not included under active repair time and logistic time.

Free time is time during which operational use of the system is not required. This time may or may not be downtime, depending on whether or not the system is in operable condition.

Storage time is time during which the system is presumed to be in operable condition, but is being held for emergency—i.e., as a spare.

Alert time is that element of uptime during which the system is awaiting a command to engage in its mission.

Reaction Time is that element of uptime needed to initiate a mission, measured from the time the command is received.

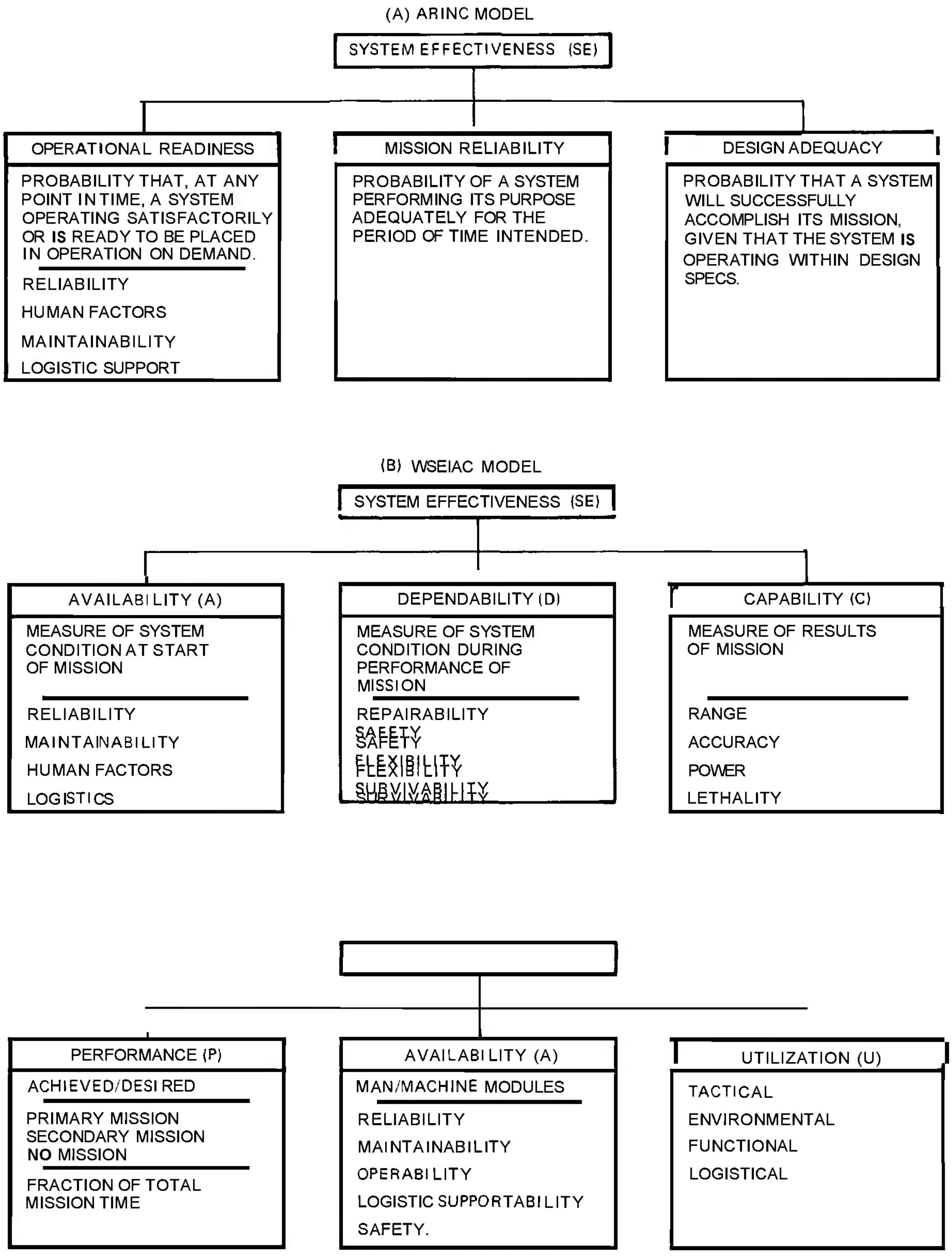


Figure 2-2. System Effectiveness Models

Reprinted from *Maintainability Principles & Practices,* by B. S. Blanchard, Jr. and E. E. Lowery. Copyright 1969, McGraw-Hill Book Co. Used with permission of McGraw-Hill Book Company.

C = capability, a measure of the ability of the system to achieve the mission objectives, given the system condition during the mission (dependability).

These are usually expressed as probabilities as follows:

1. A is a vector array of various state probabilities of the system at the beginning of the mission.

2. D is a matrix of conditional probabilities over a time interval, conditional on the effective state of the system during the previous time interval.

3. C is also a delineal probability matrix representing the performance spectrum of the system, given the mission and system conditions—expected figures of merit for the system.

The similarity of the WSEIAC definitions to the ARINC definitions should be noted.

2-2.3 THE NAVY CONCEPT

In the early 1960's, under the sponsorship of the Systems Effectiveness Branch of the Office of Naval Material, the Navy developed a system effectiveness concept (Fig. 2-2(C)), which also combines three basic system characteristics—performance, availability, and utilization (Ref. 5). It can be expressed as "a measure of the extent to which a system can be expected to complete its assigned mission within an established time frame under stated environmental conditions. It may also be defined mathematically as "the probability that a system can successfully meet an operational demand throughout a given time period when operated under specified conditions".

It has been formulated as follows:

$$E_s = PAU \tag{2-3}$$

where

E_s = index of system effectiveness

P = index of system performance—a numerical index expressing system capability, assuming a hypothetical 100% availability and utilization of performance capability in actual operation

A = index of system *availability*—numerical index of the extent to which a system is ready and capable of fully performing its assigned mission(s)

U = index of system *utilization*—a numerical index of the extent to which the performance capability of the system is utilized during the mission.

The components of the Navy model are not as readily compared as are the ARINC and WSEIAC models. The Navy has stated that "the terms PU and A are similar, respectively, to the WSEIAC terms C and AD" (Ref. 6). In this same reference, the Navy states that it "translates its terms PAU into the analytic terms P_C and P_T" in which

P_C = performance capability—a measure of adequacy of design and system degradation, and

P_T = detailed time dependency—a measure of availability with a given utilization.

Thus, the Navy model is compatible with the WSEIAC model (see Ref. 6) in the following manner:

$$f(P, A, U) = f(P_C, P_T) \approx f(A, D, C) \tag{2-4}$$

The WSEIAC, Navy, and ARINC concepts of system effectiveness are depicted in Fig. 2-2 (Refs. 7 and 8).

2-2.4 OPERATIONAL READINESS, AVAILABILITY, AND DEPENDABILITY

The terms operational readiness, availability, and dependability have similar connotations. As shown in Fig. 2-1, one concept of operational readiness includes total calendar time, while availability includes only desired use time. These are usually termed point concepts, since they refer to the ability of the system to operate at any given point in time when called upon to do so.

Mission reliability and dependability are terms used to depict the ability of the system to operate effectively for a specified "mission" time period, usually conditional on its being operable at the start of the period.

Unfortunately, there has been considerable overlap in the use of these terms during this period of intensive development of the concepts of system effectiveness, operational readiness, dependability, availability, and related ideas. The paragraphs that follow are an attempt to clear up some of this confusion.

2-2.4.1 Operational Readiness

A definition of operational readiness is put forth by ARINC:

"Operational readiness is the probability that, at any point in time, the system is either operating satisfactorily or ready to be placed in operation on demand when used under stated conditions, including stated allowable warning time" (Ref. 4).

As noted in this definition, this concept uses total calendar time as the basis for the computation of operational readiness (see Fig. 2-1). Others have used the term operational readiness in different contexts, varying from similar to, or synonymous with, dependability (not a point concept) on the one hand, to the availability of a specific number of systems composed of multiple devices on the other hand. Some detailed modeling techniques of operational readiness are presented in par. 2-4.2.1.

2-2.4.2 Availability

Availability (see Fig. 2-1) has generally been understood to include a relationship between uptime (reliability) and downtime (maintainability). In general, availability may be defined as the ratio of the total time the system is capable of performing its function (uptime) to the total time it is capable plus the time it is down for maintenance (uptime plus downtime). It is usually expressed as a percentage or a probability, for example:

"Availability is the probability that the system will operate satisfactorily at any point in time when used under stated conditions."

At least three kinds of availability have been defined. These are inherent (intrinsic) availability, achieved availability, and operational availability (Refs. 3 and 9).

Inherent or ***intrinsic availability*** A_i takes into account, in the calculation of the availability ratio, only those items which are inherent in the system design. It generally includes only active repair time items in the calculation of downtime, excluding such items as preventive maintenance and delay times due to administrative delays, personnel delays, and supply delays. Thus, it is a measure only of the intrinsic design variables controllable by the system designer.

Achieved availability *A,* is the measure of the availability of a system, including preventive maintenance in an ideal support environment (no delay time).

Operational availability *A,* is the extension to the actual operating environment and includes delay times as well.

All three cases have been discussed in par. 1-7.3 and defined by the steady-state Eqs. 1-23, 1-24, and 1-26. More sophisticated equations and modeling techniques are presented in par. 2-4.2.2. See also Fig. 2-1 and Table 2-1 for concepts and definitions associated with system effectiveness.

Because steady-state availability is basically a simple concept, it has often received more attention as a trade-off relationship and system design measure than have the other concepts.

2-2.4.3 Dependability

Although availability is a simple and appealing concept at first glance, it is a point concept, i.e., it refers to the probability of a system being operable at a random ***point*** in time. However, the ability of the system to continue to perform reliably for the duration of the desired operating (mission) period is often more significant. Operation over the desired period of time depends then on clearly defining system operating profiles. If the system has a number of operating modes, then the operating profile for each mode must be considered.

The term ***mission reliability*** has been used by some to denote the system reliability requirement for a particular interval of time. Thus, if the system has a constant failure rate region, so that its reliability R can be expressed as

$$R = \exp(-At) \quad (2\text{-}5)$$

where

A = failure rate = $1/MTBF$
t = time for mission

then mission reliability R_M for a mission duration of T is expressed as

$$R_M = \exp(-AT) \quad (2\text{-}6)$$

This reliability assessment, however, is conditional upon the system being operable at the beginning of its mission, or its (point) availability.

In order to combine these two concepts, the word "effectiveness" is sometimes utilized. If the system is operating within its design specifications so that $P_{DA} = 1$, then system effectiveness may be construed simply as the product of the probabilities that the system is operationally ready and that it is mission reliable.

If A is the mean availability of a system at any point in time t_o when we want to use the system, and if R_M is the system reliability during mission time T, then system effectiveness E, not including performance, may be defined as

$$E = AR_M \tag{2-7}$$

Thus, A is a weighting factor, and E represents an assessment of system ability to operate without failure during a randomly chosen mission period.

One concept of dependability, developed for the Navy (Ref. 10), takes into account the fact that, for some systems, a failure which occurs during an operating period t_1, may be acceptable if the failure can be corrected in a time t_2 and the system continues to complete its mission. According to this concept, dependability may be represented by

$$D = R_M + (1 - R_M)M_0 \tag{2-8}$$

where

D = system *dependability*—or the probability that the mission will be successfully completed within the mission time t_1 providing a downtime per failure not exceeding a given time t_2 will not adversely affect the overall mission.

R_M = mission *reliability*—or the probability that the system will operate without failure for the mission time t_1.

M_o = operational *maintainability*—or the probability that when a failure occurs, it will be repaired in a time not exceeding the allowable downtime t_2.

This definition is useful for some long duration naval missions in which system or equipment failures do not necessarily result in catastrophic events or cause mission aborts.

If we assume that the capability part of the system effectiveness formulation is 1, then we can write that

$$E = AD = A[R_M + (1 - R_M)M_0] \tag{2-9}$$

In the case where no maintenance is allowed during the mission ($t_2 = 0$ or $M_o = 0$), as in the case of a missile, then this reduces to Eq. 2-7.

$$E = AD = AR_M \tag{2-9a}$$

This concept of dependability is compatible with the WSEIAC model and, indeed, can be taken into account in the dependability state transition matrices.

There are cases in which availability or dependability, or even capability, become the dominant factors with regard to the specific system and its mission requirements. There are complex cases in which the system has multiple mode missions. There are other cases in which the system is essentially one of single mode missions. In these cases, the effectiveness model used can and should be kept simple. The versatility of the concepts previously discussed is that they can be generally applied to a complex system. The transformation to system worth, if done properly, will accomplish such simplifications.

2-2.5 PERFORMANCE, UTILIZATION, CAPABILITY, AND DESIGN ADEQUACY

It should be readily apparent that, in the context of system effectiveness definitions, these words are generally similar in their meaning and application. They may be separated into two notions:

1. The capability of the system to perform its tasks as originally specified.

2. The capability of the system to meet new requirements, such as longer range, higher accuracy, and/or different environments (higher or lower altitudes, different terrain, more severe weather, shock or vibration, or new threats or tactics).

Design adequacy, for example, is the probability that the system will perform its mission, conditioned on the fact that it is operating within design specifications. It is intended, in its original definition, to indicate the degradation of capability that may exist when a system is called upon to perform outside of its design performance envelope or design environments.

MIL-STD-721 (Ref. 11) has greatly contributed to the unification of the three system effectiveness concepts discussed in pars. 2-2.1, 2-2.2, and 2-2.3, and to the clarification of the definitions of the system characteristic terms used in these effectiveness concepts.

In this Military Standard ***system effectiveness*** is defined as "A measure of the degree to which an item can be expected to achieve a set of specific mission requirements, and which may be expressed as a function of availability, dependability, and capability".

In turn, ***availability*** is defined as "A measure of the degree to which an item is in the operable and committable state at the start of the mission, when the mission is called for at an unknown (random) point in time." ***Dependability*** is defined as "A measure of the item operating condition at one or more points during the mission, including the effects of reliability and maintainability, given the item condition(s) at the start of the mission.'' It may be stated as the probability that an item will (a) enter or occupy any one of its required operational modes during a specified mission, and (b) perform the functions associated with those operational modes. ***Capability*** is defined as "A measure of the ability of an item to achieve mission objectives given the conditions during the mission".

This system effectiveness definition comes conceptually closest to the WSEIAC concept, though it does not necessarily adopt the WSEIAC mathematical model. However, it states clearly that system effectiveness has to be viewed as a function of availability, dependability, and capability. This concept of system effectiveness has been adopted also by AMCR **11-1** (Ref. **12).**

2-2.6 TOTAL PACKAGE PLANNING AND SYSTEM EFFECTIVENESS

The key to system effectiveness lies in total package planning, sometimes called ***system engineering–the*** application of the system approach to total system design (Ref. **13).** Total package planning is concerned with the system life cycle (par. **3-2.1.1)** and the integration of all system elements into an effective whole, as depicted in Fig. **1-1.**

Maintainability requirements, as part of total package planning, must be derived from system mission requirements and logistic support concepts during the early life cycle phases. (See par. **3-2.1.1** and Fig. **3-8.)** These, in turn, dictate the operational readiness portions of system effectiveness described in this chapter in terms of system availability or dependability, and their related reliability, maintainability, and integrated logistic support requirements (see also par. 1-2).

Maintainability considerations, including both corrective and preventive maintenance, are primarily dependent upon the system mission profiles and system effectiveness requirements for each mission and mission mode. These, in turn, are derived from an analysis of the various operational states in which the system may be at any given time. Par. **2-2.6.1** discusses system operational states.

2-2.6.1 Operational States

The ARINC concept of operational readiness and the WSEIAC concept of system effectiveness specifically mention the states of the system as parameters of interest. In the case of operational readiness, the states are concerned with the various time periods into which the system operational demands may be classified. A discussion of the operational state considerations in terms of time periods is, therefore, in order.

A system may be considered to be in one of three operational states that may be defined as:

1. ***Inactive period*** is that period of time when the system is not required for use and is essentially shut down. It is possible for maintenance to be performed during this period.

2. ***Scheduled downtime*** is that period of time when preventive maintenance is performed. It is possible for deferred corrective maintenance to be performed during this period also.

3. ***Operational demand*** is that period of time during which the system must be available for performing operational missions. It is critical to system effectiveness. Operational demand may be partitioned into standby, alert, reaction, mission, and deactivation time periods.

Ideally, a system should be able to start performing its mission immediately upon receipt of the command to do ***so*** and to return to its designated nonmission state similarly upon command as shown in Fig. **2-3. As** a practical matter, there always exists some transient period of time before the system is fully activated (performing its mission at or above threshold effectiveness level) or deactivated as shown in Fig. **2-4.**

The definitions of these partitions of operational demand time are:

1. ***Standby*** is that fraction of operational demand during which a system is available for a mission, but requires relatively minor action to be performed before a mission can be initiated.

2. ***Alert time*** (as defined in MIL-STD-721) is that element of uptime during which an item is thought to be in specified operating condition and is awaiting a command to perform its intended mission.

3. ***Reaction time*** (as defined in MIL-STD-721) is that element of uptime needed to initiate a mission, measured from the time the command is **received.** It is the transient time between nonmission and mission states of the system. Since the command to initiate a

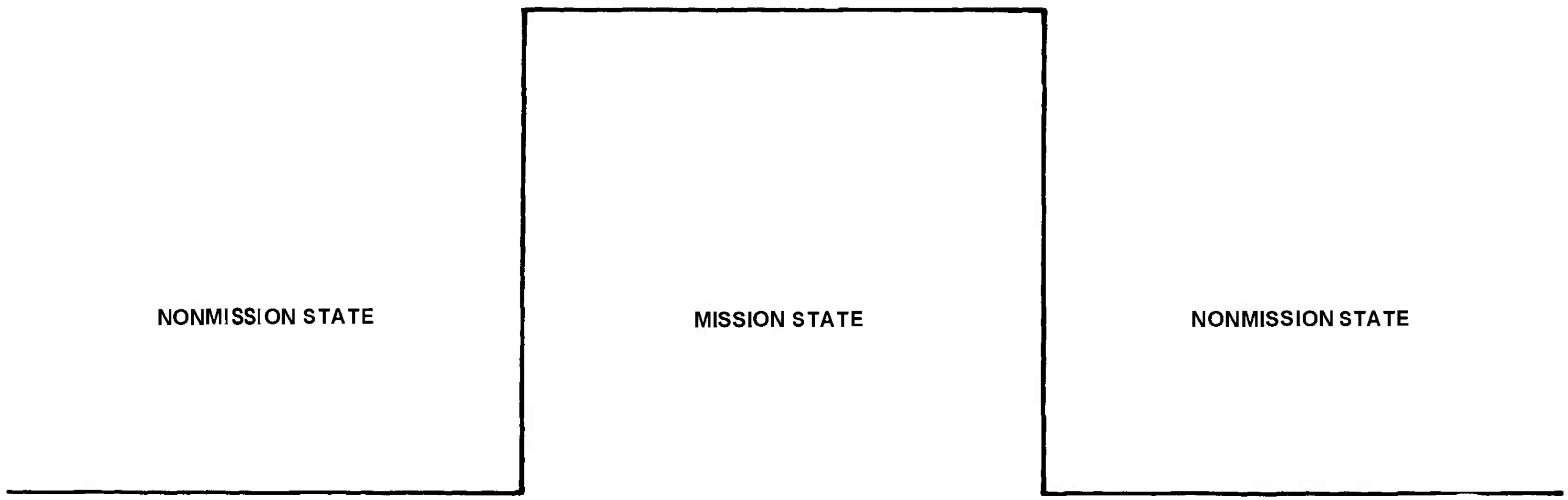


Figure 2-3. Ideal Mission State Profile

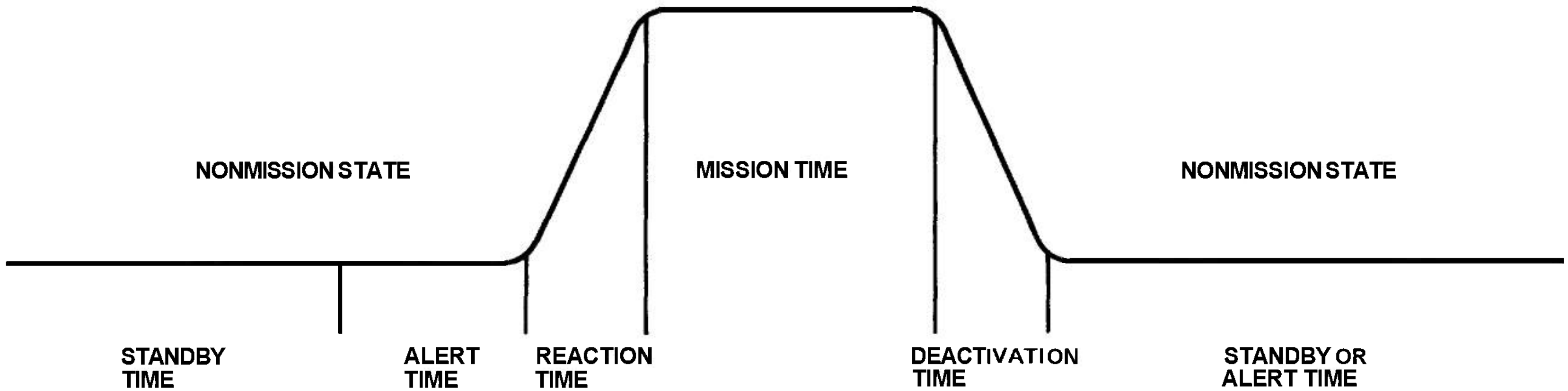


Figure 2-4. Real Life Mission Profile

mission may be given any time during operational demand, reaction time depends on the period the system is in when such command is given.

4. ***Deactivation time*** is that fraction of operational demand time required to shut down the system and return it to standby or alert portions of operational demand or to inactive or scheduled downtime periods.

5. ***Mission time*** is that element of uptime during which the item is performing its designated mission (MIL-STD-721).

6. ***Unscheduled downtime*** is the fraction of time during which the system is known to possess faults or is undergoing corrective maintenance.

7. ***Uptime*** is the fraction of operational demand during which the system is not undergoing maintenance.

In many cases, the reaction time and deactivation time may be so small as compared with the other time periods as to be considered negligible. In other cases, especially in the case of reaction time, this time period might be of prime importance to successful mission accomplishment. Generally, deactivation time is not of significance since it occurs at the end of a mission period when the system is no longer required to be operating.

Reaction time, on the other hand, may be variable depending upon whether the command for system operation occurs when the system is in an inactive state, in scheduled downtime, in standby, or in alert periods. These periods represent varying degrees of readiness for a mission. For example, during inactive time, dust covers may be on, power is off, and operating personnel may be unavailable to operate the system; during scheduled downtime, the system may be partially disassembled for servicing; during standby, dust covers are off, standby power is on, but safety switches are in "safe" position, and operating personnel, although available, are not at duty station; during alert, power is on, safety switches are in armed or go position, personnel are at their duty stations, and only the signal to start the mission is required to initiate the transient between mission and nonmission status. At the end of a mission, the system can be returned to any of these states. An illustration of the application of these is given in Ref. 14.

2-2.6.2 Effect of Logistic Support on System Effectiveness

In par. 2-2.4.2, several availability measures are described. Of these, ***inherent availability*** includes only those items of the system design which are normally design controllable; ***achieved availability*** assumes an ideal support environment. Only ***operational availability*** takes into account administrative and logistic support effects upon system effectiveness. It is also the effectiveness measure that is most difficult to demonstrate.

The lack of repair parts, spares, consumables, and proper tools; inadequate test and support equipment, maintenance facilities, and maintenance and supply information; and deficiencies in trained operating and maintenance personnel can easily negate the best reliability and maintainability design effort. Thus, while inherent and achieved availability goals may be met, actual operational availability can be readily compromised as a result of poor logistic support planning and implementation.

Experience with a number of fielded systems has shown that significant improvements in operational capability can be achieved without a system hardware redesign effort if proper attention is given to the logistic support factors during the early system planning phases. It is for these reasons, as well as the significant cost of logistic support, that such great emphasis has been given in recent years to ***integrated logistic support*** as an element of system design and effectiveness and to reduce life cycle cost, as detailed in such documents as AMCR 750-15 (Ref. 15) and TM 38-703-1 (Ref. 1).

SECTION II

SYSTEM EFFECTIVENESS METHODS

2-3 SYSTEM EFFECTIVENESS MEASURES

Every system is designed to accomplish some designated objectives, such as a specified function(s) or mission@). "System Effectiveness" is the name given to any measure which quantitatively describes how well the system will do its intended job. This ability of a system depends to a large degree on its performance or capability, but other factors such as reliability and maintainability must also be considered. A system with even the best performance designed in will not do itsjob well if it lacks reliability. If it lacks maintainability, it will require excessivedowntimes for maintenance when failures occur and will not'beoperationally ready every time when it is needed. Jointly, reliability and maintainability determine system operational readiness or availability.

To predict system effectivenessquantitatively, mathematical models that combine system performance, reliability, and maintainability into one or more effectiveness measures are used. Before the mathematical models are formulated, the mission and mission profiles must be defined and appropriate measures of system effectiveness must be selected. These measures should be both system and mission oriented so that when the numerical answers are obtained through the exercise of the mathematical models, these answers will quantitatively relate the expected response of the system with regard to the requirements and objectives of the mission.

The selection or development of appropriate effectiveness measures is not always easy. In some cases, such as communication systems, one must be content with using submodels yielding different kinds of measures that cannot be combined into a single overall effectiveness measure or a single figure of merit (Ref. 16, p. 2-24).

F. H. Kranz (Ref. 17, p. 11) states that there are three different viewpoints concerning the problem of system effectiveness quantification: "The first viewpoint is to quantify everything and to consider everything quantifiable into a figure of merit. The result is a numerical decision aid that usually has some undesirable attributes such as oversimplification, nonsensitivity to critical parameters, hidden calculations, and difficulty in exercising the model. This technique is characterized by mathematical models, computer programs, and attempted optimizations". The second viewpoint is to consider effectiveness as specified and concentrate on cost reduction, which has the danger of formulating all technicalproblems in terms of cost. The third viewpoint distrusts the "numbers game" and sticks to management actions that have, in the past, yielded cost-effective products. The result may be a well-run project yielding a product less than satisfactory for mission success. In the further text of his work, Kranz advocates a blending of all three viewpoints and suggests how to achieve this so as to assure management decisions resulting in a product with high probability of mission accomplishment, a program with minimum risks, and product and program costs within acceptable values of resources expended.

This appears to be a very reasonable approach to achieving system effectiveness.The first viewpoint is to quantify everything that is quantifiable in terms of system effectivenessmeasures and to use these numbers as inputs into system effectiveness models. However, since not everything is quantifiable, and because constraints on costs and schedules always exist, the advantageous parts of the second and third viewpoints must be integrated into the overall approach to system effectiveness. Thus, there is a need for mathematical models that are compatible with the selection effectiveness measures, and are used to allocate and define design criteria; a need to control the program so that the established design criteria are met with minimum cost; and a need to apply management methods that have proven successful on previous programs carried out without overruns and slippages.

2-3.1 TYPICAL EFFECTIVENESS MEASURES

The philosophy of choosing an appropriate measure of system effectiveness related to the system mission can be illustrated by considering an essentially continuously operating commercial system, such as a passenger airliner (Ref. 18). Its effectiveness can be measured in terms of expected seat miles flown per annum, considering delays, aborted flights, navigational deviations, emergency landings, turnaround times, etc. This number could be compared with an "ideal" airplane of the same type with normal, scheduled turnaround times and with no delays due to failures, unscheduled maintenance, or degraded operation.

Going one step further, one could look at passenger miles flown per annum, as against seat miles. This reflects utilization and revenue, and thus cost-effectiveness, when the expected annual maintenance and operating costs are added to the prorated acquisition costs.

The same or similar considerations in choosing effectiveness measures can apply to many other systems. Some typical measures are:

1. Expected number of ton-miles transported and delivered per unit time
2. Expected number of miles travelled per unit time
3. Expected number of bits processed per unit time
4. Expected number of message units transmitted per unit time
5. Expected number of kilowatt hours produced per unit time

The "unit time" may be any time measure appropriate for the specific system operation, such as mission time, hour,, day, battlefield day, month, or year.

As to weapon systems, typical system effectiveness measures are:

1. Expected number of targets destroyed per system per mission
2. Expected area destroyed per system per mission
3. Expedted area reconnoitered per system per mission
4. Expected amount of damage inflicted per system per mission
5. Rate of area destruction
6. Rate of payload delivery
7. Kill rate
8. Sweep rate
9. Single-shot kill probability
10. Probability of mission success.

There are many other system effectiveness measures that are appropriate for specific weapon systems and specific types of tactical and strategic situations. In general, three types of system effectiveness measures are most frequently used:

1. Probabilities
2. Expected values
3. Rates

Probability measures are used when a mission can be exactly defined with a unique and definite objective for its outcome. Such is the case of a ballistic missile aimed at a definite target or a bomber aircraft sent to destroy a specific target.

Expected values are more appropriate for missions with more general objectives and no specific single objective defined or even definable. For example, a reconnaissance mission may have the objective of surveying some enemy-held territory. In this case it is more meaningful to select as an appropriate measure the expected area surveyed, rather than a probability of mission success. This is not a case where the mission outcome can be only one of two possible occurrences—success or failure. In such missions one is more concerned with degrees of success, such as how much of the area gets surveyed or destroyed, or how much damage is done to the enemy.

Rates are appropriate measures of system effectiveness for weapons required to continuously repeat one and the same defined action. This would be the case of a gun, or battery of guns, required to fire into a specific area to prevent enemy penetration or infiltration while no specific targets are identified. Of course, the terms "expected values" and "rates" are often synonymous.

Whichever measure is properly chosen, the mathematical model for system effectiveness must then be geared to that measure and be capable of giving quantitative, i.e., numerical, answers in terms of the chosen measure. As already stated, performance, reliability, and maintainability are important inputs for such a model. For military systems, additional factors which will enter the picture are vulnerability, survivability, penetrability, lethality, countermeasures, enemy capability, and many other factors according to the nature of the system and the job it is intended to perform. For instance, to determine the effectiveness of a fighter aircraft, the capabilities and characteristics of the enemy aircraft it will meet in air-to-air combat must be included in the system effectiveness model. Ref. 19 is a very useful classical text for the readers who want to gain more insight into the complexities of evaluating the ***performance*** effectiveness of various weapon systems.

2-3.2 ASSOCIATED MEASURES

Reliability, maintainability, and performance have their own measures that need to be considered in system effectiveness modeling, and that provide numerical inputs into the system effectiveness models. First, let us mention some performance measures that in most cases will vary greatly with the nature of the system to be analyzed and with the mission or missions it is intended to perform. In the case of reliability and maintainability the situation is much simpler since their characteristic measures, such as MTBF and *MTTR*, are always the same.

2-3.2.1 Performance Measures

Examples of a few typical performance measures are speed, range without refuelling, and load-carrying capacity for vehicles and transports in general; caliber, range, accuracy, and firing rate for guns and similar weapons; accuracy, speed of processing, and storage capability for computers; range, signal-to-noise ratio, and bandwidth for communication systems. Obviously, each category or class of systems will have its own characteristic performance measures. When the numerical values of all important performance measures of a system are known, it is possible to calculate how well a system can perform its job or the missions for which it was designed, if all essential system elements perform satisfactorily. In the past, system effectiveness evaluations were based on performance measures only, without considering reliability or maintainability. This gave an overly optimistic picture. The hardware produced was sometimes unreliable and difficult to maintain. In the new concept of system effectiveness, reliability and maintainability play a very important role.

2-3.2.2 Modes of Operation

Reliability determines the capability of a system to sustain its performance over specific periods of time of interest, such as the mission time. If the system can operate in an alternate or degraded mode when partial failures occur, again reliability determines in which state, or operating mode, the system can be expected to operate during various phases of a mission. If some maintenance can be performed during the mission, the maintainability designed into the system will co-determine in which state the system will operate. That is, when during the mission the system transits into a degraded mode due to a failure, and the failure can be fixed—say, by replacing a failed module with a good available spare—then the system will transit again into its full capability mode of operation. And when the system requires higher level maintenance due to incapacitating failures that cannot be fixed by the operator, maintainability determines how long it will take to make the system operational again. How well the logistic support has been planned and executed becomes important at this point.

2-3.2.3 Reliability Measures

Typical reliability measures are mean-time-between-failures (*MTBF*), failure rate, renewal rate, and the probability of no failure in a given time interval. In the simplest case, when the system is of a serial configuration and all components exhibit an exponential distribution of time-to-failure, the reliability $R(t)$ of such system is given by

$$R(t) = \exp(-\lambda t) \tag{2-10}$$

where λ is the sum of the failure rates of all system components and t is the mission time. The MTBF of the system is then given by

$$MTBF = 1/\lambda \tag{2-11}$$

This applies strictly to the exponential case only. In the more general case, the MTBF is given by

$$MTBF = \int_{t=0}^{\infty} R(t)\,dt \tag{2-12}$$

which applies also to systems containing various forms of redundancy and also to nonexponentially failing components. For instance, many mechanical and related components are known to have Weibull or gamma distributions of time-to-failure.

2-3.2.4 Maintainability Measures

Typical maintainability measures are the mean-time-to-repair (*MTTR*), the mean corrective maintenance time ($\bar{M}_c$), the mean preventive maintenance time ($\bar{M}_p$), the mean corrective and preventive maintenance time ($\bar{M}$), the equipment repair time (*ERT*), the geometric mean-time-to-repair ($MTTR_G$), the maximum maintenance time (M_{MAX}), and the probability of accomplishing repair in a given time interval t, called the maintainability function $M(t)$. These maintainability measures are discussed in detail in Chapter 1, pars. 1-6 and 1-7.

2-3.2.5 Failure Rate Concepts

A measure common to both reliability and maintainability is the "failure rate". Unfortunately, this term is used in different contexts by different people, and this can easily lead to confusion (Ref. 16, p. 4-15).

In reliability the term failure rate $\lambda(t)$ is defined as

$$\lambda(t) = f(t)/R(t) \tag{2-13}$$

where $f(t)$ is the probability density function (pdf) of time-to-failure, and $R(t)$ is the reliability function, i.e., the probability of no failure in time t. This failure rate $\lambda(t)$ may be constant, decreasing, or increasing with the age of a component. It is a measure of the instantaneous hazard of one and the same component to fail as it ages with operating time t. It is also referred to as the hazard rate or instantaneous failure rate. In the case of exponentially failing components, the failure rate is constant at a given stress level. This is often the case with electronic components. However, many nohelectronic components exhibit a nonconstant, mostly increasing, failure rate. The reliability Eq. 2-10 applies only to the case when the failure rate is constant. When a component has a nonconstant failure rate $\lambda(t)$, the reliability equation becomes

$$R(T, t) = \exp\left[-\int_{x=T}^{T+t} \lambda(x)\,dx\right] \tag{2-14}$$

where t is the mission time and T is the operating age of the component at the start of the mission. Obviously, in this case the *MTBF* cannot be expressed as a reciprocal of $\lambda(t)$, which is a variable, but is computed by means of Eq. 2-12. Chapters 4 and 6, Ref. 20, discuss the failure rate concept in detail.

In maintainability the concept of failure rate is used to compute the frequencies at which components fail and must be replaced so as to restore a system to its full operational capability. However, in this concept of failure rate one is not concerned with the variable $\lambda(t)$ as defined by Eq. 2-13, since the maintainability engineer is usually not interested in the instantaneous hazard of a component to fail as it ages, but rather is interested in the successive rate at which a component must be replaced by a good item (renewed) as it fails, and again replaced upon second failure, etc. He calls this the "failure rate", and uses this concept to evaluate the *frequency* at which a component will have to be replaced in a succession of failures as they occur over a long period of time. This "failure rate" appears in the literature on mathematical statistics and stochastic processes as the renewal density $h(t)$ or renewal rate (Ref. 21), and is the rate at which a component is replaced.

Fig. 2-5 shows the renewal rate $h(t)$ and the hazard rate $\lambda(t)$ of a component which has a normal distribution of time-to-failure.

As stated before, the hazard rate $\lambda(t)$ is a measure of the hazard of one component to fail as it accumulates operating hours and therefore ages. As may be seen in Fig. 2-5, in the case of the normal distribution, $\lambda(t)$ climbs indefinitely, which means that the older the component is, the greater is the hazard that it will fail. The renewal rate $h(t)$ is a measure of the *frequency* at which the component fails and must be replaced (renewed). In other words it is the replacement rate of the component. It may greatly fluctuate initially. But as replacements repeatedly occur, $h(t)$ stabilizes with time to a steady-state value of

$$h = 1/MTBF \tag{2-15}$$

It is this steady-state constant value of the renewal rate which in maintainability is called the failure rate λ and which is used to compute component failure frequencies, and to compute equipment mean-time-to-repair (*MTTR*) and other maintainability factors. Even though the steady-state value h of the renewal rate is numerically equal to the constant λ of Eq. 2-11, the two concepts are quite different, as previously explained.

2-4 SYSTEM EFFECTIVENESS MODELS

2-4.1 THE PURPOSE OF MODELS

To design and produce effective military systems that will achieve the intended objectives, an integrated approach to system effectiveness methodology and management has become mandatory. The effectiveness of a system starts to shape up early in the conceptual and design phases. There are many elements that jointly impact on system effectiveness in the design and development phase, and additional elements have an impact on the operational effectiveness of field deployed systems. All these elements must be considered to produce systems of balanced design that will perform effectively in their operational environment and can be properly supported from a logistic support viewpoint in order to maintain their effective performance in the field.

Some of the basic elements of system effectiveness have been discussed in the preceding paragraphs. The performance, reliability, and maintainability designed into a system determine what the system is inherently

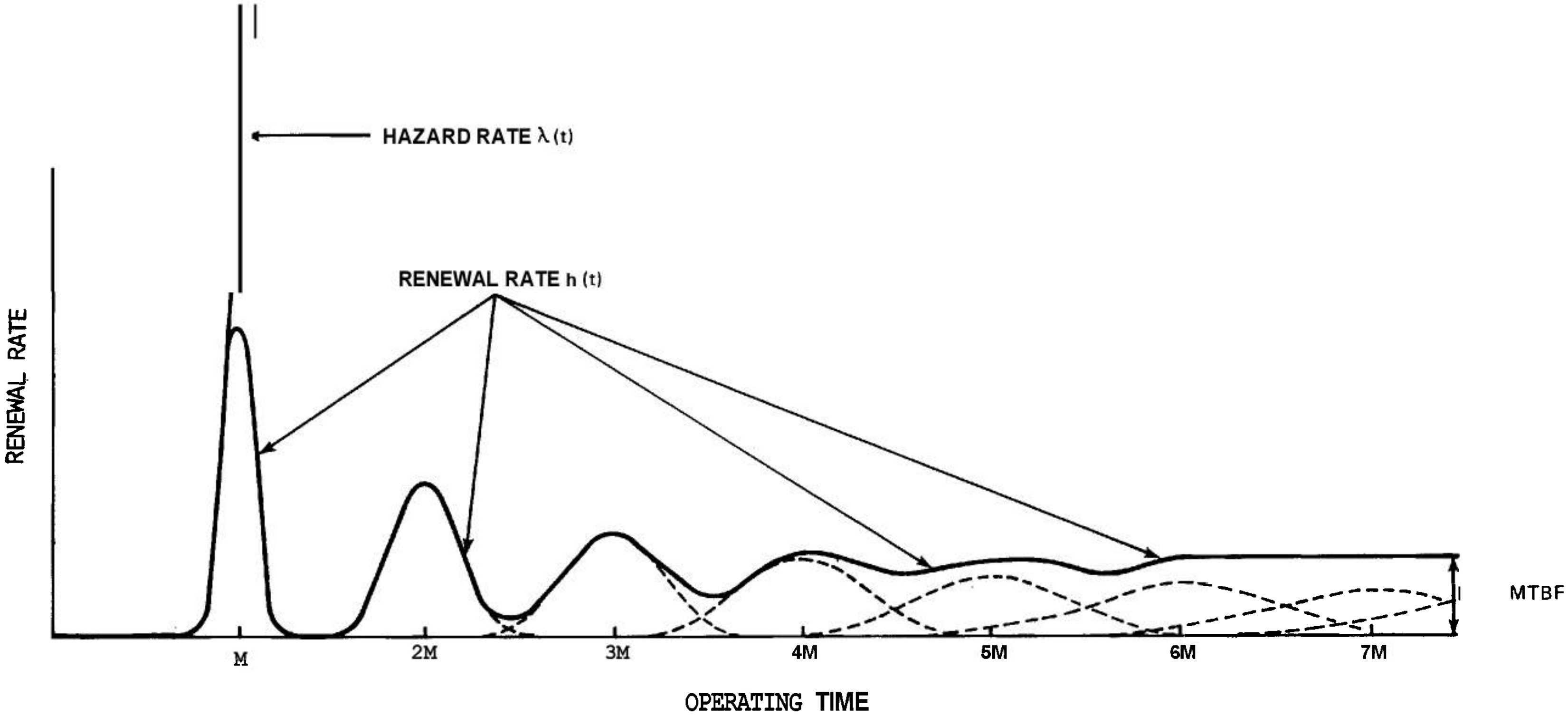


Figure 2-5. Renewal Rate and Hazard Rate

capable of doing. When, via mathematical models, one compares this inherent ability of a system to perform the specific operational tasks or mission requirements that it is intended to accomplish in a specified environment, the effectiveness of the system comes to light. One may call this the inherent system effectiveness. Also, when one considers the logistic support aspects and changing operational environment that often have a tendency to degrade operational effectiveness in the field, one may speak of an *operational* system effectiveness. This is quite similar to the concept of inherent and operational availability.

C. B. Moore (Ref. 22, p. 3) defines for an aircraft the aspects a system effectiveness analyst must take into consideration in an integrated approach to operational effectiveness analysis, namely:

1. The military objectives, requirements, and operational employment

2. The vehicle characteristics and associated set of subsystems required to accomplish the operational tasks

3. The operational environment (missions, defenses, penetration, survival, basing, tactics, etc.)

4. The men, training, and material required for support (personnel, skills, spares, equipment, safety, etc.).

He further specifically cites the elements which impact on system operational effectiveness: reliability, maintainability, performance, survivability, environment, safety, weapon delivery, operational postures, operational usage, logistics, personnel, repair parts, training, time, and cost. The cost element, of course, plays a big role in cost-effectiveness studies and in making decisions. Given all these elements involved in determining the operational system effectiveness of a complex system, it is easy to see that no single mathematical model will handle all the elements simultaneously, but rather several models are needed to perform the computations, trade-offs, sensitivity analyses, and optimizations.

Further complications arise when a system, such as a fighter-bomber, is designed to perform several different kinds of missions or mission mixes. For each kind of mission, different mathematical models may be needed and the system effectiveness will usually be numerically different for different missions. Also, the selected measure of system effectiveness may change according to mission objectives. For example, an aircraft will have different measures of effectiveness in attacking targets of different sizes and kinds.

The duration of different missions also affects system effectiveness. In longer flights reduced reliability, larger navigational errors, and possibly more exposure to enemy action lessen the probability of mission success. But even much simpler systems, such as a howitzer, will have different effectiveness against targets of different hardness and size, and at different distances—the greater the range, the greater the projectile dispersion. .

With the realization that everything must be considered and is involved in system effectiveness modeling and evaluation, it becomes obvious that system effectiveness is not the concern of a single activity, such as the reliability or maintainability organization, but rather is the concern of everybody associated with a specific project. Therefore, an integrated approach is a must. As stated in the WSEIAC Final Summary Report (Ref. 2, p. 1), "What was once merely considered desirable is now considered mandatory—an integrated methodology of system management using all available data both to pinpoint problem areas and to provide a numerical estimate of system effectiveness during all phases of the system life-cycle".

To provide such numerical estimate, or better stated, numerical estimates, system effectiveness mathematical models must be developed for the specific systems and missions. It is usually the task of the operations research groups and operation analysts to develop and exercise the mathematical models and to present the numerical answers. In the process of system synthesis and design optimization, the system effectiveness models serve several purposes:

1. To evaluate the effectiveness of a system of a specific proposed design to accomplish various operations (missions) for which it is designed and to calculate the effectiveness of other competing designs, so the decision maker can select that design which is most likely to meet specified requirements.

2. To perform trade-offs among system characteristics, performance, reliability, maintainability, etc. in order to achieve the most desirable balance among those which result in highest effectiveness.

3. To perform parametric sensitivity analyses in which the numerical value of each parameter is varied in turn, and to determine its effect on the numerical outputs of the model. Parameters that have little or no effect can be treated as constants and the model simplified accordingly. Parameters to which the model outputs show large sensitivity are then examined in detail, since small improvements in the highly sensitive parameters may result in substantial improvements in

system effectiveness at very acceptable cost (Ref. 13, p. 21).

4. To "flag" problem areas in the design which seriously limit the ability of the design to achieve the desired level of system effectiveness.

The evaluation of system effectiveness is an iterative process that continues through all life cycle phases of a system; i.e., concept development, validation, production, and operation. In each of these phases system effectiveness is continually being "measured" by exercising the system effectiveness models. In the early design stage, system effectiveness predictions are made for various possible system configurations. When experimental hardware is initially tested, first real-life information is obtained about performance, reliability, and maintainability characteristics and this information is fed into the models to update the original prediction and to further exercise the models in an attempt to improve the design. This continues when advanced development hardware is tested to gain assurance that the improvements in system design are effective or to learn what other improvements can still be made before the system is fully developed, type classified, and deployed for operational use. Once in operation, field data start to flow in and the models are then used to evaluate the operational effectiveness of the system as affected by the field environment, including the actual logistic support and maintenance practices provided in the field. The models again serve to disclose or "flag" problem areas needing improvement.

One may summarize the need for system effectiveness models as follows. First of all they provide insight, make an empirical approach to system design and synthesis economically feasible, and are a practical method for circumventing a variety of exterior constraints. Further, the models aid in establishing requirements, provide an assessment of the odds for successful mission completion, isolate problems to definite areas, and rank problems in their relative seriousness of impact on the mission. They also provide a rational basis for evaluation and choice of proposed system configurations and proposed solutions to discovered problems (Ref. 23).

Thus, system effectiveness models are an essential tool for the quantitative evaluation of system effectiveness and for designing effective weapon systems. Fig. 2-6 identifies eight principal tasks involved in system effectiveness evaluation (Ref. 2, p. 170). Task 1 is mission definition, Task 2 is system description, Task 3 is selection of figures of merit or, in a more general sense, the selection of appropriate system effectiveness measures, and Task 4 is the identification of accountable factors that impose boundary conditions and constraints on the analysis to be conducted. After completing these four tasks, it becomes possible to proceed with Task 5, the construction of the mathematical models. To obtain numerical answers from the models, numerical values of all parameters included in the models must be established or estimated (Task 7). To do this, good and reliable data must first be acquired from data sources, tests, etc. (Task 6). The final Task 8 exercises the models by feeding in the numerical parametric values to obtain system effectiveness estimates and perform optimizations. Fig. 2-7 (Ref. 7) illustrates in more detail the whole process of system effectiveness evaluations, beginning with the military operational requirements and leading, through the exercising of the system effectiveness model(s), to the decision-making stage.

2-4.2 MODELING TECHNIQUES

As discussed in par. 2-4.1, system effectiveness models integrate a number of system characteristics with the mission objectives, the mission profiles and environments, and the logistic support. The main characteristics of the system are, in the broadest sense, its performance, reliability, and maintainability. They jointly determine system capability.

Reliability and maintainability define system availability and/or operational readiness. Reliability determines the state probabilities of the system during the mission, i.e., the system dependability. If repairs can be performed during the mission, maintainability also becomes a factor in dependability evaluations; this case is often referred to as "reliability with repair". Then, there is the impact of logistic support on the downtime and turnaround time of the system since shortcomings in the logistic support may cause delays over and above the maintenance time as determined by the system maintainability design. Finally, there are the performance characteristics of the system that are affected by the state in which the system may be at any point in time during a mission, i.e., by the system dependability.

Before system effectiveness models can be constructed, a great deal of submodeling must be done. Availability, operational readiness, downtime distributions, dependability, etc., require in most cases their own modeling to obtain the numerical answers that may be fed into an overall system effectiveness model, if such can be constructed. Some of these submodeling techniques will now be discussed.

2-4.2.1 Operational Readiness Models

Availability, being defined as the uptime ratio, is not always a sufficient measure to describe the ability of a

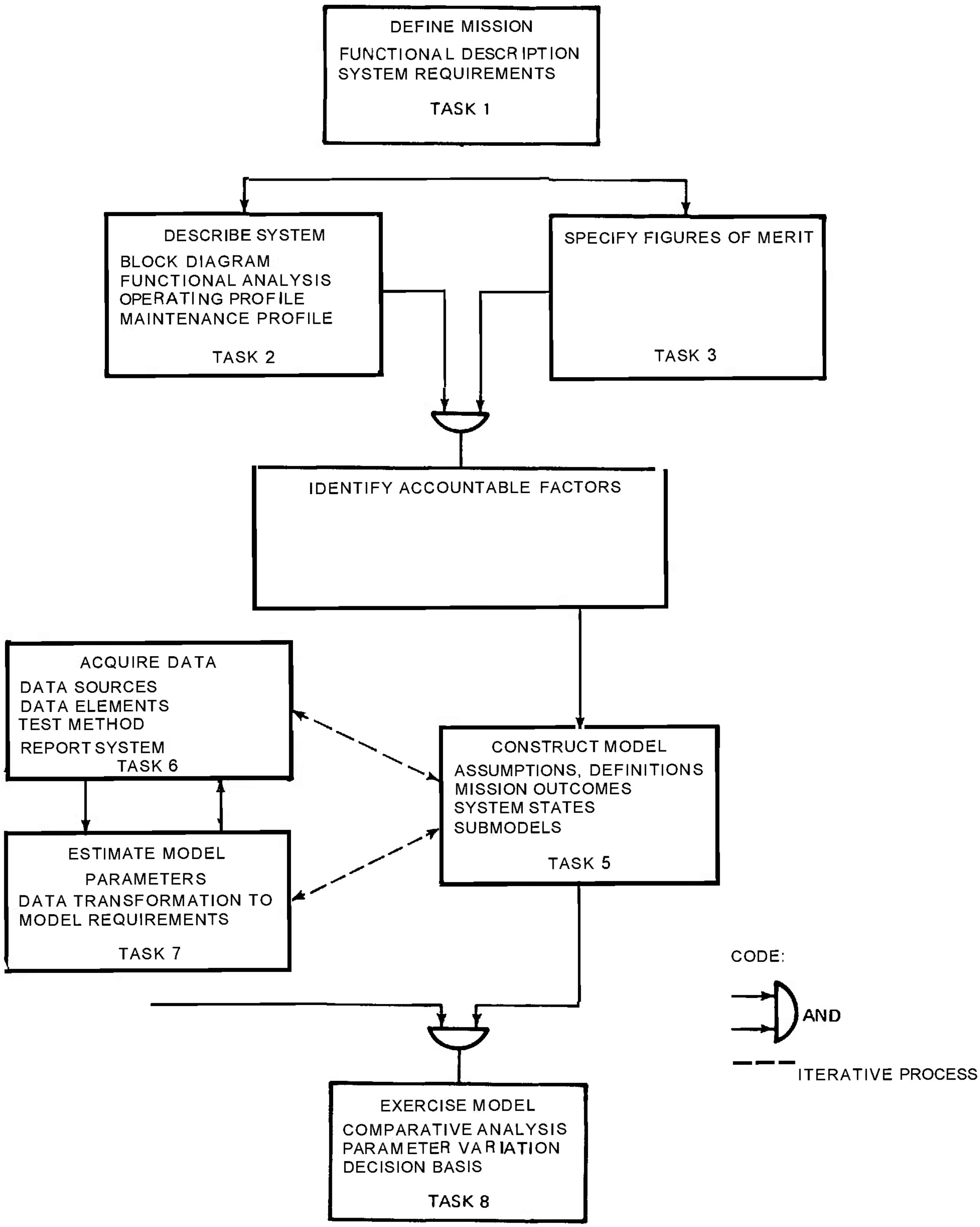
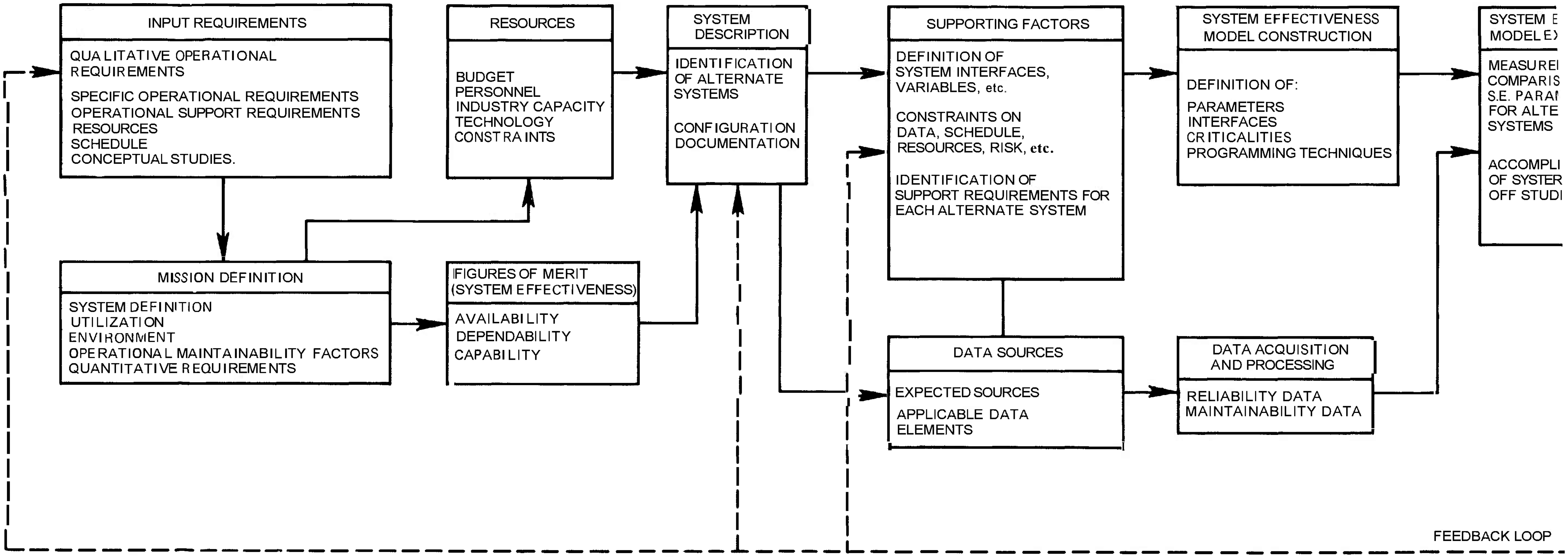


Figure 2-6. The Principal Tasks in System Effectiveness Evaluation


INPUT REQUIREMENTS
QUALITATIVE OPERATIONAL REQUIREMENTS
SPECIFIC OPERATIONAL REQUIREMENTS
OPERATIONAL SUPPORT REQUIREMENTS
RESOURCES
SCHEDULE
CONCEPTUAL STUDIES.
RESOURCES
BUDGET
PERSONNEL
INDUSTRY CAPACITY
TECHNOLOGY
CONSTRAINTS
SYSTEM DESCRIPTION
IDENTIFICATION OF ALTERNATE SYSTEMS
CONFIGURATION DOCUMENTATION
SUPPORTING FACTORS
DEFINITION OF SYSTEM INTERFACES, VARIABLES, etc.
CONSTRAINTS ON DATA, SCHEDULE, RESOURCES, RISK, etc.
IDENTIFICATION OF SUPPORT REQUIREMENTS FOR EACH ALTERNATE SYSTEM
SYSTEM EFFECTIVENESS MODEL CONSTRUCTION
DEFINITION OF:
PARAMETERS
INTERFACES
CRITICALITIES
PROGRAMMING TECHNIQUES
SYSTEM E
MODEL E
MEASURE
COMPARIS
S.E. PARA
FOR ALTE
SYSTEMS
ACCOMPLI
OF SYSTE
OFF STUD
MISSION DEFINITION
SYSTEM DEFINITION
UTILIZATION
ENVIRONMENT
OPERATIONAL MAINTAINABILITY FACTORS
QUANTITATIVE REQUIREMENTS
FIGURES OF MERIT (SYSTEM EFFECTIVENESS)
AVAILABILITY
DEPENDABILITY
CAPABILITY
DATA SOURCES
EXPECTED SOURCES
APPLICABLE DATA ELEMENTS
DATA ACQUISITION AND PROCESSING
RELIABILITY DATA
MAINTAINABILITY DATA
FEEDBACK LOOP

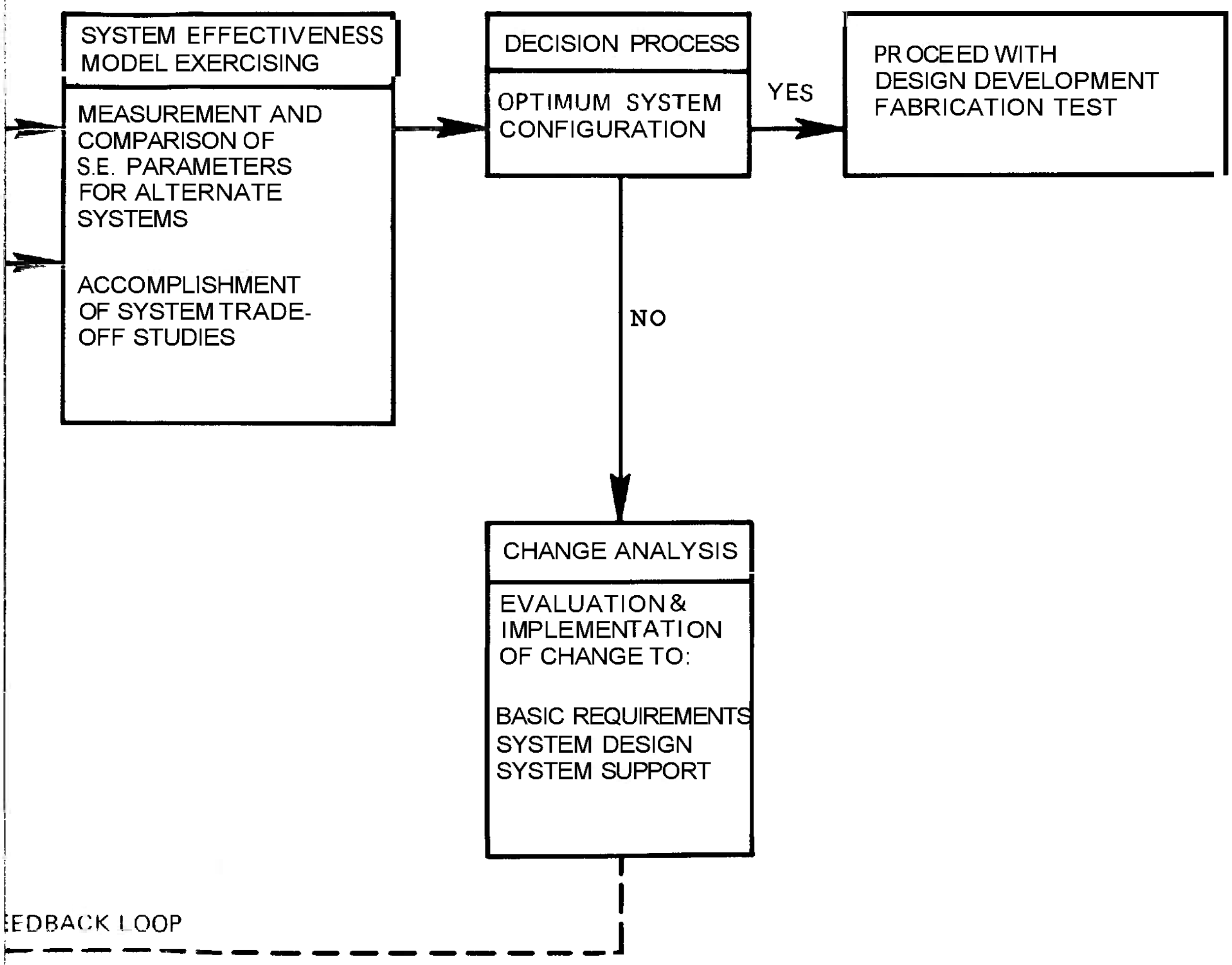


Figure 2-7. The Process of System Effectiveness Evaluation

system to be committed to a mission at any arbitrary time. In many practical military operations the concept of operational readiness serves this purpose better. We here define operational readiness as the probability that a system is in an operable condition, i.e., ready to be committed to perform a mission when demands for its use arise. The difference as well as the similarity between availability and operational readiness will become clear by comparing the models developed subsequently with the availability models discussed in the preceding paragraph.

In the development of operational readiness models one has to consider the usage and the maintenance of the system; i.e., its operating, idle, and repair times. When a call arrives for the system to engage in a mission, the system, at such time, may be in a state of perfect repair and ready to operate immediately. But it may also be in need of maintenance and not ready. Its state, when called upon to operate, depends on the preceding usage of the system—i.e., on its preceding mission, in what condition it returned from that mission, and how much time elapsed since it completed the last mission. Many models can be developed for specific cases, and some are discussed in the following paragraphs.

1. Model A

In this model the assumption is made that if no failures needing repair occurred in the preceding mission, the system is immediately ready to be used again; and, if such failures did occur, the system will be ready for the next mission only if its maintenance time is shorter than the time by which the demand for its use arises. The operational readiness P_{OR} may then be expressed as

$$P_{OR} = R(t) + Q(t) \times P(t_m < t_d) \qquad (2\text{-}16)$$

where

$R(t)$ = probability of no failures in the preceding mission

$Q(t)$ = probability of one or more failures in the preceding mission

t = mission duration

$P(t_m < t_d)$ = probability that if failures occur the system maintenance time t_m is shorter than the time t_d at which the next demand or call for mission engagement arrives.

The calculations of $R(t)$ and $Q(t) = 1 - R(t)$ are comparatively simple using standard reliability equations; however, all possible types of failures that need fixing upon return, in order to restore in full the system reliability and combat capability, must be considered, including any failures in redundant configurations.

As for $P(t_m < t_d)$, one needs to know the probability distributions of the system maintenance time and of call arrivals. Denoting by $f(t_m)$ the probability density function of maintenance time, and by $g(t_d)$ the probability density function of time to the arrival of the next call, counted from the instant the system returned from the preceding mission in a state requiring repair, the probability that the system will be restored to its full opertional capability before the next call arrives is

$$P(t_m < t_d) = \int_{t_m=0}^{\infty} f(t_m)\left[\int_{t_d=t_m}^{\infty} g(t_d)\,dt_d\right]dt_m \qquad (2\text{-}17)$$

The integral in the square brackets on the right side of this equation is the probability that the call arrives at t_d *after* a variable time t_m. When this is multiplied by the density $f(t_m)$ of the duration of maintenance times and integrated over all possible values of t_m, we get $P(t_m < t_d)$.

Now assume that maintenance time t_m and time to next call arrival t_d are exponentially distributed, with M_1 being the mean time to maintain the system and M_2 the mean time to next call arrival. The probability density functions are thus

$$f(t_m) = [\exp(-t_m/M_1)]/M_1 \qquad (2\text{-}18)$$

$$f(t_d) = [\exp(-t_d/M_2)]/M_2 \qquad (2\text{-}19)$$

We then obtain

$$P(t_m < t_d) = \int_0^{\infty} M_1^{-1}\exp(-t_m/M_1) \times\left[\int_{t_m}^{\infty}\frac{1}{M_2}\cdot\exp(-t_d/M_2)\,dt_d\right]dt_m$$

$$= \int_0^{\infty} M_1^{-1}\exp\left[-(1/M_1 + 1/M_2)t_m\right]dt_m$$

$$= M_2/(M_1 + M_2) \qquad (2\text{-}20)$$

In this *exponential* case, system operational readiness becomes

$$P_{OR} = R(t) + Q(t)[M_2/(M_1 + M_2)] \tag{2-21}$$

As a numerical example let us look at a system with a probability of $R = 0.8$ of returning from a mission of say $t = 1$-hr duration without requiring repair, and, therefore, had a probability of $Q = 0.2$ that it will require repair. If system mean maintenance time is $M_1 = 1$ hr and the mean time to next call arrival is $M_2 = 2$ hr, the operational readiness of the system becomes

$$P_{OR} = 0.8 + 0.2(2/3) = 0.933$$

Comparing this result with the conventional steady-state availability concept and assuming that the system has a mean maintenance time of $M_1 = 1$ hr and a mean time to failure of $M_2 = 5$ hr (roughly corresponding to the exponential case of $R = 0.8$ for a one-hour mission), we obtain a system availability of

$$A = M_2 / (M_1 + M_2) = 5/6 = 0.833$$

which is a result quite different from $P_{OR} = 0.933$.

An equation for $P(t_m < t_d)$, which yields identical results as Eq. 2-17, can be derived by convolution when introducing a new random variable, $z = t_d - t_m$, which is the difference between the time a call arrives t_d and the time when system maintenance is completed t_m. Whenever z is positive, the system is operationally ready. The density $p(z)$ of z is the joint density of the difference of two random variables t_d and t_m, given by

$$p(z) = \int_{t_d = z}^{\infty} g(t_d) f(t_d - z)\, dt_d \tag{2-22}$$

where $g(t_d)$ is the density of time to next call and $f(t_m) = f(t_d - z)$ is the density of system maintenance time since t_m can be substituted by $t_d - z$, which follows from $z = t_d - t_m$. The integration limits go from $t_d = z$ to $t_d = \infty$, since for $t_d < z$, $t_d - z$ becomes negative and $p(z) = 0$ by definition, i.e., system is not ready. The probability $P(t_m < t_d)$ is then

$$P(t_m < t_d) = \int_{z=0}^{\infty} p(z)\, dz \tag{2-23}$$

Returning to the exponential case treated in Eq. 2-20, and using Eq. 2-22, we obtain

$$\begin{aligned} p(z) &= \int_{t_d=0}^{\infty} [M_2^{-1} \exp(-t_d/M_2)] \times \{M_1^{-1} \exp[-(t_\alpha - z)/M_1]\}\, dt_d \\ &= (M_1 M_2)^{-1} \exp(z/M_1) \times \int_{t_d = x}^{\infty} \exp(-t_d/M_2) \cdot \exp(-t_d/M_1)\, dt_d \\ &= [1/(M_1 + M_2)] \exp(-z/M_2) \end{aligned} \tag{2-24}$$

The probability $P(t_m < t_d)$ is then, according to Eq. 2-23,

$$\begin{aligned} P(t_m < t_d) &= [1/(M_1 + M_2)] \int_{z=0}^{\infty} \exp(-z/M_2)\, dz \\ &= M_2/(M_1 + M_2) \end{aligned} \tag{2-25}$$

which agrees with the result of Eq. 2-20.

2. Model B

The operational readiness model of Eq. 2-16 can be extended to the case when mission duration time t is not the same for each mission but is distributed with a density $q(t)$. We then get

$$P_{OR} = \int_0^{\infty} R(t) q(t)\, dt + P(t_m < t_d) \int_0^{\infty} Q(t) q(t)\, dt \tag{2-26}$$

Since the integrals in Eq. 2-26 are fixed numbers, we may write

$$\begin{aligned} R &= \int_0^{\infty} R(t) q(t)\, dt \\ Q &= \int_0^{\infty} Q(t) q(t)\, dt \end{aligned} \tag{2-27}$$

and using the symbol P for $P(t_m < t_d)$, i.e., $P = P(t_m < t_d)$, Eq. 2-26 may be written in the form

$$P_{OR} = R + QP \tag{2-28}$$

In this equation R is the probability that the system returns without failures from the last mission; $Q = 1 - R$ is the probability that one or more failures developed in the last mission; and P is the probability that the system will be repaired before the next call arrives if it developed failures. The mission times are variable here with density $q(t)$.

3. Model *C*

A further extension of the P_{OR} model of **Eq. 2-28** is possible by considering the case when so called ***dormant*** failures may develop after the system checks out O.K. and before the next call to engage in a mission arrives. That is the period when the system is dormant, i.e., not operating, and is ***believed*** to be ready.

Denoting by R_D the probability of no failure(s) in the dormancy state, by t_d the variable time to next call arrival counted from the time of return from the preceding mission, and by P_D the probability that the system will be repaired before the next call arrives and will not fail in dormancy, operational readiness may be written as

$$P_{OR} = RR_D + QP_D \tag{2-29}$$

where R and Q are defined in the same way as in Eq. **2-27** in Model **B,** and R_D is the probability of no failures in the state of dormancy in the waiting period t_d.

In the computation of P_D we make use of **Eq. 2-23,** conditioning it by the requirement that no dormant failure occurs after t_m in the interval $z = t_d - t_m$, the probability of which is given by $R_D(z)$. With this condition we get

$$P_D = \int_0^\infty p(z)R_D(z)\,dz \tag{2-30}$$

Thus P_D is the probability that maintenance time t_m is smaller than the time t_d to next call arrival ***and*** no dormant failure occurs after completing maintenance at t_m.

In the computation of R_D, we use the fact that dormancy time t_d is the same as the time to next call arrival which has the density $g(t_d)$ of Model A, i.e.,

$$R_D = \int_0^\infty g(t_d)R_D(t_d)\,dt_d \tag{2-31}$$

which is the probability of no dormant failure(s) occurring in the variable dormancy time t_d when at $t_d = 0$ the system returned without failures.

As an exercise let us assume that all variable times are ***exponentially*** distributed with the following means:

M_1 = mean maintenance time of the system
M_2 = mean time to next call arrival
M_3 = mean time of $R(t)$, i.e., of the probability of no failure(s) occurring in mission time t
M_4 = mean mission time when mission durations are distributed with density $q(t)$
M_5 = mean time of R_D, i.e., of the probability that no dormant failures occur.

We compute first R and Q of **Eq. 2-28** as follows:

$$R = \int_{t=0}^\infty [\exp(-t/M_3)\cdot\exp(-t/M_4)/M_4]\,dt$$
$$= M_3/(M_3 + M_4) \tag{2-32}$$
$$Q = 1 - R = M_4/(M_3 + M_4)$$

Next we compute P_D, using **Eqs. 2-30** and **2-24:**

$$P_D = [1/(M_1 + M_2)]\int_0^\infty \exp(-z/M_2)\cdot\exp(-z/M_5)\,dz$$
$$= \left(\frac{M_2}{M_1 + M_2}\right)\times\left(\frac{M_5}{M_2 + M_5}\right) \tag{2-33}$$

Finally we compute R_D, using **Eq. 2-31:**

$$R_D = \int_0^{\infty} [\exp(-t/M_2) \cdot \exp(-t/M_5)/M_2]\,dt$$

$$= M_5/(M_2 + M_5) \qquad (2\text{-}34)$$

Eq. 2-29 then assumes the form:

$$P_{OR} = \left(\frac{M_3}{M_3 + M_4}\right) \times \left(\frac{M_5}{M_2 + M_5}\right) \times \left(\frac{M_4}{M_3 + M_4}\right) \times \left(\frac{M_2}{M_1 + M_2}\right) \times \left(\frac{M_5}{M_2 + M_5}\right)$$

$$= \frac{M_5}{M_2 + M_5}\left(\frac{M_3}{M_3 + M_4} + \frac{M_2}{M_1 + M_2} + \frac{M_4}{M_3 + M_4}\right) \qquad (2\text{-}35)$$

And when no dormant failures can occur, i.e., when $R_D = 1$:

$$P_{OR} = \frac{M_3}{M_3 + M_4} + \left(\frac{M_2}{M_1 + M_2}\right) \times \left(\frac{M_4}{M_3 + M_4}\right) \qquad (2\text{-}36)$$

For a numerical example assume $M_1 = 1$, $M_2 = 2$, $M_3 = 5$, $M_4 = 1$, and $M_5 = 10$. We obtain from Eq. **2-35:**

$P_{OR} = (10/12)[(5/6) + (2/3)(1/6)] = 0.787$

And from Eq. **2-36,** i.e., no dormant failures, we obtain:

$P_{OR} = 5/6 + (2/3)(1/6) = 0.833 + 0.111$
$= 0.944$

In this example we took the mean time to dormant failure M_5 to be twice that of mean time to failure when operating M_3. In reality we would expect M_5 to be at least ten times M_3, or even more. In some systems dormant failures may not occur at all.

4. Model **D**

Models A through *C* are "strict" in the sense that no allowance for turn-around time or for alert time is made. The models can be "relaxed" if a minimum turn-around time t_o is ***allowed*** for refueling, checkout, etc.; i.e., if a call arrives within t_o after the system returns from a mission, it will not count as a failure to be ready, and if an alert is given at t_d, there is an alert time of t_a ***allowed*** for pre-mission checkout and correcting anything that needs repairing before t_a expires, i.e., a second chance is given to make the system ready by $t_d + t_a$.

In the model which follows we assume that a minimum alert time of t_a is allowed ***after*** a call arrives at t_d. Operational readiness may then be expressed as:

$$P_{OR} = R[R_D + (1 + R_D)M_a] + Q[P_D + (1 + P_D)M_a] \qquad (2\text{-}37)$$

In this equation the first term on the right is the probability that the system returned without need for repair from the preceding mission *(R)* and checks **O.K.** at t_d (R_D) ***or*** does not check **O.K.** at t_d $(1 - R_D)$ but can be fixed in t_a (M_a). The second term is the probability that the system needed repair after returning from the preceding mission $(Q = 1 - R)$ and is ready at t_d (P_D) ***or*** is not ready at t_d $(1 - P_D)$ but can be fixed in t_a (M_a). The equations for R, Q, P_D, and R_D are the same as given in Models **B** and *C*, and the equation for M_a, the probability of repairing the system in t_a, is given by

$$M_a = \int_{t_m = 0}^{t_a} f(t_m)\,dt_m \qquad (2\text{-}38)$$

where $t_m = 0$ at t_d when alert is sounded, and $f(t_m)$ is the probability density of the maintenance time t_m.

5. Model E

Returning to Eq. **2-16** we "decompose" the term $Q(t)P(t_m < t_d)$ into its "constituent elements", and write

$$P_{OR} = R + \sum_j Q_j P_j \qquad (2\text{-}39)$$

This decomposition makes the model more tractable for complex systems that may include redundancies of the parallel and standby type. The system may return from the preceding mission in a variety of states requiring repair, say j states, where the jth state occurs with probability P_j in which case ***fit,*** $< t_d) = P_j$. The sum of all $(Q_j P_j)$ is then the term $QP(t_m < t_d)$ of **Eq. 2-16**

Assume a system consists of a subsystem with the ith subsystem having a probability of q_i that it will need repair and a probability of r_i that it will not need repair. The number of states j in which any one or more subsystems will require repair or replacement is then the combination of any one out of n, any two out of n, any three out of n, etc., subsystems requiring repair, i.e.,

$$j = \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n} = 2^n - 1 \qquad (2\text{-}40)$$

For instance the system consists of three subsystems a, b and c (such as black boxes or system replaceable units) and each of these can suffer some kind of failure or malfunction in a mission independently of the other subsystems, the number of states requiring repair is

$$j = \binom{3}{1} + \binom{3}{2} + \binom{3}{3} = 3 + 3 + 1 = 2^3 - 1 = 7 \qquad (2\text{-}41)$$

Each of these seven states occurs with a certain probability. The first three states are: a needs repair (b and c do not) which occurs with probability $q_a\, r_b\, r_c$; b needs repair (a and c do not) which occurs with probability $r_a\, q_b\, r_c$; and c needs repair (a and b do not) which occurs with probability $r_a\, r_b\, q_c$. The following tabulation lists all seven cases:

State Number j	Units Failed	Probability of state Q_j	P_j (ready by t_d/given Q_j)
1	a	$Q_1 = q_a r_b r_c$	P_1
2	b	$Q_2 = r_a q_b r_c$	P_2
3	c	$Q_3 = r_a r_b q_c$	P_3
4	a, b	$Q_4 = q_a q_b r_c$	P_4
5	a, c	$Q_5 = q_a r_b q_c$	P_5
6	b, c	$Q_6 = r_a q_b q_c$	P_6
7	a, b, c	$Q_7 = q_a q_b q_c$	P_7

(2-42)

As to the computation of the P_j terms, the maintenance time distributions of the three subsystems at the system maintenance level must be known, i.e., $f_a(t_m)$, $f_b(t_m)$, and $f_c(t_m)$. We get then P_1, P_2, and P_3 by substituting $f_a(t_m)$, $f_b(t_m)$, and $f_c(t_m)$, respectively, into Eq. 2-17. As to P_4, P_5, etc., there are two or more subsystems to be repaired. The distribution of system downtime in those cases depends on the maintenance policy that may be of the sequential maintenance type (i.e., one at a time) or of the parallel type (i.e., simultaneous maintenance).

In the sequential case we form the joint densities of downtime by convolutions and obtain

$$f_{ab} = f_a * f_b;\quad f_{ac} = f_a * f_c;\quad f_{bc} = f_b * f_c;$$
$$f_{abc} = f_a * f_b * f_c \qquad (2\text{-}43)$$

We use, then, these joint densities in Eq. 2-17 to get P_4 to P_7.

In the case of parallel repair, where each subsystem has its own crew to work on it, we compute the P_i's as follows. We transform Eq. 2-17 by changing the integration limits and get

$$P(t_m < t_d) = \int_{t_d=0}^{\infty} g(t_d)\left[\int_{t_m=0}^{t_d} f(t_m)\,dt_m\right]dt_d$$
$$= \int_{t_d=0}^{\infty} g(t_d)\,M(t_d)\,dt_d \qquad (2\text{-}44)$$

since $M(t_d)$, the maintainability function, is given by

$$M(t_d) = \int_{t_m=0}^{t_d} f(t_m)\,dt_m \qquad (2\text{-}45)$$

For the jth state we get then P_j as

$$P_j = \int_{t_d=0}^{t_d} g(t_d)\,M_j(t_d)\,dt_d \qquad (2\text{-}46)$$

where M_j is the maintainability function, i.e., the system downtime distribution for either one, or two or more subsystems in simultaneous repair. Thus

$$M_1 = M_a(t_d);\; M_2 = M_b(t_d)$$

and

$$M_3 = M_c(t_d)$$

for single repair of subsystems a, b, and c treated in the previous example and further, for simultaneous repairs, we get

$$M_4 = 1 - (1 - M_1)(1 - M_2)$$
$$M_5 = 1 - (1 - M_1)(1 - M_3)$$
$$M_6 = 1 - (1 - M_2)(1 - M_3) \qquad (2\text{-}47)$$
$$M_7 = 1 - (1 - M_1)(1 - M_2)(1 - M_3)$$

The terms M_4 through M_7 in Eq. 2-47 are maintainability functions, i.e., M_4 is the probability that system maintenance will be completed by t_d when subsystems a and b need repair, etc., and M_7 is the probability that system maintenance will be completed on time when all three subsystems need repair or must be replaced.

To recapitulate, in the case of sequential repair we form the joint density functions of the subsystems requiring repair, while in the case of parallel repair we form the maintainability functions M_1 through M_7 by simple probability equations of simultaneously proceeding maintenance events (i.e., in parallel).

Many other operational readiness models can be developed and the models here discussed can be further refined, for instance, by considering imperfect checkout so that some failures are not detected, etc. The purpose followed here was to introduce the reader to certain modeling techniques so he can develop his own models to suit specific weapons, missions, and maintenance policies. In par. 2-4.2.2 we show some availability modeling techniques, following the same purpose.

2-4.2.2 Availability Models

The concept of availability was originally developed for repairable systems that are required to operate continuously, i.e., round-the-clock, and are at any random point in time either operating or are "down" because of failure and are being worked upon so as to restore their operation in minimum time. In this original concept a system is considered to be in only two possible states-operating or in repair—and availability is defined as the probability that a system is operating satisfactorily at any random point in time t when subject to a sequence of "up" and "down" cycles which constitute an alternating renewal process (Ref. 21, pp. 80-86).

1. Model A

Consider first a single unit system or a strictly serial system that has a stationary reliability $R(t)$; its availability $A(t)$ that it will be in an "up" state (i.e., will be operating) at time t when it started in an "up" condition at $t = 0$, is given by the renewal equation

$$A(t) = R(t) + \int_{y=0}^{t} R(t - y)n(y)\,dy \qquad (2\text{-}48)$$

In this equation the first term on the right, $R(t)$, is the probability that the system does not fail at all up to time t, and the second term is the probability that the system, prior to t, failed one or more times, was last restored to operation at time $y < t$ and survived without failure the following period $(t - y)$ so that at time t it is operating. The function $n(t)$ is the renewal rate of the system, i.e., the rate at which the system enters the "up" state as it undergoes a series of operation-failure-repair-and "up" again cycles. This renewal rate is given by the renewal equation

$$n(t) = h(t) + \int_{y=0}^{t} h(t - y)n(y)\,dy \qquad (2\text{-}49)$$

where $h(t)$ is the joint probability density function (pdf) of the sum of two random variables, i.e., the time-to-failure and the time-to-repair. If we denote the pdf of time-to-failure (i.e., of the "up" time) by $g(t)$, and the pdf of the time-to-repair (i.e., of the "down" time) by $f(t)$, we get $h(t)$ by convoluting $g(t)$ with $f(t)$, i.e.,

$$h(t) = \int_{y=0}^{t} g(t - y)f(y)\,dy \qquad (2\text{-}50)$$

As an example assume a system with the following exponential density functions for uptime and downtime

$$g(t) = \lambda \exp(-\lambda t) \qquad (2\text{-}51)$$

$$f(t) = \mu \exp(-\mu t) \qquad (2\text{-}52)$$

where λ is the failure rate and μ is the repair rate.

To solve for $A(t)$, we take the Laplace transforms of Eqs. 2-48 through 2-52 and get

$$A^*(s) = R^*(s) + R^*(s)n^*(s) \qquad (2\text{-}53)$$

$$n^*(s) = g^*(s)f^*(s) \qquad (2\text{-}54)$$

$$h^*(s) = g^*(s)f^*(s) \qquad (2\text{-}55)$$

Substituting Eq. 2-55 into Eq. 2-54 we get

$$n^*(s) = g^*(s)f^*(s)/[1 - g^*(s)f^*(s)] \quad (2\text{-}56)$$

which in turn we substitute into **Eq. 2-53,** and realizing that $R^*(s)$ can be written in terms of the density function $g^*(s)$ as

$$R^*(s) = (1/s)[1 - g^*(s)] \quad (2\text{-}57)$$

Eq. **2-55** obtains the form

$$A^*(s) = [1 - g^*(s)]/\{s[1 - f^*(s)g^*(s)]\} \quad (2\text{-}58)$$

Since the Laplace transforms of the density functions $g(t)$ and $f(t)$ are

$$g^*(s) = \lambda/\lambda + s \quad (2\text{-}59)$$

$$f^*(s) = \mu/(\mu + s) \quad (2\text{-}60)$$

We substitute these into Eq. **2-58** and transforming into the time domain we get

$$A(t) = \mu/(\lambda + \mu) + [\lambda/(\lambda + \mu)]\exp[-(\lambda + \mu)t] \quad (2\text{-}61)$$

Fig. **2-8** shows the availability function $A(t)$ for the case when the repair rate μ is four times that of the failure rate. Stated differently, the MTBF is four times that of the MTTR; or the maintenance time ratio, defined as ***MTTR/MTBF,*** is **1:4.**

We may write Eq. **2-61** also in terms of the reciprocal values of the failure and repair rates—i.e., in terms of the MTBF and the *MTTR*—remembering, however, that both time-to-failure and time-to-repair must be exponentially distributed for the equation to hold:

$$A(t) = \frac{MTBF}{MTBF + MTTR} + \frac{MTTR}{MTBF + MTTR} \times \exp\left[-\left(\frac{1}{MTBF} + \frac{1}{MTTR}\right)t\right] \quad (2\text{-}62)$$

When we study this equation we see that as t increases the second term on the right diminishes and that availability in the limit becomes a constant, i.e.,

$$\lim_{t \to 0} A(t) = A = \frac{MTBF}{MTBF + MTTR} \quad (2\text{-}63)$$

We call this the steady-state or equilibrium availability of a serial system. It is equivalent to the intrinsic availability of Eq. **1-23,** Chapter **1.**

We may see in Fig. **2-8** that an exponentially failing and exponentially repaired system with a maintenance time ratio of **1:4** approaches the steady state rather rapidly, in a calendar time of just over one-half of the system *MTBF.* For lower maintenance time ratios the process stabilizes even more rapidly.

Looking again at Eq. **2-63,** we may divide the numerator and the denominator by the MTBF and write the steady-state availability in terms of the maintenance ratio

$$A = 1/(1 + \alpha) \quad (2\text{-}64)$$

where α = *MTTR/MTBF,* the maintenance time ratio (*MTR*). Thus the availability A does not depend on the actual values of the *MTBF* or MTTR but only on their ratio. A system with an MTBF of say **4** hr and an MTTR of 1 hr will have the same steady-state availability of 80% as a system with an MTBF of 100 hr and an MTTR of **25** hr. But from a mission accomplishment viewpoint it may make all the difference whether the system has an MTBF of **100** hr or **4** hr!

An availability of 80%, shown in Fig. **2-8,** is in most practical cases not adequate. Much higher availabilities can be achieved when properly designing for reliability and maintainability. High reliabilities are required for mission accomplishments and, with modular design for maintenance where failed items can be quickly replaced as modules, much better maintenance time ratios should be achievable.

2. Model B

What we discussed so far is the concept of the so-called *pointwise* availability which, as in Fig. 2-8, shows us the probability that a system is "up" and operating

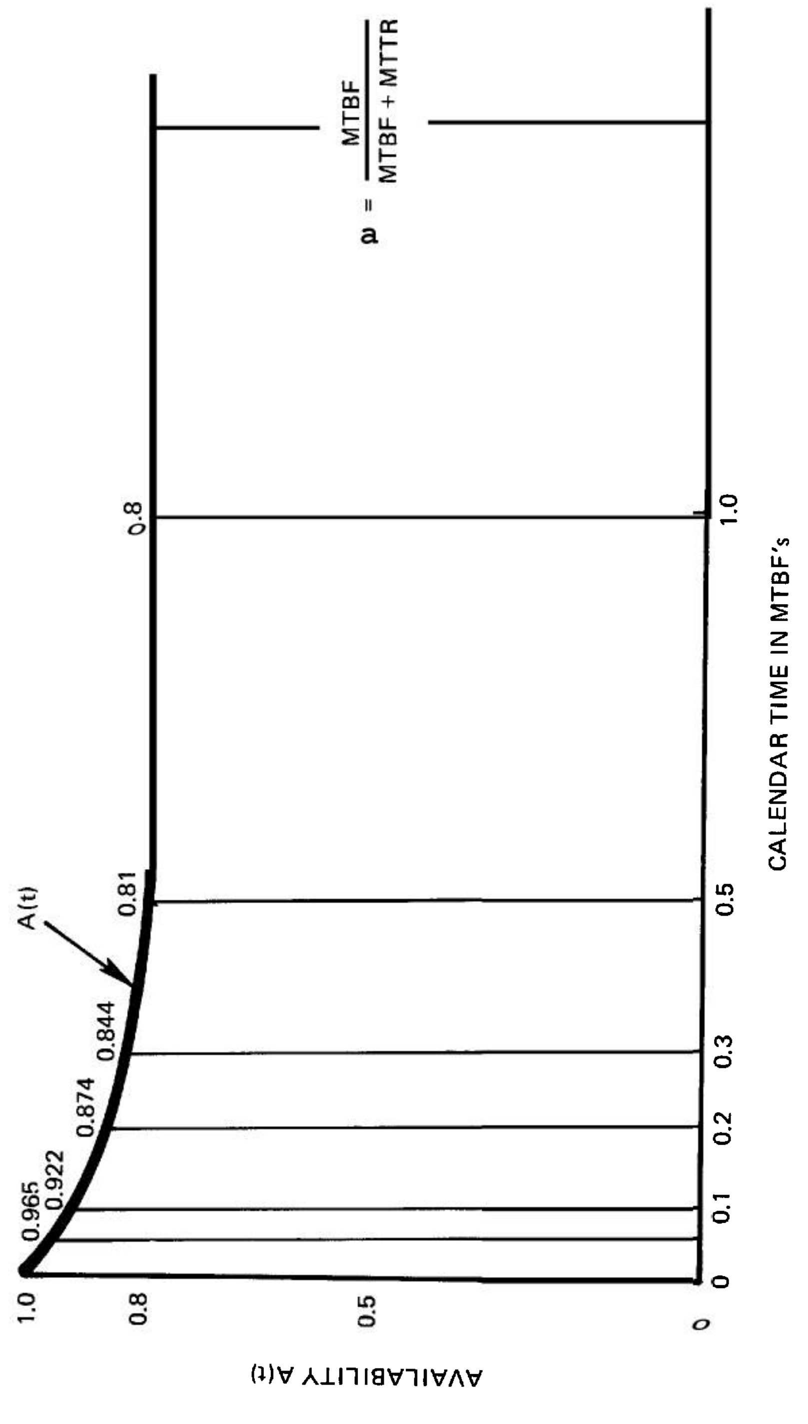


Figure 2-8. Availability of System With Maintenance Time Ratio 1:4

t any point in time. Often, however, one may be interested in knowing what percent or fraction of a time interval $[a,b]$ a system can be expected to operate. This is called the interval or average availability A_{AV} of a system and is given by the time average of the availability function $A(t)$, averaged over the interval $[a,b]$:

$$A_{AV}(a,b) = [1/(b-a)] \int_a^b A(t)\,dt \qquad (2\text{-}65)$$

For instance, if we want to know the fraction of time a system such as shown in Fig. 2-8 will be operating counting from $t = 0$ to any time t, we substitute $A(t)$ of Eqs. 2-61 or 2-62 into Eq. 2-65 and perform the integration. The result is (Ref. 24, p. 34)

$$A_{AV}(t) = \frac{1}{t}\left\{ \int_0^t \frac{\mu}{\lambda+\mu}\,dt + \int_0^t \frac{\lambda}{\lambda+\mu} \times \exp[-(\lambda+\mu)t]\,dt \right\}$$

$$= \frac{\mu}{\lambda+\mu} + \frac{\lambda}{t(\lambda+\mu)^2}\{1 - \exp[-(\lambda+\mu)t]\} \qquad (2\text{-}66)$$

Fig. 2-9 shows the relationship of $A(t)$ to $A_{AV}(t)$ for the exponential case. Note that in the limit, in the steady state, we again get the availability A of Eq. 2-63, i.e.,

$$\lim_{t\to\infty} A_{AV}(t) = \mu/(\lambda+\mu)$$

$$= MTBF/(MTBF + MTTR) \qquad (2\text{-}67)$$

But in the transient state of the process, as shown in the figure for an interval $[0,T]$, before equilibrium is reached, $A_{AV}(t)$ is in the exponential case larger than $A(t)$, for an interval $[0,t]$. This is not true for all distributions since $A(t)$ and $A_{AV}(t)$ may be subject to very large fluctuations in the transient state.

From Eq. 2-66 we may also get the average or expected "on" time in an interval $[0,t]$ by multiplying $A_{AV}(t)$ and t, the length of the time interval of interest. Ref. 25, pp. 74-83, contains an excellent mathematical treatment of the pointwise and interval availability, and related concepts. Earlier work in these areas is found in Refs. 26 and 27.

3. Model *C*

When a series system consists of N units that are separately repairable or replaceable whenever the system goes down because of any one unit failing, the steady-state availability of such a series system is given by

$$A = \left(1 + \sum_{i=1}^{N} \alpha_i\right)^{-1} \qquad (2\text{-}68)$$

where α_i is the maintenance time ratio of the ith unit in the system, i.e.,

$$\alpha_i = (MTTR)_i/(MTBF)_i \qquad (2\text{-}69)$$

Caution is necessary in computing α_i, since Eq. 2-68 applies to the availability of the whole system. Thus, when the units are replaceable as line replaceable units or system replaceable units, the $MTTR_i$ is the mean time required to replace the unit by a good one at the system maintenance level and is not the mean repair time of the failed removed unit. On the other hand if failed units are not replaced but are repaired at the system level, $MTTR_i$ is the mean-time-to-repair the unit, which becomes also the downtime for the system. Thus, when computing the α's of the units and the availability A of the system, all $MTTR$'s must be those repair times that the system experiences as its own downtime. The $MTTR_i$ of the ith unit is thus the system mean repair time when the ith unit fails.

If we compare Eq. 2-68 with Eq. 2-64 in Model A we find that they are identical. The system maintenance time ratio is

$$\alpha = MTTR/MTBF \qquad (2\text{-}70)$$

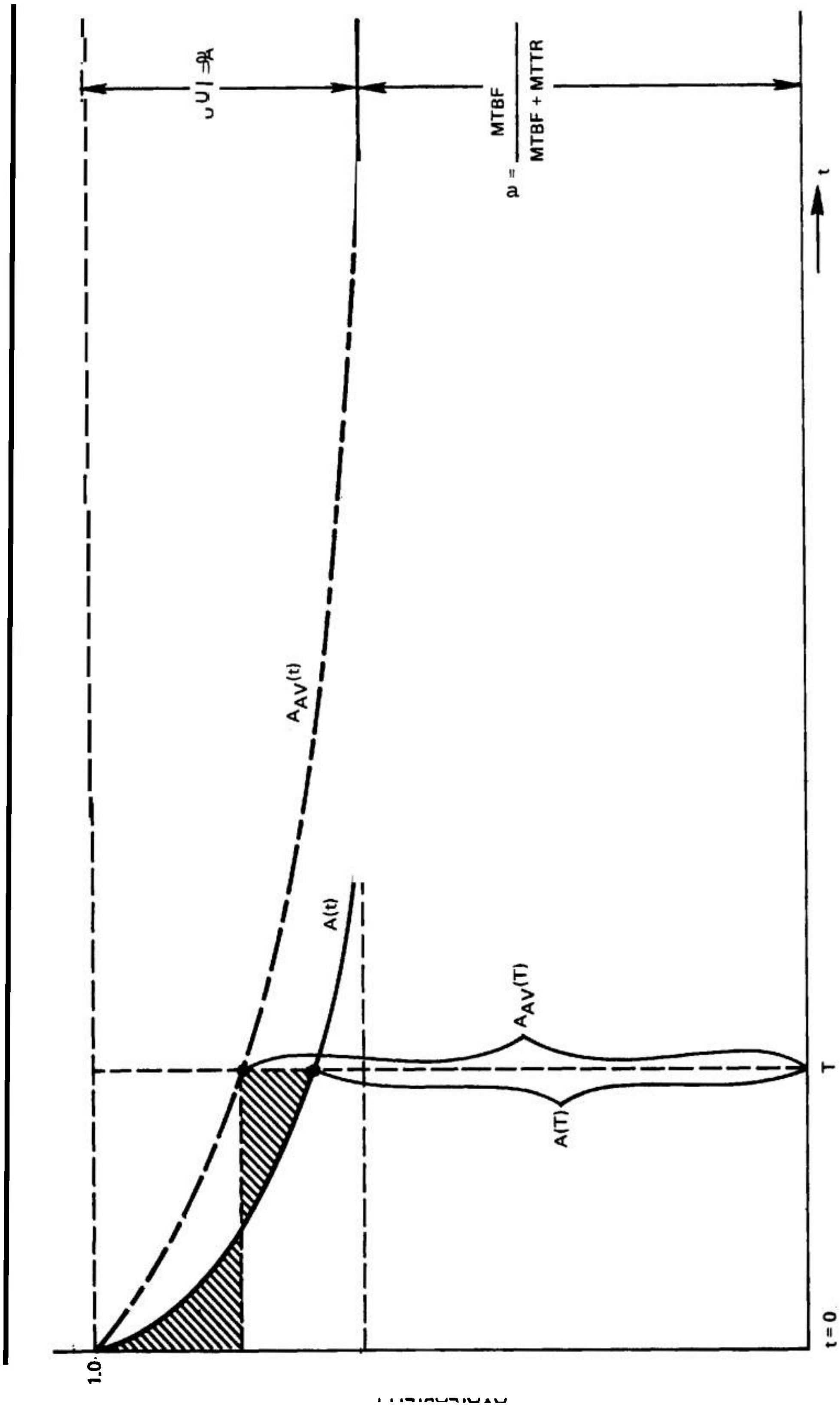
But the serial system MTTR as shown in Chapter 1 is given by

$$MTTR = \sum \lambda_i t_i \Big/ \sum \lambda_i \qquad (2\text{-}71)$$

while its MTBF is

$$MTBF = \left(\sum \lambda_i\right)^{-1} \qquad (2\text{-}72)$$

The ratio α is, therefore, also


A_AV(t)
A(t)
A_AV(T)
A(T)
MTBF
MTBF + MTTR
1.0
t = 0
T
t

$$\alpha = \left(\sum \lambda_i t_i\right)\left(\sum \lambda_i\right) / \sum \lambda_i = \sum \lambda_i t_i = \sum \alpha_i \quad (2\text{-}73)$$

where $\lambda_i = 1/MTBF_i$ and $t_i = MTTR_i$

The maintenance time ratio (MTR) is actually the average system downtime per system operating hour. Conceptually, it is very similar to the maintenance ratio (MR) defined as maintenance manhours expended per system operating hour. The difference is that in the MTR one looks only at system downtime in terms of clock hours of system repair, whereas in the MR one looks at all maintenance manhours expended at all maintenance levels to support system operation.

4. Model D

In this model the availability of some redundant systems is considered (Ref. 24, pp. 36-38). First we deal with two equal, independent units in a parallel redundant arrangement with each unit being separately repairable or replaceable while the other unit continues operating. Thus, the system is "up" if both or any one of the two units operates.

If we define the unavailability U of a unit as

$$U = 1 - A = MTTR/(MTBF + MTTR) \quad (2\text{-}74)$$

then the probability that the system is unavailable is the probability that both units are down at the same time, which is

$$U_{system} = U^2 \quad (2\text{-}75)$$

and system availability is

$$A_{system} = 1 - U^2 \quad (2\text{-}76)$$

Further, using the binomial expansion

$$(A + U)^2 = A^2 + 2AU + U^2 \quad (2\text{-}77)$$

we find that we may write Eq. 2-76 also in the form

$$A_{system} = A^2 + 2AU \quad (2\text{-}78)$$

which gives us the probability A^2 that both units are operating at any point in time, and the probability $2AU$ that only one unit is working. Over a period of time T, the system will, on the average, be for a time TA^2 operating with both units up, while for $2TAU$ only one unit will be up. If the performance of the system is P_1 when both units are up, but only P_2 when only one unit is up, the system output or effectiveness SE over T is expected to be

$$SE = P_1 TA^2 + 2P_2 TAU \quad (2\text{-}79)$$

Assume a ship with two engines which are subject to on-board repair when they fail. When both engines work the ship speed is 30 kt, and when only one engine works it is 20 kt. Let engine $MTBF$ be 90 hr and let its MTTR be 10 hr, so that the availability of an engine is $A = 0.9$ and its unavailability is $U = 0.1$. Over a 24-hr cruise the ship will be expected to travel on the average

$SE = 30 \times 24 \times .81 + 2 \times 20 \times 0.9 \times 0.1 = 583.2 + 86.4 = 669.6$ nmi.

The expected time for the ship to be found idle with both engines out for a 24-hr cruise is

$$T_{idle} = 24U^2 = 24(0.01) = 0.24 \text{ hr} \quad (2\text{-}79a)$$

For three units in parallel we get

$$(A + U)^3 = A^3 + 3A^2U + 3AU^2 + U^3 \quad (2\text{-}80)$$

If the system goes down only if all three units are down, system availability is

$$A_{system} = A^3 + 3A^2U + 3AU^2 \quad (2\text{-}81)$$

but if at least two units are needed for system operation since a single unit is not sufficient, system availability becomes

$$A_{system} = A^3 + 3A^2U \qquad (2\text{-}82)$$

In general, for a system with n equal, redundant units, we expand the binomial term $(A + U)^n$ which yields the probabilities of being in any one of the possible states. Then, by adding the probabilities of the ***acceptable*** states, we obtain the availability of the system. As stated earlier, the units must be independent of each other, both in terms of their failures and in terms of their repairs or replacements, with no queuing up for repair.

Ref. **28** contains, throughout the text, extensive tabulations of availability and related measures of multiple parallel and standby redundant systems for cases of unrestricted as well as restricted repair when failed redundant units must queue up and wait until their turn comes to get repaired.

Returning briefly to Eq. **2-75,** when the two redundant units are not equal but have unavailabilities $U_1 = 1 - A_1$ and $U_2 = 1 - A_2$, system unavailability becomes

$$U_{system} = U_1U_2 \qquad (2\text{-}83)$$

and availability

$$A_{system} = 1 - U_1U_2 \qquad (2\text{-}84)$$

Again, we may expand the multinomial

$$(A_1 + U_1)(A_2 + U_2) = A_1A_2 + A_1U_2 + A_2U_1 + U_1U_2 \qquad (2\text{-}85)$$

and may write system availability in the form

$$A_{system} = A_1A_2 + A_1U_2 + A_2U_1 \qquad (2\text{-}86)$$

For n-unequal units we expand the term

$$\prod_{i=1}^{n} (A_1 + A_2) = 1 \qquad (2\text{-}87)$$

and add together the probabilities of acceptable states to obtain system availability and other effectiveness measures, as illustrated in the ship engines example.

5. Model E

A very different situation in availability modeling is encountered when system "uptime" is not measured in hours of operation or any time parameter, but rather in terms of number of rounds fired, miles travelled, actuations or cycles performed, etc. The reliability parameter is then no longer expressed in terms of MTBF, but rather in mean-rounds-between-failures (*MRBF*), mean-miles-between-failures (*MMBF*), mean-*cycles*-between-failures (*MCBF*), etc. The failure rate then also is expressed in number of failures per round, per mile, per cycle but not in number of failures per operating hour.

For straightforward reliability calculations this poses no problem since the same reliability equations apply as in the time domain, except that the variable time t in hours is replaced by the variable number of rounds, number of miles, etc. We may then calculate the reliability of such systems for one, ten, one hundred, or any number of rounds fired or miles travelled, as we wish. The maintainability calculations remain as before, since downtime will always be measured in terms of time, and the parameter of main interest remains the ***MTTR***.

However, when it comes to availability, which usually combines two time parameters, i.e., the MTBF and the ***MTTR*** into a probability of the system being up at some time t, a difficult problem arises when the time t is replaced by rounds or miles since the correlation between time and rounds or time and miles is quite variable.

An equation for the steady-state availability of machine guns is given in Ref. 29. This equation is based on a mission profile that at discrete times t_1, t_2, t_3, etc., requires the firing of N_1, N_2, N_3, etc., bursts of rounds. When the gun fails during a firing, say at time t_3, it fires only f rounds instead of N_3 rounds and must undergo repair during which repair time it is not available to fire, i.e., fails to fire let's say a required N_4 rounds at t_4 and a further N_5 rounds at t_5 before becoming again available (see Fig. 2-10). Its availability A based on the rounds not fired during repair may be expressed, for the described history, as

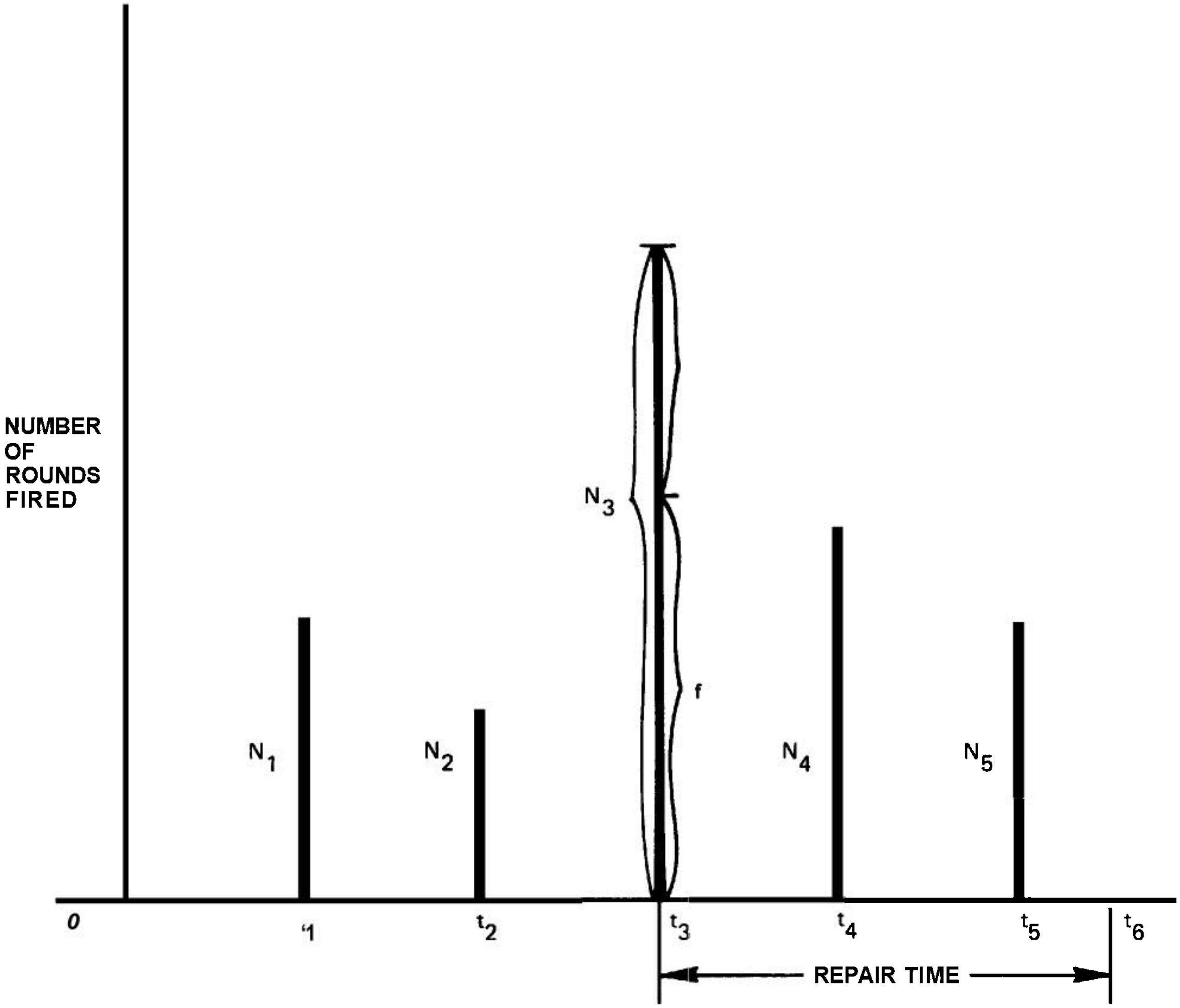


Figure 2-10. Hypothetical History of Machine Gun Usage

$$A = (N_1 + N_2 + f)/(N_1 + N_2 + N_3 + N_4 + N_5) \quad (2\text{-}88)$$

Such sequence of rounds fired followed by rounds missed (not fired) constitutes a renewal process in terms of rounds fired, as shown in Fig. 2-11, where the gun fails after firing x rounds, fails to fire $\gamma(x)$ rounds in the burst of rounds during which it failed and also misses firing the required bursts of rounds while in repair for an MTTR = M Assume that the requirements for firing bursts of rounds arrives at random according to a Poisson process with rate r and the average number of rounds per burst is N, then, the limiting availability of the gun may be expressed as

$$A = MRBF/(MRBF + N + rMN) \quad (2\text{-}89)$$

where $MRBF$ is the mean number of rounds to failure. The derivation of this formula developed by R. E. Barlow, is contained in the Appendix of Ref. 29. To calculate A from Eq. 2-89 one must know the $MRBF$ and **MTTR** of the gun, the average rounds Nfired per burst, and the rate r at which requirements for firing bursts of rounds arrive.

Similar availability equations can be developed for other types of weapons, and also for vehicles where the renewal process is in terms of miles travelled. Other approaches to calculating the availability of guns, as well as vehicles, are found in Ref. 30, and are based on calculating, from historical field data, the maintenance ratios and, via regression analysis, the maintenance time ratios (called the "maintenance clock hour index") that are in turn used in the conventional time-based equations of inherent, achieved, and operational availability.

For example, consider a machine gun system in a tank on which historical data are available, showing that 0.014 corrective maintenance manhours are expended per round fired, and that per year 4800 rounds are fired while the vehicle travels for 240 hr per yr. The maintenance ratio (MR) for the gun system is then computed as (Ref. 30, pp. 36-38):

$$MR_{Gun} = \frac{MMH}{Round} \times \frac{Number\ of\ Rounds\ Fired\ per\ Annum}{Vehicle\ Operating\ Hours\ per\ Annum}$$

$$= 0.014 \times (4800/240) = 0.28 \quad (2\text{-}90)$$

The dimensions for 0.28 is gun system maintenance manhours per vehicle operating hour. The corrective maintenance time ratio α (called maintenance clock hour index Ω) is, according to this example, given by

$$\alpha_{Gun} = 0.628(0.28)^{0.952} = 0.187 \quad (2\text{-}91)$$

The numbers 0.628 and 0.952 are the intercept and the regression coefficients, respectively, obtained by regression analysis as developed in Ref. 30, p. 18, Table 1. The dimension for α_{Gun} is gun system downtime per vehicle operating hour. The inherent availability of the gun system is then, according to the conventional time equation, Eq. 2-64,

$$A_i = (1 + \alpha_{Gun})^{-1} = (1.187)^{-1} = 0.842 \quad (2\text{-}92)$$

This may be interpreted as the gun system being available for 84.2% of the vehicle operating time. Caution is required in using this approach for weapon availability calculations since in the case where the vehicle would have to be stationary and the gun would still fire rounds, MR and α would become infinitely large and the inherent availability of the gun system would become zero.

2-4.3 COMPLEX MODELS

In complex system effectiveness mathematical models an attempt is made to relate the impact of system reliability, maintainability, and performance to the mission profiles, scenario, use, and logistic support. Only in simple situations can a meaningful single model be developed that will relate all these parameters and yield a single quantitative measure of system effectiveness. Numerous complex models exist and, as a matter of fact, every major company in the aerospace business has developed a multitude of such models, claimed to be unique and the only meaningful ones, and uses them primarily as sales tools. In the following paragraphs we discuss some of these models which have achieved a certain popularity and a degree of acceptance.

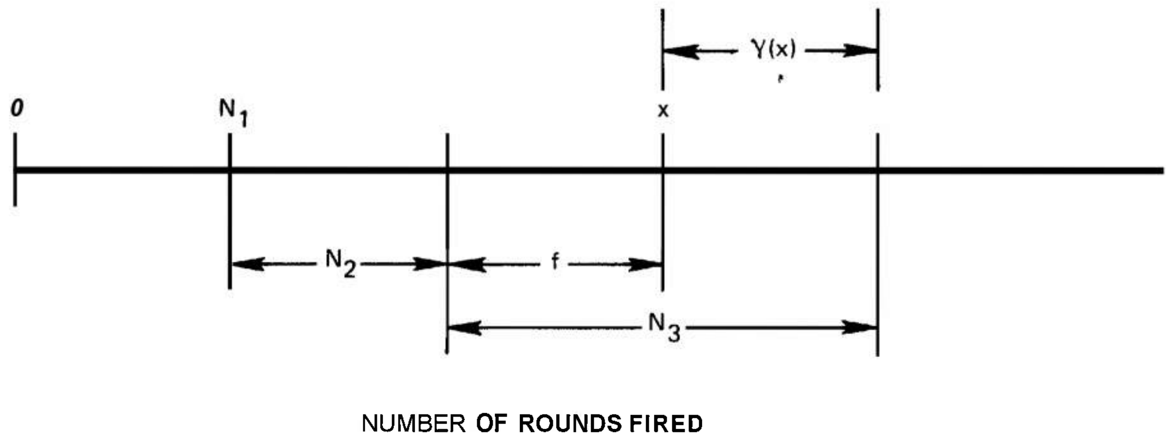


Figure 2-11. Renewal Process in Terms of Rounds Fired

2-4.3.1 The WSEIAC Model

This model is briefly introduced in par. 2-2.2 and is developed in the reports of the Weapon System Effectiveness Industry Advisory Committee (Ref. 2). Basically, the model is a product of three matrices: the Availability row vector $\vec{A}$, the Dependability matrix D, and the Capability matrix C. In the most general case assume that a system can be in different states, and at any given point in time is in either one or the other of the states. The availability row vector is then

$$\vec{A} = [a_1 a_2 a_3 \cdots a_i \cdots a_n] \tag{2-93}$$

where a_i is the probability that the system is in state i at a random mission beginning time. Since the system can be in only one of the n states and n is the number of all possible states it can be in, including the down states in which the system cannot start a mission, the sum of all the probabilities a_i in the row vector must be unity, i.e.,

$$\sum_i^n a_i = 1 \tag{2-94}$$

The dependability matrix D is defined as a square n X n matrix

$$\begin{bmatrix} d_{11} & d_{12} & d_{13} \cdots d_{1n} \\ d_{21} & d_{22} & d_{23} \cdots d_{2n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ d_{n1} & d_{n2} & d_{n3} \cdots d_{nn} \end{bmatrix} \tag{2-95}$$

where the meaning of the element d_{ij} is defined as the expected fraction of mission time during which the system will be in state j if it was in state i at the beginning of the mission. If system output is not continuous during the mission but is required only at a specific point in the mission (such as over the target area), d_{ij} is defined as the probability that the system will be in state j at the time when output is required if it was in state i at mission start.

When no repairs are possible or permissible during a mission, the system, upon failure or partial failure cannot be restored to its original state during the mission and can at best remain in the state i in which it started the mission, or will degrade into lower states, or fail completely. In the case of no repairs during the mission some of the matrix elements become zero. If we define state 1 as the highest state (i.e., everything works perfectly), and n the lowest state (i.e., complete failure), the dependability matrix becomes triangular with all entries below the diagonal being zeros.

$$D = \begin{bmatrix} d_{11} & d_{12} & d_{13} \cdots d_{1n} \\ 0 & d_{22} & d_{23} \cdots d_{2n} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ 0 & 0 & 0 \cdots d_{nn} \end{bmatrix} \tag{2-96}$$

If the matrix is properly formulated the sum of the entries in each row must equal unity. For example, for the first row we must have

$$d_{11} + d_{12} + \cdots + d_{1n} = 1 \tag{2-97}$$

and the same must apply to each subsequent row. This provides a good check when formulating a dependability matrix.

The capability matrix C describes system performance or capability to perform while in any of the n possible system states. If only a single measure of system effectiveness is of importance or of interest, C will be a one column matrix with n elements, such as

$$C = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{bmatrix} \tag{2-98}$$

where c_j represents system performance when the system is in state j.

System effectiveness SE, in the WSEIAC model is then defined as

$$SE = [a_1 \quad a_2 \cdots a_n] \times \begin{bmatrix} d_{11} & d_{12} \cdots d_{1n} \\ d_{21} & d_{22} \cdots d_{2n} \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ d_{n1} & d_{n2} \cdots d_{nn} \end{bmatrix} \times \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ \cdot \\ c_n \end{bmatrix}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_i d_{ij} c_j \qquad \text{(2-99)}$$

in matrix notation.

Ref. 2 contains several numerical examples of how to perform system effectiveness calculations using the WSEIAC model. Also Ref. 31, Chapter VII, discusses this model at length and provides numerical examples.

2-4.3.2 Other Models

System effectiveness analyses, in conjunction with life cycle costing, provide a tool for the decision maker to use in determining which design approach to choose out of a number of alternatives. A single mathematical model that would be all-inclusive is seldom possible to construct, except in simple situations. In most cases multiple models will be needed. An example of such a multiple model approach is given in Ref. 22. The models described there were developed for design decisions on the F-1 11, and are briefly discussed here. They fall into four major categories:

1. Maintenance Analysis and Review Technique (MART)
2. Logistic Assets Requirements Models (LARM)
3. Related Effectiveness Models
4. Cost and Cost/Effectiveness Models

The relationship of these models is shown in Fig. 2-12.

The MART group of maintenance models consists of:

1. Subsystem Simulation Model (SSM) which establishes for each subsystem of the aircraft the probability and time distributions for maintenznce, skills, equipment, and facilities.

2. Network Analysis Model (NAM) which evaluates the turn-around sequence and defines critical activities for the maintenance required on subsystems.

3. Base Maintenance and Operations Model (BMOM) which simulates a fleet of aircraft in a real-world base environment, i.e., subject to constraints of schedules and assets.

The LARM group of logistic models consists of:

1. Shop Maintenance Model (SHMM) which simulates maintenance shop activities, including flight line maintenance.

2. Inventory Policy Model (IPM) which computes composition of base inventory and maintenance kits for maximum fill rate, minimum cost, and minimum weight.

3. Spares Provisioning and Requirements Effectiveness Model (SPAREM) which determines the spare requirements and delay times for varying logistic policies, operational loads, and flight programs.

The group of Related Effectiveness Models is used to simulate missions and to measure system effectiveness. The principal effectiveness models are:

1. Tactical Air-to-ground Effectiveness Model (TAGEM) which basically considers different target types, multiple mission types, and variations in environment.

2. Weapon Delivery Model (WDM) which considers weapon types, weapon delivery methods, system reliability, accuracy, survivability, etc.

3. Effectiveness Simulation Model (ESM) which simulates aircraft fleet deployment and combat operations, considering availability, in flight reliability, re-sortie capability, etc.

Two additional effectiveness models are the Naval Air Effectiveness Model (NAEM) and the Air Battle Model (ABM).

Also, supporting models are used in support of the above described models, such as the Maintenance Analysis Requirements Model (MARM), the Reliability Requirements Analysis Model (RRAM), and others.

Finally the Cost and Cost/Effectiveness Models and the Incremental Cost/Effectiveness Model (ICEM) are used to determine total program costs, relate these to the effectiveness parameters obtained from the preceding models, and evaluate effects of design changes on an incremental basis in an extensive trade-off and sensitivity analysis. For more details on these models the reader is referred to the original work (Ref. 22).

Fig. 2-13 shows in a simplified way the relation of the models described here to the WSEIAC model.

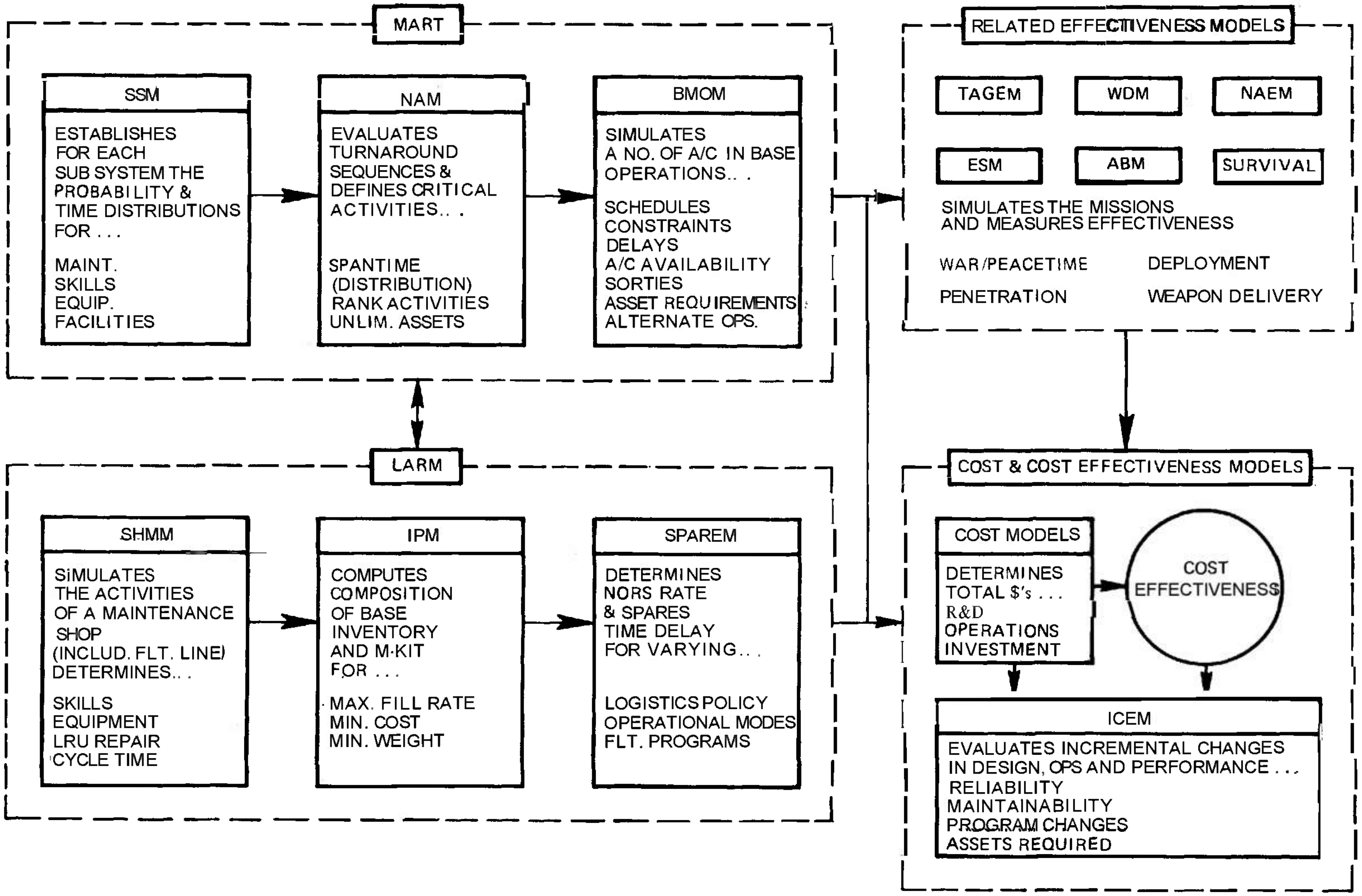


Figure 2-12. Interrelationships of Effectiveness Mudels
Adapted from C. B. Moore, *An Integrated Approach to Determining Operational or System Effectiveness*, Report No. MR-0-128, 2 March 1966. Used with permission of General Dynamics, Fort Worth, Texas.

AVAILABILITY	MAINTENANCE: MART (SSM, NAM, BMOM) REQUIREMENTS: MARM, RRAM LOGISTICS: LARM (SHMM, IPM, SPAREM)
DEPENDABILITY AND CAPABILITY	RELATED EFFECTIVENESS MODELS (TAGEM, ESM, WDM) PENETRATION AND SURVIVAL MODELS MISSION RELIABILITY MODELS

Figure 2-13. Complex Models Relationship

2-5 TRADE-OFF TECHNIQUES

2-5.1 GENERAL

System effectiveness and cost/effectiveness models provide the best tools for performing trade-off studies on the system level. Through the computerized models any changes in any of the multitude of reliability, maintainability, performance, mission profile, logistic support, and other parameters can be immediately evaluated as to their effect on the effectiveness and total cost of a system. Thus cost effectiveness modeling and evaluation, besides being used for selecting a specific system design approach from among several competing alternatives, is a very powerful tool for performing parametric sensitivity studies and trade-offs down to component level when optimizing designs to provide the most effective system for a given budgetary and life-cycle cost constraint or the least costly system for a desired effectiveness level.

At times, however, especially in the case of the more simple systems, trade-offs may be limited to achieving a required system availability while meeting the specified reliability and maintainability requirements. Comparatively simple trade-off techniques can then be used as shown'in the following paragraph. This is then followed by a discussion and explanation of linear programming as a general mathematical tool for certain trade-off situations. The maintainability design trade-off aspects and the cost-oriented trade-offs are discussed at length in Chapters 5 and 7.

2-5.2 RELIABILITY VS MAINTAINABILITY

As stated earlier in this chapter and in Chapter 1, reliability and maintainability jointly determine the inherent availability of a system. Thus, when an availability requirement is specified, there is a distinct possibility of trading off between reliability and maintainability since, in the steady state, availability depends only on the ratio or ratios of *MTTR/MTBF* that in par. 2-4.2.2 we referred to as maintenance time ratio (*MTR*) and used the symbol α, i.e.,

$$\alpha = MTTR/MTBF \tag{2-100}$$

so that the inherent availability equation assumed the form

$$A_i = 1/(1 + \alpha) \tag{2-101}$$

Now, obviously, innumerable combinations of MTTR and *MTBF* will yield the same α and, therefore, the same availability A, However, there is usually also a mission reliability requirement specified and also a maintainability requirement. Both of these requirements must also be met in addition to the availability requirement.

Ref. 32 provides a trade-off example that is repeated here, for convenience, in a somewhat different form. Fig. 2-14 represents a system consisting of five major subsystems in a series arrangement. The *MTBF* of this system is

$$MTBF = \left(\sum \lambda_i\right)^{-1} = (0.0775)^{-1} = 12.9\text{hr} \tag{2-102}$$

and its *MTTR* is

$$MTTR = \sum \lambda_i (MTTR)_i \Big/ \sum \lambda_i = 0.33(0.0775)^{-1} = 4.26\text{hr} \tag{2-103}$$

Since the maintenance time ratio equals

$$\alpha = 4.26(12.9)^{-1} = 0.33 \tag{2-104}$$

which is the sum of the maintenance ratios of the five serial subsystems

$$\alpha = \sum \alpha_i = 2/100 + 1/200 + 5/25 + 5/50 + 2/400 = 0.33 \tag{2-105}$$

then

$$A = (1 + 4.26/12.9)^{-1} = 0.752 \tag{2-106}$$

By inspection of Eq. 2-105 we see that Subsystems 3 and 4 have the highest maintenance time ratios, i.e., 0.2 and 0.1, and therefore are the "culprits" in limiting system availability to 0.752 which may be completely unacceptable.

If, because of state-of-the-art limitations it is not possible to increase the *MTBF*'s of these two subsystems and their *MTTR*'s cannot be reduced by repackaging, the first recourse could be the adding of a parallel redundant subsystem to Subsystem 3. Now two cases may have to be considered (a) the case where no repair of a failed redundant unit is possible until both

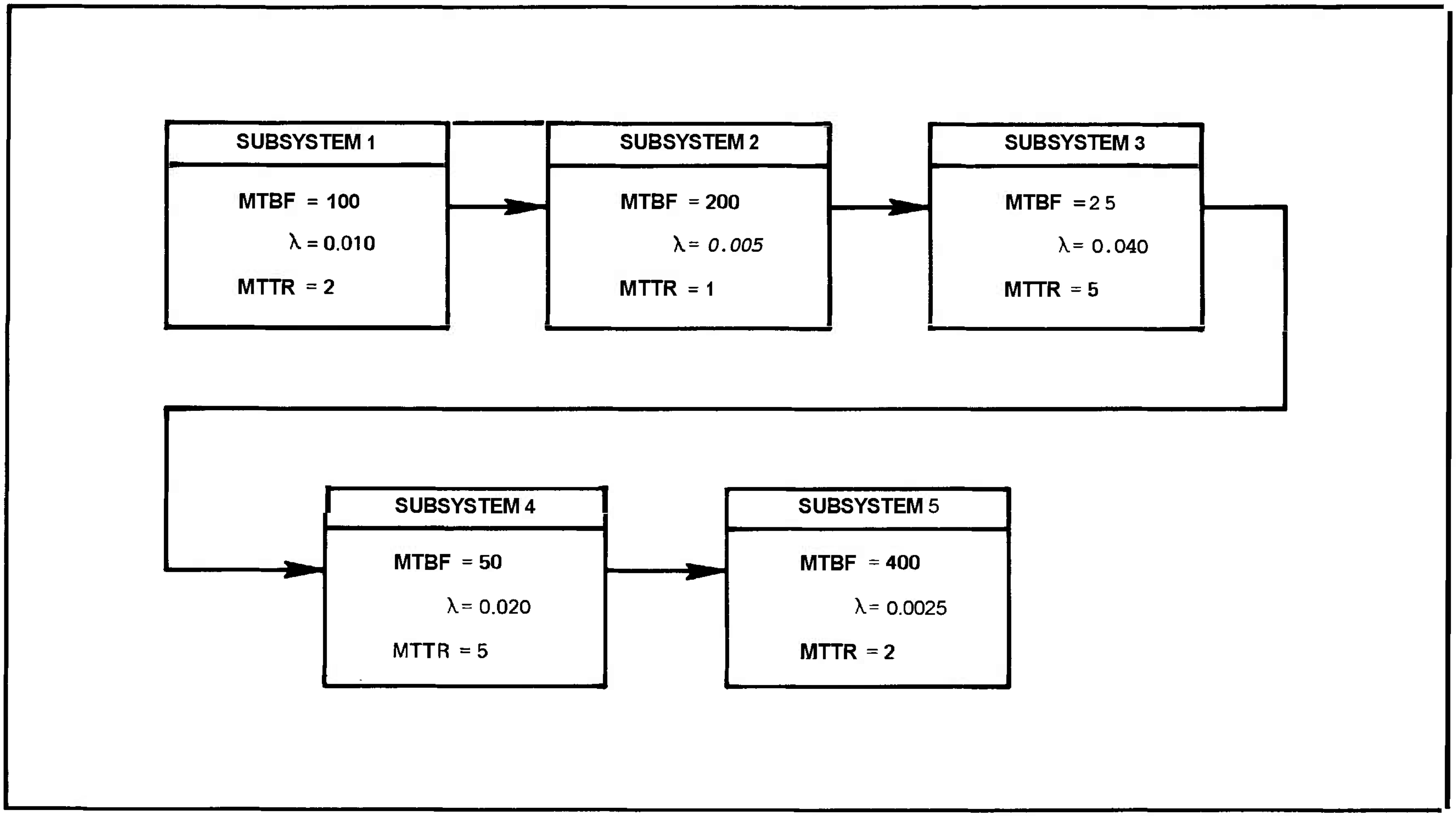


Figure 2-14. Block Diagram of a Series System

fail and the system stops operating, or (b) repair is possible while the system is operating.

In the first case the MTBF of Subsystem 3, which now consists of two parallel units, becomes 1.5 times that of a single unit, i.e., 1.5 X 25 = 37.5 hr. With both units failed, both must be repaired. If a single crew repairs both in sequence, the new MTTR becomes 2 hr and availability actually drops. If two repair crews simultaneously repair both failed units, and repair time is assumed exponentially distributed, the MTTR of both units is again 1.5 times that of a single unit, or, 1.5 hr, and system availability remains the same as before, with nothing gained. But if repair of a failed redundant unit is possible while the system operates, the steady-state availability of Subsystem 3 becomes (Ref. 33, p. 133 and Ref. 28, p. 123)

$$A_3 = (\mu^2 + 2\lambda\mu)/(\mu^2 + 2\lambda\mu + \mu^2) \qquad (2\text{-}107)$$

for a single repair crew. Since, for a single unit in this subsystem the failure rate A = 0.04 and the repair rate $\mu = 1/5 = 0.2$, we get

$$A_3 = (0.04 + 2 \times 0.04 \times 0.2)(0.04 + 2 \times 0.04 \times 0.02 + 2 \times 0.0016)^{-1}$$
$$= 0.056(0.0592)^{-1} = 0.946 \qquad (2\text{-}108)$$

as compared to 0.833 when no redundancy was used. The value of $A_3 = 0.946$ of the redundant configuration corresponds to a maintenance time ratio of

$$\alpha_3 = (1 - A_3)A_3^{-1} = 0.054(0.946)^{-1} = 0.057 \qquad (2\text{-}109)$$

The whole system maintenance time ratio now becomes

$$\alpha = \sum \alpha_i = 0.02 + 0.005 + 0.057 + 0.1 + 0.005 = 0.187 \qquad (2\text{-}110)$$

and system availability A is

$$A = (1 + 0.187)^{-1} = (1.187)^{-1} = 0.842 \qquad (2\text{-}111)$$

as compared with 0.752 without redundancy in Subsystem 3. If this new value of availability is still not acceptable, redundancy would also have to be applied to Subsystem 4. But to achieve these gains in availability, repair of failed redundant units must be possible while the system is operating. This is called availability with repair. Otherwise, redundancy will not increase availability and may even reduce it, even though it increases system reliability.

A different method of straightforward trade-off between reliability and maintainability is shown in Fig. 2-15 (Ref. 34, p. 81). The specific trade-off example shown in this figure is based on a requirement that the inherent availability of the system must be at least A = 0.99, the *MTBF* must not fall below 200 hr, and the MTTR must not exceed 4 hr. The trade-off limits are within the shaded area of the graph, resulting from the equation for inherent availability

$$A_i = MTBF/(MTBF + MTTR) \qquad (2\text{-}112)$$

The straight line for A = 0.99 goes through the points (200,2) and (400,4), the first number being the MTBF and the second number being the *MTTR*. Any system with an MTBF larger than 200 hr and an MTTR smaller than 4 hr will meet or exceed the minimum availability requirement of A = 0.99. If there are several system design alternatives that comply with the specification requirements, the design decision is made by computing the life-cycle costs of each alternative and usually selecting the least expensive system, unless substantial gains in system effectiveness are achieved which would warrant increasing the expenditures.

2-5.3 LINEAR PROGRAMMING

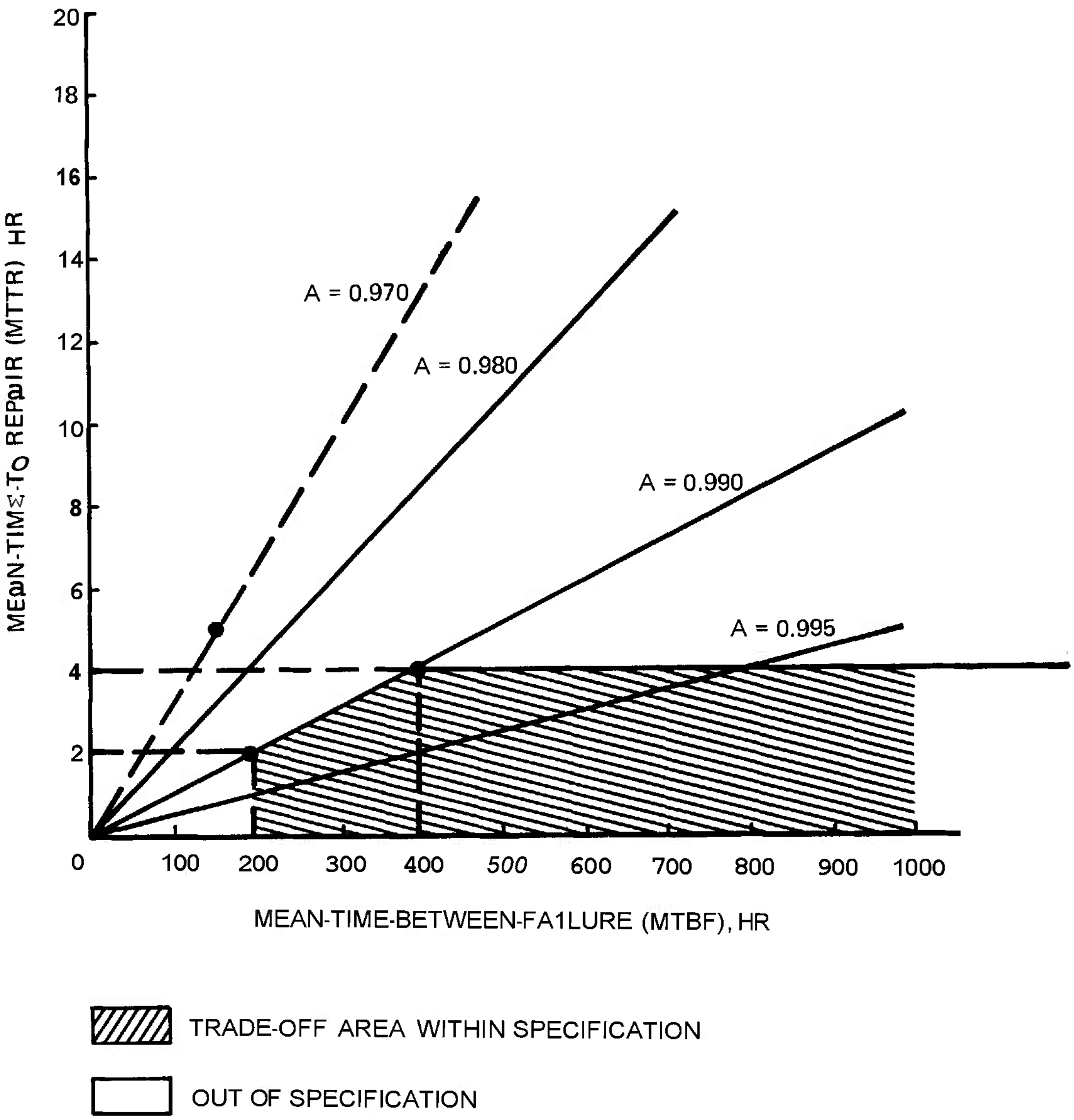
Linear programming is a mathematical technique for devising an optimal allocation of scarce resources among competing activities in an optimal manner.

The adjective "linear" means that the variables that appear in the problem, in both the objective function and the constraints, do so as linear functions. The word "programming" is used in the sense of planning, not in the sense of preparing instructions for a computer—although, of course, computers are often used in the solution of linear programming problems.

A general statement of the linear programming problem is: find values of $x_1, x_2, \ldots, x_n$ which maximize the linear function

$$Z = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n$$

subject to the conditions



REQUIREMENT

A = 99%

MTBF = 200 HR MIN

MTTR = 4 HR MAX

Figure 2-15. Reliability-Maintainability Trade-offs

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \le b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \le b_2$$
$$\cdots \qquad (2\text{-}113)$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \le b_m$$
$$x_1 \ge 0,\ x_2 \ge 0, \ldots, x_n \ge 0$$

The function whose maximum value is being sought is called the *objective function;* the variables $x_1, x_2, \ldots, x_n$ are called *decision variables,* the conditions under which the maximum value of the objective function is sought are called *constraints*—in particular, those appearing on the last line are called the *non-negativity constraints.* The quantities c_j, a_{ij}, and b_i are known constants. The b_i are required to be positive.

The quantities $x_1, x_2, \ldots, x_n$ represent the level of the n-competing activities; Z is the measure of effectiveness by which the optimality of the allocation of resources among the activities is judged; c_j represents the change in Z per unit increase in x_j. There are m-scarce resources with b_i representing the amount of the ith resource available for allocation among the n activities; a_{ij} is the amount of the ith resource needed for one unit of the jth activity—thus, ith constraint is the amount of the ith resource used up. The non-negativity constraints state the condition that no activity level can be negative.

Any n-tuple $(x_1, x_2, \ldots, x_n)$ that simultaneously satisfies all the constraints is called a *feasible solution.* The set of all feasible solutions is sometimes called the *feasible region.* One seeks, then, to find the maximum value of the objective function over the feasible region.

2-5.3.1 General Features of a Linear Programming Problem and an Example

The specification of a linear programming problem just given will now be illustrated and a method of solution will be outlined. In addition, other forms of linear programming problems will be treated. These other forms involve the minimization, rather than the maximization, of the objective function; also, some or all of the constraints may involve greater-than-or-equal-to signs ($\ge$) or equalities ($=$), rather than less-than-or-equal-to ($\le$).

1. Example

A small machine shop manufactures two models, standard and deluxe, of an unspecified product. Each standard model requires 4 hr of grinding and 2 hr of polishing; each deluxe model requires 2 hr of grinding and 5 hr of polishing. The manufacturer has two grinders and three polishers; in his 40-hr week, therefore, he has 80 hr of grinding capacity and 120 hr of polishing capacity. He makes a profit of $3 on each standard model and $4 on each deluxe model. He can sell all he can make of both.

How should the manufacturer allocate his production capacity to standard and deluxe models; i.e., how many of each model should he make in order to maximize his profit? (This example comes from Ref. 35.)

This verbal description must be converted to an algebraic one. Let x_1 denote the number of standard models produced, x_2 the number of deluxe models produced, and P the profit in dollars.

The grinder is used 4 hr for each standard model and 2 hr for each deluxe model, so $4x_1 + 2x_2$ is the number of hours of grinder time used. This cannot exceed the number of hours of grinder time available, 80 hr. Thus the grinder constraint can be stated as

$$4x_1 + 2x_2 \le 80 \qquad (2\text{-}114)$$

Note that a less-than-or-equal-to sign is appropriate here, not an equality sign. This comes about because an optimal (i.e., maximum profit) allocation may leave some grinder capacity unused.

The polisher is used 2 hr for each standard model and 5 hr for each deluxe model, so $2x_1 + 5x_2$ is the number of hours of polisher time used. This cannot exceed the number of hours of polisher time available, 120 hr. Thus, the polisher constraint is

$$2x_1 + 5x_2 \le 120 \qquad (2\text{-}115)$$

Again, the less-than-or-equal-to sign is proper because an optimal solution may leave some polisher capacity unused.

Negative amounts of production of standard and deluxe models do not make sense, so that we have the non-negativity constraints: $x_1 \geq 0$, $x_2 \geq 0$.

Each standard unit sold contributes \$3 to profit, while each deluxe unit sold contributes **\$4.** Thus, the total profit is $P = 3x_1 + 4x_2$.

Stated mathematically, then, the problem of optimal (i.e., maximum profit) allocation of resources (grinder and polisher) between activities (production of standard units and production of deluxe units) is:

a. maximize $P = 3x_1 + 4x_2$

b. subject to:

$$4x_1 + 2x_2 \text{ I } 80$$
$$2x_1 + 5x_2 \text{ I } 120$$
$$x_1 \geq 0, x_2 \geq 0.$$

Since only two variables, x_1 and $x,,$are involved in this formulation, it is possible to show the feasible region graphically as well as to superimpose various equal-profit lines on this feasible region and, thereby, to determine the maximum profit. Later, we will present and discuss purely analytical procedures for solving linear programming problems. Such methods are, of course, needed as the general linear programming will typically involve more than two variables and graphical methods will not be sufficient.

Returning now to the example, we first note that the non-negativity constraints require that the feasible region be in the first quadrant of the x_1 x, plane. The set of points satisfying the grinder constraint, $4x_1 + 2x_2 \leq 80$, are those which lie on and below the line $4x_1 + 2x_2 = 80$.

The set of points which satisfy an inequality $ax, + bx_2 \leq c$ lies on, and on one side of, the line $ax_1 + bx,$ I c. Which side of the line can be determined by checking some point off the line to see if it satisfies the inequality. If it does, so do all points on the same side of the line; if not, the other side of the line is appropriate. The origin (0,0) is a convenient test point unless the line passes through the origin. (This happens if and only if $c = 0$.) In that event, one of the points (1,0) or (0,1) may be a suitable test point. The set of points satisfying the polisher constraint, $2x_1 + 5x_2$ I **120,** lies on and below the line $2x_1 + 5x_2 = 120$. The separate constraint sets and the feasible region (shaded), where all constraints are satisfied simultaneously, are shown in Fig. **2-16.**

It should be noted that the feasible region is a convex polygon. A set of points is convex if the line segment joining any pair of points in the set lies entirely within the set. It is a fact that the feasible region for any linear programming problem is a convex polygon. If the linear programming problem involves more than two decision variables, then the feasible region will be a multidimensional polygon, not a planar (i.e., two-dimensional) polygon as in the example.

In Fig. **2-17** we superimpose equal-profit lines on the feasible region. That is, we plot $3x_1 + 4x_2 = P$for various values of P. (Note that this produces a family of parallel lines.) Our goal is to increase Pas much as possible while still having the line $3x_1 + 4x_2 = P$ retain contact with (i.e., intersect) the feasible region. (Note that as P increases, the line $3x_1 + 4x_2 = P$ moves away from the origin.) The largest value that P can have, compatible with the requirements, is seen to be **110.** This value is attained at the vertex (10,20) of the feasible region. (The coordinates of this vertex can be read off from a carefully prepared graph or, more generally, by solving simultaneously the pair of equations describing the lines which intersect at the vertex. These equations are: $4x_1 + 2x3 = 80$ and $2x_1 + 5x_2 = 120$.)

We have now reached a solution to our linear programming problem. The maximum profit attainable, subject to the stated constraints on grinder and polisher time, is **\$1 10.** This maximum profit is achieved when the manufacturer produces **10** standard models and **20** deluxe models per week.

While in this example the optimum allocation of the two resources, grinder time and polisher time, used all that was available of each, this will not always occur, i.e., an optimum allocation of resources may leave some amount of resources unused.

We call attention to the fact that the optimum value of the objective function in this example was found at a vertex of the feasible region. This was not a fortuitous occurrence, but is a general property of linear programming problems. It obtains also if one seeks the minimum of a linear objective function, or if some of the constraints involve equalities or greater-than-or-equal-to inequalities. The reader can gain further insight into this fact by considering families of equal profit lines of differing slopes and observing that in each case the

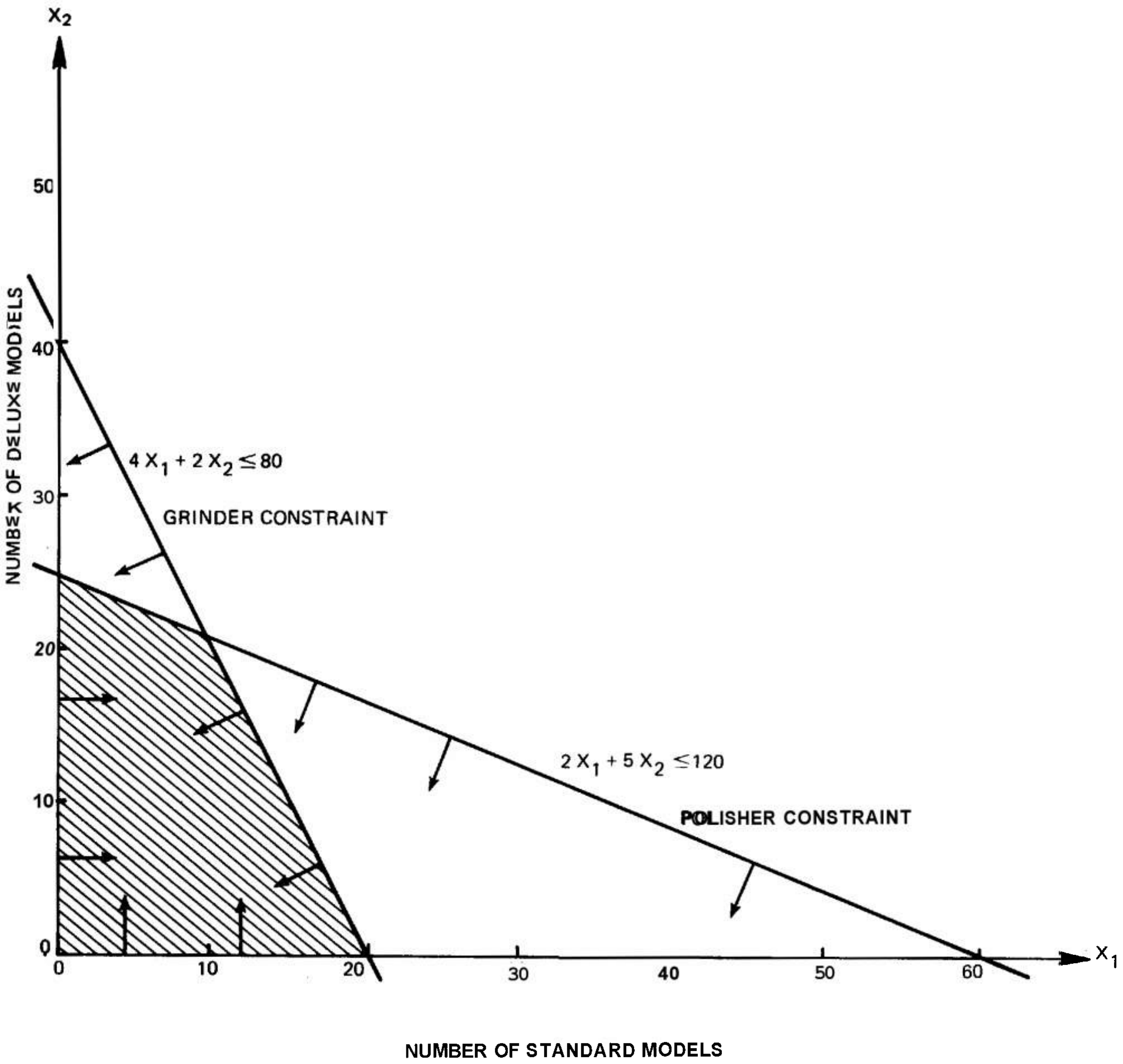


Figure 2-16. Constraints and Feasible Region

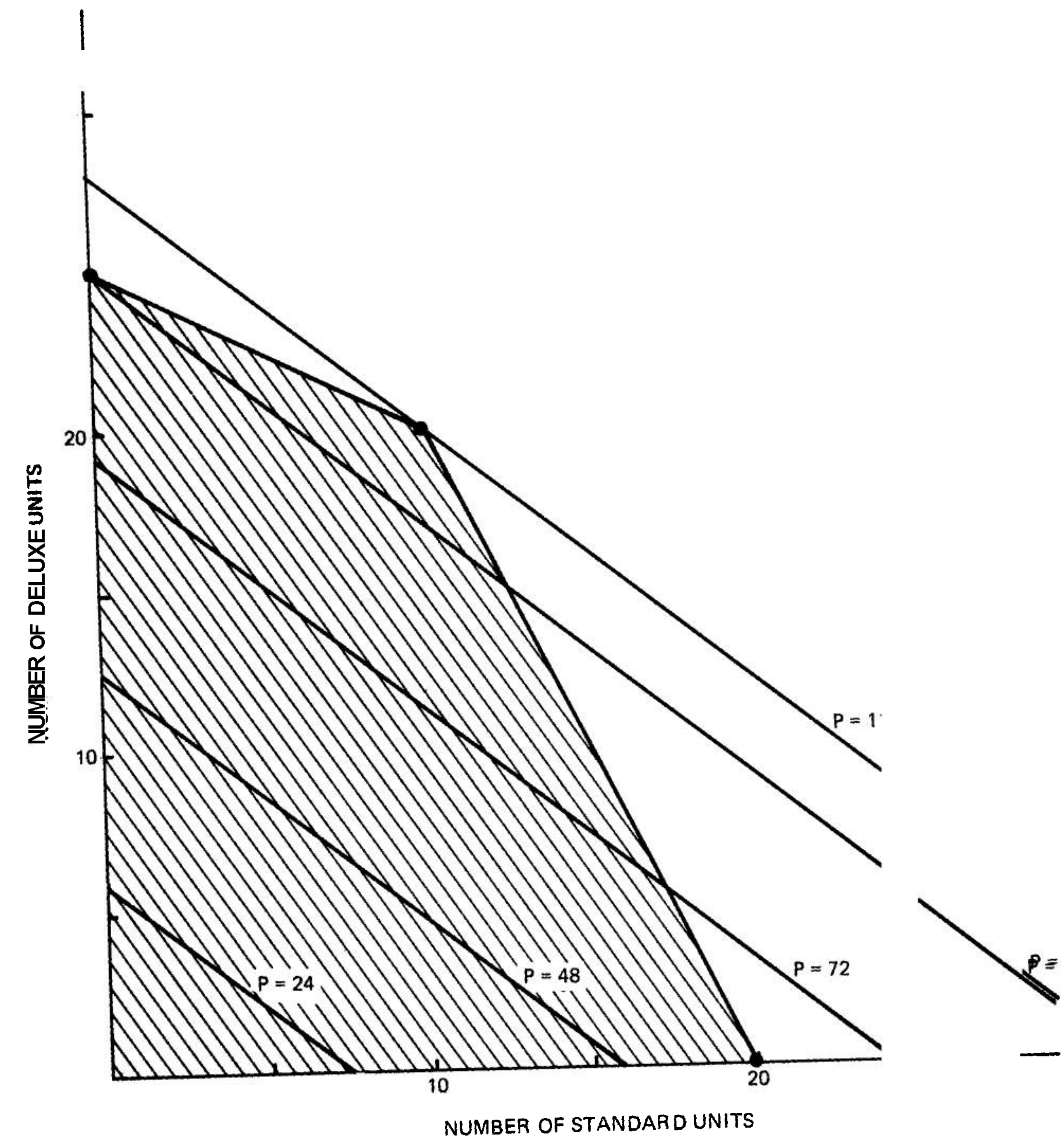


Figure 2-17. Equal [illegible]

maximum occurs at a vertex of the feasible region. Additionally, the same conclusion would obtain if the minimum of a linear objective function were sought.

2-5.3.2 Preliminaries to the Simplex Method

The preceding example has served to highlight the important point that in seeking the optimum value of the objective function in a linear programming problem, one need only compare the value of the objective function at the vertices of the feasible region. This reduces the set of points to be examined from an infinite to a finite set. However, this finite set may contain a large number of vertices, and it could be a considerable amount of work to find the coordinates of each vertex and to evaluate the objective function at each. Thus, a systematic procedure to find the maximum value of the objective function, without completely enumerating all vertices and evaluating the objective function at each, one, is highly desirable. Fortunately, such a procedure exists. It is called the simplex method and we will describe it in the next paragraph. First, we want to define some variables that, in addition to the decision variables, play a role in linear programming problems. These are the so-called "slack variables".

Consider the constraints in the general formulation of the linear programming problem. (See Eq. 2-113.) These less-than-or-equal-to constraints can be converted to equality constraints by adding to the left-hand side of each inequality the difference (the "slack") between the right-hand side and it. Denote by x_{n+i} the slack variable introduced into the ith constraint. Then the set of the formulas of Eq. 2-113 can be rewritten as

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + x_{n+1} &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + x_{n+2} &= b_2 \\
\cdots & \qquad (2\text{-}116) \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + x_{n+m} &= b_m
\end{aligned}
$$

This ith slack variable x_{n+i} can be interpreted as the amount of the ith resource which goes unallocated. By their definition, the slack variables are non-negative. Thus, the general linear programming problem can be reformulated as:

a. maximize $P = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n$

subject to the set of Eqs. 2-116 and with

b. $x_1 \geq 0, x_2 \geq 0, \ldots, x_n \geq 0,$
$x_{n+1} \geq 0, \ldots, x_{n+m} \geq 0.$

Note that the slack variables do not enter the objective function. However, it is convenient to think of them as being in the objective function with zero coefficients. The coefficients appearing in this formulation of the linear programming problem play an important role in the simplex method. Particular attention is paid to the n-component row vector $\vec{c} = (c_1, c_2, \ldots, c_n)$ of coefficients in the objective function and to the m-component column vectors of coefficients

$$
\vec{a}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ \cdot \\ a_{m1} \end{bmatrix}, \quad
\vec{a}_2 = \begin{bmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ \cdot \\ a_{m2} \end{bmatrix}, \ldots, \quad
\vec{a}_n = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ \cdot \\ a_{mn} \end{bmatrix}
$$

$$
\vec{a}_{n+1} = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \quad
\vec{a}_{n+2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \ldots, \qquad (2\text{-}117)
$$

$$
\vec{a}_{n+m} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix}, \quad
\vec{b} = \begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix}
$$

which appear in the constraint equations. Column vectors such as $\vec{a}_{n+1}, \vec{a}_{n+2}, \ldots, \vec{a}_{n+m}$ that have all entries equal to zero except for a single entry of one, are called unit vectors. A set of m unit vectors, each having its nonzero entry (i.e., a one) in a different location than all the others, is called a *basis*. Note that the column vectors of coefficients of the slack variables constitute a basis.

TABLE 2-2.
SIMPLEX TABLEAU

				c_1	. . .	c_n	0	. . . .	0
Row	Vectors $\vec{B}$ in Basis	$\vec{c}_B$	$\vec{b}$	$\vec{a}_1$	. . .	$\vec{a}_n$	$\vec{a}_{n+1}$		$\vec{a}_{n+m}$
1	B_1	c_{B1}	y_{10}	y_{11}		y_{1n}	$y_{1,n+1}$		$y_{1,n+m}$
2	B_2	c_{B2}	y_{20}	y_{21}		y_{2n}	$y_{2,n+1}$		$y_{2,n+m}$
•									
•									
•									
m	B_m	c_{Bm}	y_{m0}	y_{m1}		y_{mn}	$y_{m,n+1}$		$y_{m,n+m}$
m + 1			$z_0 = y_{m+1,0}$	$z_1 - c_1 = y_{m+1,1}$		$z_n - c_n = y_{m+1,n}$	$z_{n+1} = y_{m+1,n+1}$		$z_{n+m} = y_{m+1,n+m}$

2-5.3.3 The Simplex Method

The simplex method of solving linear programming problems is an iterative technique. Each iteration examines the value of the objective function at a vertex of the feasible region. The computations at each iteration indicate if the optimum solution has been reached and, if so, what the maximum value of the objective function is and for what values of the decision variables and slack variables this optimum is attained. If the optimum solution has not been reached, the simplex method will, at the next iteration, examine for optimality an adjacent vertex at which the value of the objective function is at least as large as it is at the present iteration. The process will always terminate in a finite number of steps. (There is a circumstance in which the standard simplex procedure must be adjusted to assure this finite termination. However, this situation-called degeneracy—is so rare, we will not discuss it in this handbook. Those interested in pursuing the point may consult, e.g., Hadley (Ref. 36).)

The computations at each stage of the simplex process can conveniently be arranged in a tabular form called a ***simplex tableau*** as shown in Table 2-2.

The quantities z_i, $i \geq 1$, appearing in row $m + 1$ of the tableau are the inner product of the $\vec{c}_B$ column with the $\vec{a}_i$ column. (By the inner product of two columns is meant the sum of products of corresponding entries in the two columns.) The term z_o is the inner product of the $\vec{c}_B$ column with the b column. The entries in the c_B column are the coefficients from the objective function corresponding to vectors in the basis. The vectors in the basis, $\vec{B}_1, \ldots, \vec{B}_m$, are a subset of the vectors $\vec{a}_1, \ldots, \vec{a}_{n+m}$.

The basis vectors $\vec{B}_1, \ldots, \vec{B}_m$ determine a vertex of the feasible region; z_o is the value of the objective function at that vertex. If all the quantities in row $m + 1$ to the right of the double line are non-negative, then the optimal solution has been found and the maximum value of the objective function is z_o. The solution $(x^*_1, \ldots, x^*_{n+m})$ is read from the $\vec{b}$ column; i.e., if the vector $\vec{a}_i$ is in the basis, then x^*_i is the corresponding entry in the $\vec{b}$ column. If the vector $\vec{a}_i$ is not in the basis, then $x^*_i = 0$.

If at least one entry in row $m + 1$ to the right of the double line is negative, then another vertex of the feasible region must be examined for optimality. In moving to another vertex, the simplex method changes one vector in the basis in such a manner that the value of the objective function at the new vertex is at least as large as at the previous vertex. Changing one vector in the basis involves two decisions: which new vector will enter the basis, and which vector currently in the basis will leave.

The entering vector is determined by considering those vectors $\vec{a}_j$ for which $z_j - c_j < 0$. The vector $\vec{a}_k$ will enter the basis if

$$z_k - c_k = \min_j (z_j - c_j), z_j - c_j < 0.$$

Simply stated, one chooses as the entering vector the one for which the corresponding quantity $z_k - c_k$ is the most negative of all the $z_j - c_j$ terms. If there is a tie among two or more vectors to enter, any one of them may be chosen. One possible rule is to choose as entering vector the one with the lowest index (i.e., subscript j) among those eligible.

Having determined the vector to enter the basis, one must next determine the vector to leave the basis. This is accomplished by considering the ratios of terms in the $\vec{b}$ column to corresponding *positive* terms in the $\vec{a}_k$ column (the column of the entering vector). The leaving vector, the one in row r, is the one for which this ratio is least, i.e., the vector in row r is removed from the basis if

$$\frac{y_{r0}}{y_{rk}} = \min_i \left[\frac{y_{i0}}{y_{ik}},\ y_{ik} > 0\right] \tag{2-118}$$

It is possible that there is no positive entry in column k (i.e., $y_{ik} \le 0$ for $i = 1, \ldots, m$). In that case there is an unbounded solution to the linear programming problem; i.e., the objective function can be made arbitrarily large. While this is a mathematical possibility, it is seldom the case for a real-life, nontextbook problem and, should this situation arise, one should check carefully the formulation of the problem to make sure it is correct.

Assuming that at least one $y_{ik} > 0$, it is possible that there is a tie for the leaving vector. Again, one may choose arbitrarily among the eligible vectors to determine a leaving vector.

Having chosen a vector to enter the basis and another vector to leave the basis, it becomes necessary to transform the entries in the tableau. We will use a prime (′) to designate the new entries; unprimed letters refer to the old tableau. It is convenient to refer to the entry y_{rk} as the pivot element. This is the entry in column of the entering vector (column k) and the row of the leaving vector (row r). The transformation equations are for $j = 0, 1, \ldots, m + n$,

$$y'_{rj} = y_{rj}/y_{rk} \tag{2-119}$$

$$y'_{ij} = y_{ij} - \frac{y_{ik}}{y_{rk}} y_{rj},\quad i \neq r,\ i = 1, \ldots, m + 1 \tag{2-120}$$

The first equation states that all old entries in the row in which the pivot element appears are divided by the pivot element to obtain the new entries in that row. The second equation states that old entries in any nonpivot row are replaced by the existing entry minus a certain multiple of the corresponding element in the pivot row. This multiple is the ratio of the entry in the row being transformed and the pivot column to the pivot element. This transformation process will be illustrated when we carry through the solution of a linear programming problem by the simplex method. Before we do this, however, we must indicate the form that the initial tableau takes. It will be described by giving values to the entries in Table 2-2. They are:

$$\begin{aligned}
&\vec{B}_1 = \vec{a}_{n+1},\ \vec{B}_2 = \vec{a}_{n+2}, \ldots,\\
&\vec{B}_m = \vec{a}_{n+m};\\
&c_{B1} = 0,\ c_{B2} = 0, \ldots, c_{Bm} = 0;\\
&y_{10} = b_1,\ y_{20} = b_2, \ldots,\\
&y_{m0} = b_m,\ y_{m+1,0} = 0;\\
&y_{11} = a_{11},\ y_{21} = a_{21}, \ldots,\\
&y_{m1} = a_{m1},\ y_{m+1,1} = -c_1;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&y_{1n} = a_{1n},\ y_{2n} = a_{2n}, \ldots,\\
&y_{mn} = a_{mn},\\
&y_{m+1,n} = -c_n;\\
&y_{1,n+1} = 1,\ y_{2,n+1} = 0, \ldots,\\
&y_{m,n+1} = 0,\ y_{m+1,n+1} = 0;\\
&\ldots\ldots\ldots\ldots\ldots\ldots\\
&y_{1,n+m} = 0,\ y_{2,n+m} = 0, \ldots,\\
&y_{m,n+m} = 1,\ y_{m+1,n+m} = 0.
\end{aligned}$$

The simplex calculations will be illustrated through the example that was previously considered from a geometrical point of view. Upon introducing slack variables, the problem is stated as:

a. maximize $3x_1 + 4x_2$

b. subject to:

$$\begin{aligned}
4x_1 + 2x_2 + x_3 &= 80\\
2x_1 + 5x_2 + x_x &= 120\\
x_1 \ge 0, x_2 \ge 0, x_3 &\ge 0, x_4 \ge 0.
\end{aligned}$$

We arrange the coefficients into an initial simplex tableau as shown in Table 2-3. The most negative entry in the last row, -4, is in the $\vec{a}_2$ column; therefore $\vec{a}_2$ will enter the basis, i.e., k = 2. To determine the leaving vector, we consider the ratios (80/2) = 40 and (120/5) = 24 and remove the vector corresponding to the minimum of these ratios. This vector is $\vec{a}_4$ which is in row 2; i.e., $r = 2$. Thus the new basis will be $\vec{a}_3$, $\vec{a}_2$ and these vectors will appear, in that order, in the "Basis" column of the second tableau (*see* Table 2-4). The corresponding coefficients from the objective function, 0 and 4, will appear in the $\vec{c}_B$ column.

The pivot element y_{22} (the entry in the column of the entering vector and the row of the leaving vector) is 5.

TABLE 2-3.
INITIAL SIMPLEX TABLEAU

				3	4	0	0
Row	Basis	$\vec{c}_B$	$\vec{b}$	$\vec{a}_1$	$\vec{a}_2$	$\vec{a}_3$	$\vec{a}_4$
1	$\vec{a}_3$	0	80	4	2	1	0
2	$\vec{a}_4$	0	120	2	5	0	1
3			0	-3	-4	0	0

TABLE 2-4.
SECOND TABLEAU

				3	4	0	0
Row	Basis	$\vec{c}_B$	$\vec{b}$	$\vec{a}_1$	$\vec{a}_2$	$\vec{a}_3$	$\vec{a}_4$
1	$\vec{a}_3$	0	32	16/5	0	1	-2/5
2	$\vec{a}_2$	4	24	2/5	1	0	1/5
3			96	-7/5	0	0	4/5

Thus, the elements in the new row 2 will be equal to the elements in the old row 2 divided by 5. Stated algebraically,

$$y'_{2j} = y_{2j}/y_{22} = y_{2j}/5, \; j = 0, \ldots, 4 \tag{2-121}$$

The new row 1 elements are given by:

$$y'_{1j} = y_{1j} - \left(\frac{y_{ik}}{y_{rk}}\right) y_{rj} = y_{1j} - \left(\frac{y_{12}}{y_{22}}\right) y_{2j}$$

$$= y_{1j} - \left(\frac{2}{5}\right) y_{2j}, \; j = 0, \ldots, 4 \tag{2-122}$$

Similarly, the new row 3 elements are given by:

$$y'_{3j} = y_{3j} - \left(\frac{y_{3k}}{y_{rk}}\right) y_{rj} = y_{3j} - \left(\frac{y_{32}}{y_{22}}\right) y_{2j}$$

$$= y_{3j} - \left(\frac{-4}{5}\right) y_{2j}$$

$$= y_{3j} + \left(\frac{4}{5}\right) y_{2j}, \; j = 0, \ldots, 4 \tag{2-123}$$

Specifically, we have:

$$y'_{10} = 80 - \left(\frac{2}{5}\right)120 = 32, \quad y'_{11} = 4 - \left(\frac{2}{5}\right)2 = 16/5,$$
$$y'_{12} = 2 - \left(\frac{2}{5}\right)5 = 0, \quad y'_{13} = 1 - \left(\frac{2}{5}\right)0 = 1$$
$$y'_{14} = 0 - \left(\frac{2}{5}\right)1 = -2/5;$$
$$y'_{20} = 120/5 = 24, \quad y'_{21} = 2/5, \qquad \text{(2-124)}$$
$$y'_{22} = 5/5 = 1, \quad y'_{23} = 0/5 = 0,$$
$$Y'_{24} = 1/5;$$
$$y'_{30} = 0 + \left(\frac{4}{5}\right)120 = 96, \quad y'_{31} = -3 + \left(\frac{4}{5}\right)2 = -7/5,$$
$$y'_{32} = -4 + \left(\frac{4}{5}\right)5 = 0, \quad y'_{33} = 0 + \left(\frac{4}{5}\right)0 = 0,$$
$$y'_{34} = 0 + \left(\frac{4}{5}\right)1 = 4/5.$$

The second tableau, then, is as shown in Table **2-4.**

There is yet a negative entry in the last row to the right of the double line in the second tableau. Since there is only one such, the corresponding vector $\vec{a}_1$, enters the basis. Considering the ratios (32/16/5) = 10 and **(24/2/5)** = 60, we remove the vector corresponding to their minimum. This is $\vec{a}_3$ which is in row 1; i.e., $r = 1$. Thus the new basis will be $\vec{a}_1$, $\vec{a}_2$ and the vectors will appear, in that order, in the "Basis" column of the third tableau. The corresponding coefficients from the objective function, **3** and **4,** will appear in the $\vec{c}_B$ column. The pivot element y_{11} (the entry in the column of the entering vector and the row of the leaving vector) is 16/5. The new row 1 entries, then, are (5/16) times the old row 1 entries.

In equation form:

$$y'_{1j} = y_{1j}\Big/\frac{16}{5} = \left(\frac{5}{16}\right)y_{1j}, \quad j = 0, \ldots, 4 \qquad \text{(2-125)}$$

(Note that the new second tableau entries are again denoted with a prime mark. To be completely consistent, we ought perhaps to use a double prime mark, but this is really unnecessary as no confusion will result from our practice.)

The new row 2 entries are given by

$$y'_{2j} = y_{2j} - \left(\frac{2/5}{16/5}\right)y_{1j} = y_{2j} - \left(\frac{1}{8}\right)y_{1j},$$
$$j = 0, \ldots, 4 \qquad \text{(2-126)}$$

The new row **3** entries are given by

$$y'_{3j} = y_{3j} - \left(\frac{-7/5}{16/5}\right)y_{1j} = y_{3j} + \left(\frac{7}{16}\right)y_{1j},$$
$$j = 0, \ldots, 4 \qquad \text{(2-127)}$$

Numerically, we obtain

$$y'_{01} = \left(\frac{5}{16}\right)32 = 10, \quad y'_{11} = \left(\frac{5}{16}\right)\left(\frac{16}{5}\right) = 1,$$
$$Y'_{12} = \left(\frac{5}{16}\right)0 = 0, \quad y'_{13} = \frac{5}{16},$$
$$y'_{14} = \left(\frac{5}{16}\right)\left(-\frac{2}{5}\right) = -\frac{1}{8};$$
$$y'_{20} = 24 - \left(\frac{1}{8}\right)32 = 20, \quad y'_{21} = \frac{2}{5} - \left(\frac{1}{8}\right)\left(\frac{16}{5}\right) = 0,$$
$$y'_{22} = 1 - \left(\frac{1}{8}\right)0 = 1, \quad y'_{23} = 0 - \left(\frac{1}{8}\right)1 = -\frac{1}{8}, \qquad \text{(2-128)}$$
$$y'_{24} = \frac{1}{5} - \left(\frac{1}{8}\right)\left(-\frac{2}{5}\right) = \frac{1}{4};$$
$$y'_{30} = 96 + \left(\frac{7}{16}\right)32 = 110, \quad y'_{31} = -\frac{7}{5} + \left(\frac{7}{16}\right)\left(\frac{18}{5}\right) = 0,$$
$$y'_{32} = 0 + \left(\frac{7}{16}\right)0 = 0, \quad Y'_{33} = 0 + \left(\frac{7}{16}\right)1 = \frac{7}{16},$$
$$y'_{34} = \frac{4}{5} + \left(\frac{7}{16}\right)\left(-\frac{2}{5}\right) = \frac{5}{8}.$$

The third tableau, then, is as shown in Table **2-5.** Since all the entries in the last row to the right of the double line in the third tableau are non-negative, we

TABLE 2-5.
THIRD TABLEAU

Row	Basis	$\vec{c}_B$	$\vec{b}$	$\vec{a}_1$ (3)	$\vec{a}_2$ (4)	$\vec{a}_3$ (0)	$\vec{a}_4$ (0)
1	$\vec{a}_1$	3	10	1	0	5/16	-1/8
2	$\vec{a}_2$	4	20	0	1	-1/8	1/4
3			110	0	0	7116	518

have found the optimal solution. The optimal value of the objective function is 110 with the solution being $x_1 = 10, x_2 = 20, x_3 = 0, x_4 = 0$. This agrees, of course, with the solution we found earlier through a geometrical argument. However, the simplex method illustrated here can solve a linear programming problem involving any number of decision variables, while the geometrical method is usable only for a linear programming problem involving two decision variables.

2-5.3.4 Other Linear Programming Formulations

We must next look at variations of the linear programming problem we stated at the outset in Eqs. 2-113:

Variation 1. Suppose the objective function

$$Z = c_1 x_1 + c_2 x_2 + \ldots + c_n x_n$$

is to be minimized rather than maximized. Then one maximizes

$$-Z = -c_1 x_1 - c_2 x_2 - \ldots - c_n x_n,$$

and having found the maximum of $-Z$, takes its negative; i.e., min $Z = -\max(-Z)$. This fact is easily verified. Let x^* be the point in the feasible region at which the objective function takes on its minimum, and let Z^* denote that minimum value. Then, for any other value Z of the objective function, $Z \geq Z^*$. Multiplying both sides of this inequality by (-1) will change the sense of the inequality to yield $(-Z) \leq (-Z^*)$, i.e., $(-Z^*)$ is the maximum value of the negative of the objective function. We conclude that

$$\min Z = Z^* = -(-Z^*) = -[\max(-Z)] \tag{2-129}$$

Variation 2. Suppose one or more of the inequality signs in the constraints are of the $\geq$ form, instead of $\leq$. For example, the first constraint might read

$$a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n \geqslant b_1.$$

One first converts this inequality to an equality by subtracting from the left-hand side, x_{n+1}, the amount by which the left-hand side exceeds the right-hand side. (The quantity x_{n+1} is non-negative and is called a ***surplus variable.***) This yields

$$a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n - x_{n+1} = b_1.$$

Whereas in the situation in which a slack variable was added to the left-hand side to convert a $\leq$ inequality to an equality and, thereby, the unit element of a basis vector was created, this is not the case when a surplus variable is introduced. The coefficient of the surplus variable is -1, rather than the $+1$ needed for a basis vector. The $+1$ coefficient is achieved by adding in another variable, called an ***artificial variable,*** to the equation formed by the introduction of the slack variable. Denoting the artificial variable by x_{n+2}, the equation now reads

$$a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n - x_{n+1} + x_{n+2} = b_1.$$

Since equality existed prior to introducing x_{n+2}, it is clear that x_{n+2} must equal zero. This is of no concern as we only want the ***coefficient*** of x_{n+2} not x_{n+2} itself. However, we must assure ourselves that in the solution to the linear programming problem in which x_{n+2} appears, any optimal solution will have

$x_{n+2} = 0$. This is accomplished by changing the objective from:

maximize $c_1 x_1 + c_2 x_2 + \ldots + c_n x_n$

to:

maximize $c_1 x_1 + c_2 x_2 + \ldots + c_n x_n - Mx_{n+2}$,

where M is a very large positive number. If x_{n+2} were positive, the objective function would be smaller than it would be if x_{n+2} were zero. (Note that x_{n+2}, being a variable in a linear programming problem, must be greater than or equal to zero; it cannot be negative.) Thus, x_{n+2} is forced to be zero and the optimal solution is unaltered from what it was prior to the introduction of the artificial variable.

We will illustrate the preceding discussion with an example. Consider the linear programming problem:

a. minimize $Z = 5x_1 + 2x_2$

b. subject to:

$$x_1 + x_2 \leq 4$$
$$x_1 - x_2 \geq 2$$
$$x_1 \geq 0, x_2 \geq 0.$$

Introducing a slack variable x_3, a surplus variable x_4, and an artificial variable x_5 into the constraints, and subtracting Mx_5 from the objective function, the problem is transformed to:

a. maximize $5x_1 + 2x_2 - Mx_5$

b. subject to:

$$x_1 + x_2 + x_3 = 4$$
$$x_1 - x_2 - x_4 + x_5 = 2$$
$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0.$$

To solve this problem by the simplex method, we begin by setting up the initial tableau as shown in Table 2-6. Recalling that M in the initial tableau is a large positive number, we see that the most negative entry to the right of the double line is $-M - 5$. Thus the vector $\vec{a}_1$ will enter the basis. To determine the leaving vector, we consider the ratios 4/1 and 2/1 and remove the vector corresponding to the minimum ratio. The vector to be removed is $\vec{a}_5$.

It is a fact that once a vector of coefficients of an artificial variable (called an artificial vector) is removed from a basis, it will never again enter the basis. Thus, once such a vector leaves the basis, its column may be deleted from subsequent simplex tableaus. Since $\vec{a}_5$ is an artificial vector, it will not enter hereafter.

The transformation equations to the second tableau are

$$x'_{2j} = x_{2j}, \; j = 0, \ldots, 4 \qquad (2\text{-}130)$$

$$x'_{1j} = x_{1j} - x_{2j}, \; j = 0, \ldots, 4 \qquad (2\text{-}131)$$

$$x'_{3j} = x_{3j} + (M + 5)x_{2j}, \; j = 0, \ldots, 4 \qquad (2\text{-}132)$$

and the second tableau is, then, as shown in Table 2-7.

The most negative entry in the bottom row and to the right of the double line in the second tableau is -7, belonging to vector $\vec{a}_2$. This next enters the basis. The choice of the leaving vector is easy in this case as only one entry in the $\vec{a}_2$ column is positive and therefore the vector in its row, $\vec{a}_3$, is the only vector eligible for removal from the basis. The transformation equations to the third tableau are:

$$x'_{1j} = \frac{1}{2} x_{1j}, \; j = 0, \ldots, 4 \qquad (2\text{-}133)$$

$$x'_{2j} = x_{2j} + \frac{1}{2} x_{1j}, \; j = 0, \ldots, 4 \qquad (2\text{-}134)$$

$$x'_{3j} = x_{3j} + \frac{7}{2} x_{1j}, \; j = 0, \ldots, 4 \qquad (2\text{-}135)$$

The third tableau is, then, as shown in Table 2-8.

There is still a negative entry in the last row to the right of the double line in the third tableau. It is in the $\vec{a}_4$ column, so $\vec{a}_4$ is the entering vector. There is only one positive element in the $\vec{a}_4$ column and it is in the a_2 row, so $\vec{a}_2$ leaves the basis. The transformation equations to the fourth tableau are:

$$x'_{1j} = 2x_{1j}, \; j = 0, \ldots, 4 \qquad (2\text{-}136)$$

$$x'_{2j} = x_{2j} + x_{1j}, \; j = 0, \ldots, 4 \qquad (2\text{-}137)$$

$$x'_{3j} = x_{3j} + 3x_{1j}, \; j = 0, \ldots, 4 \qquad (2\text{-}138)$$

The fourth tableau is, then, as shown in Table 2-9. Since all final row entries to the right of the double line in the fourth tableau are non-negative, the optimal solution has been attained, namely a maximum value of 20 for the objective function which is achieved for $x_1 = 4$, $x_2 = 0$, $x_3 = 0$, $x_4 = 2$. Additionally, of course, $x_5 = 0$. However, one generally does not cite the fact that the artificial variables are zero (after all, that is the value they must have!), while one does give

TABLE 2-6.
INITIAL TABLEAU

Row	Basis	$\vec{c}_B$	$\vec{b}$	$\vec{a}_1$	$\vec{a}_2$	$\vec{a}_3$	$\vec{a}_4$	$\vec{a}_5$
				5	2	0	0	-M
1	$\vec{a}_3$	0	4	1	1	1	0	0
2	$\vec{a}_5$	-M	2	1	-1	0	-1	1
3			-2M	-M -5	M -2	0	M	0

TABLE 2-7.
SECOND TABLEAU

Row	Basis		$\vec{b}$	$\vec{a}_1$	a_2	$\vec{a}_3$	a_4
				5	2	0	0
1	$\vec{a}_3$	0	2	0	2	1	1
2	$\vec{a}_1$	5	2	1	-1	0	-1
3			10	0	-7	0	-5

the values of the slack and surplus variables, as well as the decision variables, in an optimal solution.

Variation 3. Another variation which can arise of the linear programming problem first described is that one or more of the b_i's may be negative. This can easily be remedied by multiplying both sides of the constraint by (− 1). This will change the sense of the inequality between the two sides in the constraint. If the inequality was initially 2, then it will become ≤ and the inclusion of a slack variable is called for. If the inequality was initially ■ then it will become ≥ and a surplus variable and an artificial variable are needed. The artificial variable must also enter the objective function, as in Variation 2 previously discussed.

Variation **4.** Another variation of the originally described linear programming problem is where one or more of the constraints appears as an equality. Again, by the inclusion of an artificial variable in each equality, one creates the necessary number of basis vectors. Each artificial variable must enter the objective function with a coefficient of $-M$, with M a large, positive number.

Variation **5.** There is one final item that will be mentioned here regarding the solution of a linear programming problem. We have written all the preceding material as though there were a unique solution. This may not always be true. Two other possibilities exist: (a) no feasible solution, or (b) multiple solutions.

Case (a) can be disposed of fairly readily by the simple expedient of checking any purported solution to make sure all constraints are satisfied. If it appears that a linear programming problem has no feasible solution, then it may be wise to re-examine the problem formulation for correctness.

The Occurrence of case (b) can be detected from the final simplex tableau. If, when optimality has been reached, there is a zero entry in the last row of a column not represented in the basis, then multiple solutions exist. Another vertex, at which the value of the objective function is the same as at the present vertex, can be found by entering into the basis any vector $\vec{a_j}$ not in the basis, but with $z_j - c_j = 0$. Any point on the line segment joining these two vertices will also be optimal. If this occurs, the problem is said to have ***alternative optima.***

Alternative optima can be illustrated by reference to the simple example given in par. **2-5.3.1.** If the problem had been to maximize $P = 2x_1 + 5x_2$ subject to:

$$4x_1 + 2x_2 \leq 80$$
$$2x_1 + 5x_2 \leq 120$$
$$x_1 \ 2 \ 0, x_2 \ 2 \ 0,$$

then the optimal value of P would be **120** and this would be achieved at any point on the line segment with end points **(0,24)** and **(10,20).** This can be seen geometrically by observing that the line $2x_1 + 5x_2 = P$ is parallel to the polisher constraint line and that when P is increased as much as possible so that the profit line lies inside or on the feasible region, it will coincide with the polisher constraint line.

TABLE 2-8.
THIRD TABLEAU

				5	2	0	0
Row	Basis	$\vec{c}_B$	$\vec{b}$	$\vec{a}_1$	$\vec{a}_2$	$\vec{a}_3$	→
1	$\vec{a}_2$	2	1	0	1	112	112
2	$\vec{a}_1$	5	3	1	0	1/2	-112
3			17	0	0	7/2	-312

TABLE 2-9.
FOURTH TABLEAU

				5	2	0	0
Row	Basis	$\vec{c}_B$	$\vec{b}$	$\vec{a}_1$	$\vec{a}_2$	$\vec{a}_3$	$\vec{a}_4$
1	$\vec{a}_4$	0	2	0	2	1	1
2	$\vec{a}_1$	5	4	1	1	1	0
3			20	0	3	5	0

The purpose of these discussions has been to define what constitutes a linear programming problem; to motivate, by an example, the solution process; and to illustrate the solution of a linear programming problem by the simplex method. The problems we examined were small scale, i.e., involved only few decision variables and few constraints. Real-life linear programming problems can involve hundreds of decision variables and constraints. For problems of such magnitude hand-calculated simplex solutions are, of course, out of the question. Fortunately there are efficient, accurate computer programs available to handle such problems.

2-5.3.5 Maintenance Applications

We close by referring to some applications of linear programming to the determination of optimal maintenance policies. The models are rather complicated and require considerable preliminaries even to state. Thus we shall merely refer the interested reader to some sources. One such is Section 5, Chapter 5 "Optimal Maintenance Policies Under Markovian Deterioration" in the book by Barlow and Proschan (Ref. 25). Another is the book by Derman (Ref. 37).

ICES

1. TM 38-703-1, ***Integrated Logistic Support: Support Integration,*** Dept. of the Army, November 1969.
2. AFSC-TR-65-6, ***Chairman's Final Report,*** Weapon System Effectiveness Industry Advisory Committee (WSEIAC), Air Force Systems Command, January 1965 (AD-467 816).
3. AMCP 706-134, Engineering Design Handbook, ***Maintainability Guide for Design.***
4. W. H. Von Alven, Ed., ***Reliability Engineering,*** Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964.
5. ***Systems Effectiveness,*** Systems Effectiveness Branch, Office of Naval Material, Washington, D.C., 1965 (AD-659 520).
6. NAVMAT P 3941 A, ***Navy Systems Performance Effectiveness Manual,*** Headquarters, Naval Material Command, 1 July 1968.
7. B. S. Blanchard, "Cost Effectiveness, System Effectiveness, Integrated Logistic Support, and Maintainability", ***IEEE Transactions*** on ***Reliability,*** R-16, No. 3, pp. 117-126, 1967.
8. B. S. Blanchard and E. E. Lowery, ***Maintainability,*** McGraw-Hill Book Company, Inc., N.Y., 1969.
9. A. S. Goldman and T. B. Slattery, ***Maintainability: A Major Element of System Effectiveness,*** John Wiley & Sons, Inc., N.Y., 1964.
10. NAVSHIPS 94324, ***Maintainability Design Criteria Handbook for Designers of Shipboard Electronic Equipment,*** Dept. of the Navy, Change 2, 1965.
11. MIL-STD-721, ***Definitions of Effectiveness Terms for Reliability, Maintainability, Human Factors, and Safety,*** 25 August 1966.
12. AMCR 11-1, ***Army Programs, Systems Analysis,*** 21 April 1970.
13. M. B. Kline and M. W. Lifson, "Systems Engineering", Chapter 2 in ***Cost-Effectiveness: The Economic Evaluation of Engineered Systems,*** J. M. English, Ed., John Wiley & Sons, Inc., N.Y., 1968.
14. M. B. Kline, ***Maintainability Considerations in System Design,*** Doctoral Dissertation, Department of Engineering, University of California Los Angeles, 1966; University Microfilms, Inc., Ann Arbor, Michigan.
15. AMCR 750-15, ***Integrated Logistics Support,*** 19 February 1969.
16. AMCP 706-191, Engineering Design Handbook, ***System Analysis and Cost-Effectiveness.***
17. ***Integration Engineering for System Effectiveness,*** EIA Effectiveness Quantification Bulletin No. 2, EIA Committee G47 with F. H. Kranz, Electronic Industries Association, Washington, D.C., March 1970.
18. Igor Bazovsky and Glen E. Benz, "Effectiveness of Ship Systems and Machinery", in ***Proceedings*** 1968 ***Annual Symposium on Reliability,*** Boston, Mass., January 1968.
19. P. M. Morse and G. E. Kimball, ***Methods of Operations Research,*** The M.I.T. Press, Cambridge, Mass., 1951.
20. Igor Bazovsky, ***Reliability Theory and Practice,*** Prentice-Hall, Inc., Englewood Cliffs, N.J., 1961.
21. D. R. Cox, ***Renewal Theory,*** John Wiley & Sons, Inc., N.Y., 1962.
22. C. B. Moore, ***An Integrated Approach to Determining Operational Effectiveness,*** Publication No. MR-0-128, General Dynamics, Fort Worth Division, March 1966.
23. A. J. Monroe, "WSEIAC Findings and Recommendations" in ***Proceedings of the First***

EIA Conference on Systems Effectiveness, Electronic Industries Association, Washington, D.C., October 1965.

24. I. Bazovsky, N. R. MacFarlane, and R. L. Wunderman, *Study of Maintenance Cost Optimization and Reliability of Shipboard Machinery,* United Control Corporation, Seattle, Wash., June 1962 (AD-283 428).
25. R. E. Barlow and F. Proschan, *Mathematical Theory of Reliability,* John Wiley & Sons, Inc., N.Y., 1965.
26. R. E. Barlow and L. C. Hunter, "Mathematical Model for System Reliability", *The Sylvania Technologist,* 13, Nos. 1 and 2, 1960.
27. J. E. Hosford, "Measures of Dependability", *Operations Research,* 8, No. 1, 1960.
28. B. A. Kozlov and I. A. Ushakov, *Reliability Handbook* Holt, Rinehart and Winston, Inc., N.Y., 1970.
29. *Mathematical Models for Availability of Machine Gun Systems,* Technical Report No. 3, Prepared by Igor Bazovsky and Associates, Inc. for the Small Arms Systems Laboratory, US Army Weapons Command, February 1970.
30. PAM 69-8, *Availability,* US Army Combat Developments Command Maintenance Agency, Aberdeen Proving Ground, Md., November 1970.
31. *Elements of Reliability and Maintainability,* DOD Joint Course Book, US Army Management Engineering Training Agency, Rock Island, Ill., 1967.
32. *Maintainability Engineering Guide,* Report RC-S-65-2, Revision A, US Army Missile Command, Redstone Arsenal, Ala., April 1967.
33. Gerald A. Sandler, *System Reliability Engineering,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.
34. *Maintainability Trade-off Techniques,* EIA Maintainability Bulletin No. 8, Electronic Industries Association, Washington, DC, July 1966.
35. Anonymous. *An Introduction to Linear Programming,* International Business Machines Corp., White Plains, N.Y., 1964.
36. G. Hadley, *Linear Programming,* Addison-Wesley Publishing Co., Inc., Reading, Mass., 1962.
37. Cyrus Derman, *Finite State Markovian Decision Process,* Academic Press, N.Y., 1970.

CHAPTER 3

MAINTAINABILITY ORGANIZATION AND MANAGEMENT

3-1 GENERAL

Where does the maintainability effort fit into the overall management organization? Should it be organized and managed within its own centralized structure? How does it interface with other organizational elements? Should it be combined with the reliability engineering or system effectiveness efforts? Is it part of the system engineering, design engineering, or integrated logistic support organization? What is its relationship to maintenance engineering? Should the structure of the contractor's organization reflect either the military customer's organization or what is perceived as the desires of the military customer? These are among the questions which concern maintainability engineers and management. It is the purpose of this chapter to examine these questions with regard to the maintainability engineering organization and its management.

Maintainability management can be discussed in several contexts. One is the management of the maintainability engineering function as an engineering discipline. A second is the organizational structure and relationships for carrying out the maintainability function. A third context has to do with the phase in the system life cycle of concern at the moment. This latter implies that there are dynamic (temporal) aspects of maintainability management and organization which may require a change of emphasis, if not of organization and responsibility, depending upon the stage and phase of the system life cycle in which the system design happens to be.

Since maintainability is defined (par. 1-1) as a **characteristic of design,** it follows that maintainability engineering is of primary concern and has its greatest impact during those phases of the system life cycle which are concerned with system and equipment design and test. Also, as its name implies, maintainability engineering belongs in the engineering (technical) organization.

3-1.1 ORGANIZATIONAL ACTIVITIES

In order to determine the organizational structure for the management of maintainability, one must first determine the functions and tasks performed by the maintainability organization. These activities may be classified into the following functions:

1. Management and administrative
2. Test and analytical
3. Design
4. Documentation
5. Coordination.

Each of these activities is described in the following paragraphs.

3-1.1.1 Maintainability Management and Administration

Maintainability management and administrative functions include those tasks concerned with performance, cost, and schedule and which give overall direction and control to the effective performance of the maintainability engineering effort aspect of program management (Ref. 1). These tasks include:

1. Preparing maintainability program plan—including milestones, schedules, and budgets—in accordance with specified program management requirements, system requirement specifications, and other management documents.

2. Preparing and issuing policies and procedures for use in the performance of the maintainability engineering function.

3. Participating in program management and design reviews which impact on maintainability.

4. Organizing and staffing the maintainability engineering effort.

5. Preparing budgets and schedules, and assigning responsibilities, tasks, and work orders for the maintainability effort.

6. Monitoring and controlling the output of the maintainability engineering organization.

7. Providing management liaison and coordination with higher level management, other related disciplines, and subcontractors.

8. Providing training and indoctrination with regard to maintainability.

9. Participating in industry/Government meetings and symposia with regard to maintainability management.

Since maintainability engineering is part of an interdisciplinary system engineering effort, the coordination and liaison aspects of maintainability with other disciplines are considerable if an optimal total system design is to be achieved. The coordination function, therefore, is listed separately in par. 3-1.1.5.

3-1.1.2 Maintainability Analysis

A significant portion of the maintainability engineering effort is concerned with the analytical aspects. These include maintainability requirements, predictions, allocations, demonstrations, and field data evaluations, as well as providing information for system engineering analyses and trade-offs.

Maintainability analysis tasks may include the following:

1. Reviewing operational and system requirement documents and specifications with regard to maintainability requirements.

2. Participating in system engineering analyses as they affect or are affected by maintainability.

3. Participating in or performing maintenance engineering analyses.

4. Performing maintainability predictions and allocations.

5. Assisting in preparation of maintainability demonstration plans and analysis of maintainability demonstration results.

6. Preparing maintainability demonstration reports.

7. Performing maintainability trade-off analyses within the maintainability engineering discipline.

8. Providing maintainability studies, data, and other information for system level trade-offs involving other disciplines, such as reliability or safety.

9. Assisting maintenance engineering in the performance of detailed maintainability studies, such as development of repair/discard criteria, level of automation studios, use of built-in test features, and maintainability skill level analyses.

10. Analyzing maintainability feedback data from the field and other sources.

11. Participating in statistical analyses with regard to maintainability and system effectiveness.

12. Participating in industry/Government meetings and symposia involving maintainability analysis.

Specific maintainability analysis techniques are discussed in Chapters 4, 6, 7, and 8.

3-1.1.3 Maintainability Design

Maintainability design is concerned with those system and equipment features and characteristics which will promote cost-effective ease of maintenance and thus will reduce logistic support requirements. Among the activities of concern are the following:

1. Monitoring and reviewing system/equipment designs with regard to maintainability features.

2. Participating in the preparation of maintainability engineering design criteria, guidelines, and handbooks for use by design engineers.

3. Providing consulting services to design engineers.

4. Reviewing and approving design drawings and data for maintainability features and compliance with specification requirements with regard to maintainability.

5. Participating in design reviews where maintainability is concerned.

6. Preparing maintainability design reports.

7. Participating in industry/Government meetings and symposia with regard to maintainability design.

Specific maintainability design characteristics and features are discussed in Chapter 5.

3-1.1.4 Maintainability Documentation

The maintainability engineering effort generates and utilizes a considerable amount of data and information. The effective and efficient handling of this information is important to the achievement of a cost-effective, coherent, total system design. Maintainability documentation includes:

1. Establishment and maintenance of a maintainability data bank and library of pertinent maintainability documents and information.

2. Preparation and maintenance of handbook data and information with regard to maintainability.

3. Preparation of maintainability data and feedback reports.

4. Documentation of maintainability trade-offs and the results of maintainability analyses.

5. Documentation of the results of maintainability design reviews.

6. Documentation of maintainability management information.

Maintainability data requirements are discussed in Chapter 9.

3-1.1.5 Maintainability Coordination

As pointed out in par. 3-1.1.1, a significant part of the maintainability management effort is concerned with coordination and liaison. The coordination effort is often one of the key elements in assuring a successful and. optimized system design. Maintainability coordination includes:

1. Interface with system engineering and other engineering disciplines, such as maintenance, design, reliability, safety, human factors, integrated logistic support, and system effectiveness.

2. Provision of maintainability training and indoctrination for all program personnel.

3. Subcontractor liaison and coordination as part of contractor responsibility, including training and indoctrination with regard to maintainability.

4. Maintainability liaison and coordination with the customer/contractor as directed by program management.

5. Liaison coordination with industry/Government advisory activities, including trade associations and professional societies.

3-1.2 ORGANIZATIONAL STRUCTURES FOR MAINTAINABILITY

Now that the maintainability management and engineering functions have been defined and described, we can address the organizational questions enumerated in par. 3-1. First, where does maintainability fit in the overall management organization? There is no unique or conventional organizational structure for maintainability. There are many versions and variations which are used by both customer and producer organizations. The structure used is often dependent upon the enterprise's overall organizational philosophy and method of doing business. There are, however, as discussed in par. 3-1.1, certain activities with which maintainability is concerned and which must be included within whatever organizational structure exists. Advanced planning and a recognition of these activities, as well as their relation to other engineering disciplines and organizational elements, will go a long way towards obtaining an efficient and effective total organization.

Considerations which should be carefully directed by top management are whether the maintainability function should be (1) an implicit rather than explicit part of the engineering organization, (2) a distinct line organizational element within the engineering department, (3) a staff function operating in an advisory capacity to project management and in an analytic and consultative capacity to designers, or (4) a part of program management or system engineering in a project or matrix organization. Such considerations are affected by the overall size of the enterprise and the project, the emphasis placed on maintainability by the customer in his system specifications, the extent to which the maintainability activities described in par. 3-1.1 are required and emphasized by both customer and producer, and the cost-effectiveness requirements of the particular project.

In small engineering organizations, maintainability tends to be an implicit part of the normal engineering design effort and is not treated analytically in any great detail.

3-1.2.1 Maintainability Engineering as a Centralized Functional Organization

The simplest explicit organizational structure places maintainability engineering as a distinct, functional, line organization within the overall engineering organization, as illustrated in Figs. 3-1 and 3-2. In this organizational form, all maintainability effort is centralized under a single manager. He has full responsibility for and control over all personnel and activities which are part of the maintainability discipline, as described in par. 3-1.1 and its subparagraphs. Such an organizational structure gives emphasis to maintainability as a design discipline. It is effective and efficient when managers and engineers recognize maintainability as a natural part of good engineering design.

The maintainability engineering manager is able to effect strong liaison and coordination with the interfacing disciplines mentioned in par. 3-1.1.5. He is able to control intra-maintainability trade-offs as well as the maintainability portions of system level trade-offs. In particular, he can maintain a proper interaction with maintenance engineering activities (par. 1-4.2).

Centralization of the maintainability engineering effort also works well when there is only one major project of concern or when there are a number of relatively small projects concerning basically similar products or customers (Ref. 2).

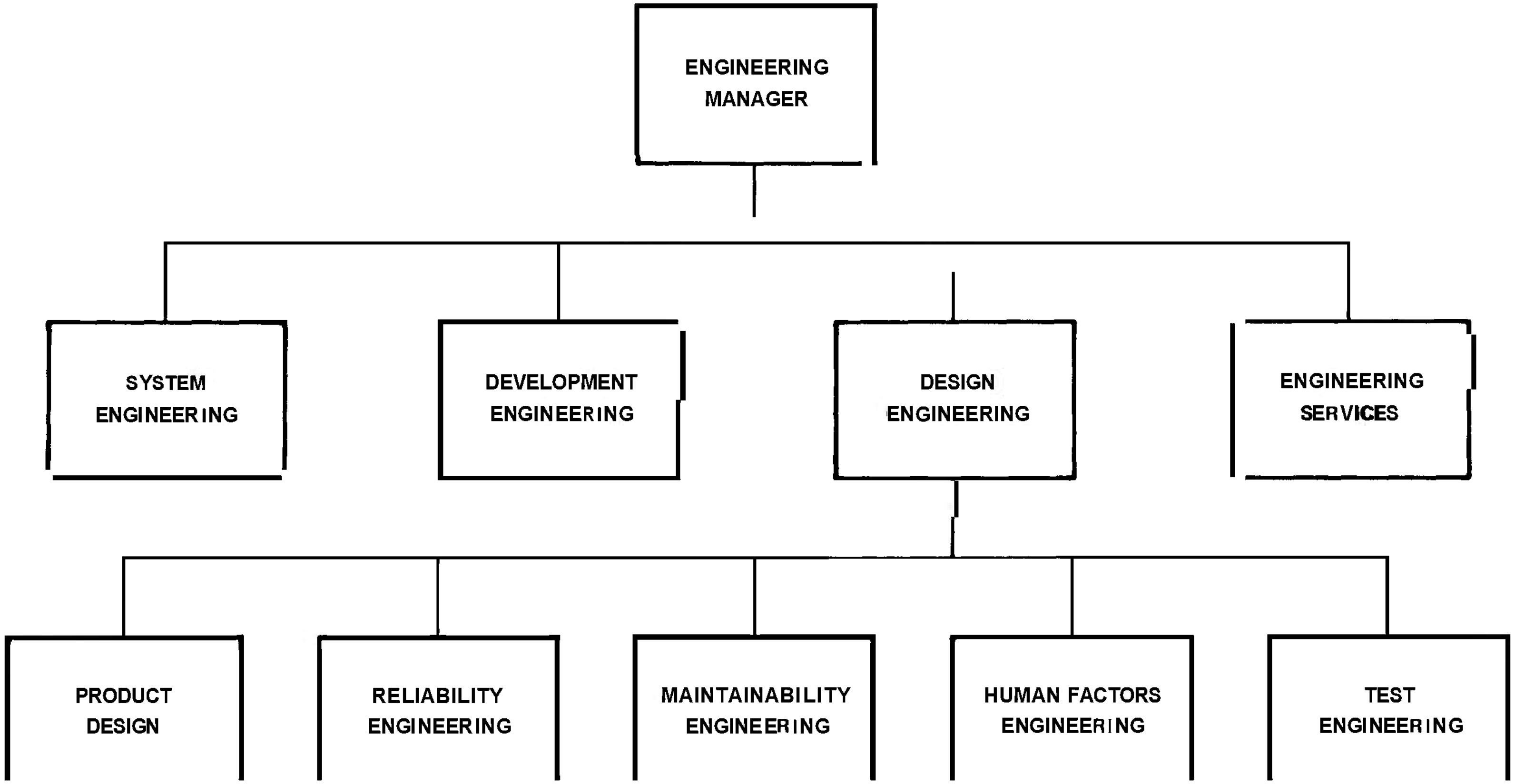


Figure 3-1. Maintainability Engineering as a Line Function

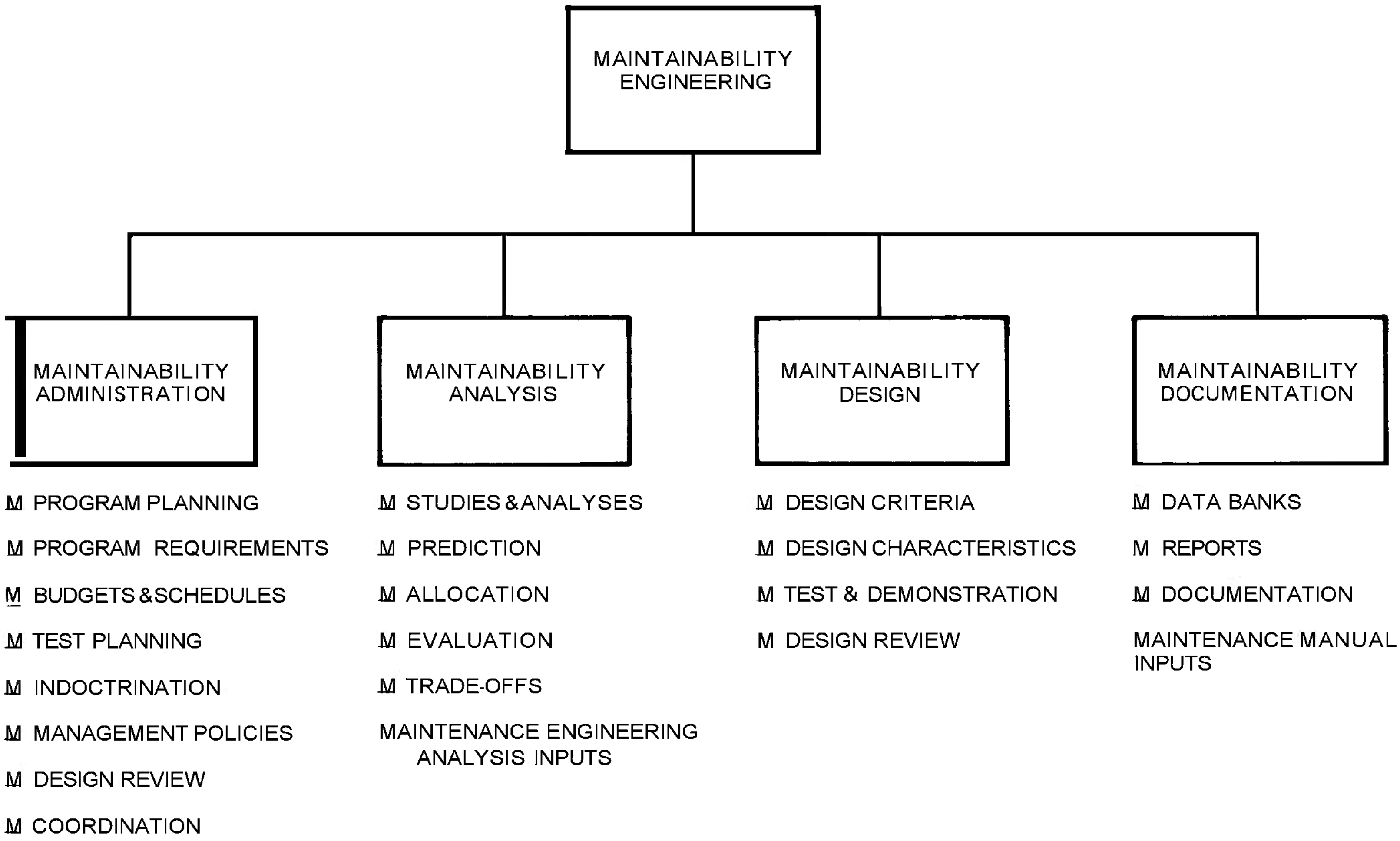


Figure 3-2. Centralized Maintainability Engineering Organization

3-1.2.2 Maintainability as a System Engineering Staff Function

In some organizations, maintainability is considered to be primarily a management staff and analytic function rather than a design function, having to do with calculating system and equipment *MTTR* and analyzing maintenance tasks and requirements in accordance with system effectiveness and system design requirement specifications. In such instances, maintainability engineering personnel are not usually design engineers, but instead are specialists who know the details of various military standards, specifications, and handbooks which deal with maintainability, or the applicable statistical and analytic techniques. They provide a staff function to the system engineers with regard to maintainability program requirements in the system specifications, perform maintainability prediction and allocations, establish maintainability demonstration requirements, and assist in system effectiveness calculations. They provide consulting services to the design engineers. A common practice in many organizations is to have maintainability and reliability engineering report to a system effectiveness manager, who in turn reports to the system engineering manager, as illustrated in Fig. 3-3, or to a design support department (Ref. 2).

The rapid growth of the analytic and statistical aspects of maintainability in the 1960's is an outgrowth of the attention given to these aspects of reliability in the 1950's. (See par. 1-1.) As a result, a common practice has been to combine maintainability and reliability into a single organizational unit under one manager. While these two disciplines are closely related, they are still different disciplines with respect to both their physical (design) and analytic aspects. The danger in putting them together in one organizational entity is to create an overbalance in one direction or the other depending upon the orientation of the manager of the unit or on the preponderance of skills in the combined unit. There are significant enough differences in these two engineering disciplines to warrant separate organizational and supervisory considerations.

3-1.2.3 Maintainability in a Decentralized Organization

In large organizations that handle many large, complex projects, the maintainability effort is often organized along the lines of the maintainability activities described in par. 3-1.1. These activities may be the responsibility of different organizational entities in the management hierarchy and are often physically separated from one another, requiring a considerable coordination effort.

In such organizations, typical of many of the aerospace companies and military organizations which use the project or matrix organizational form, there is often a small maintainability program group in the project office whose responsibility is to cover the program and coordination activities described in pars. 3-1.1.1 and 3-1.1.5. They serve as the program manager's staff experts with regard to the interpretation of program maintainability requirements and the coordination of all maintainability activities for the project manager with regard to plans and schedules, as well as the interface with the other technical disciplines. A second group, concerned with the analytical functions described in par. 3-1.1.2, may be found in a separate system effectiveness organization as part of system engineering. A third group of people, concerned with the maintainability design features described in par. 3-1-1.3, may be part of the design engineering functional organization. A fourth group, concerned with the documentation requirements described in par. 3-1.1.4, may exist in an overall documentation and data organization as part of design or project support in either a functional or project office organization. Coordination of such a highly decentralized maintainability effort is often very difficult, and various cliques tend to arise which may be devisive and may lead to a poorly executed, inefficient result.

As described in par. 1-4, maintenance engineers who look at the system from the user's viewpoint are usually concerned with maintenance engineering analysis, the analysis of maintenance tasks and maintenance resource requirements, and the preparation of maintenance instructions. Such personnel tend to reside within a support organization such as logistics or field service, separate and apart from the maintainability engineers who are then primarily concerned with maintainability analytic and design activities. This adds yet another coordination activity. Fig. 3-4 illustrates the complete decentralization of maintainability activities which can and often does occur in large organizations. (See also Refs. 2 and 3.)

Finally, note must be taken of the tendency that has arisen during recent years to group maintainability and reliability functions into an assurance organization. While there are design assurance requirements with regard to these disciplines, such assurance requirements primarily are concerned with ascertaining that the system and equipment do in fact meet customer specifications as part of an overall assurance function, and such assurance functions do in fact belong as part of a product assurance organization. However, it must

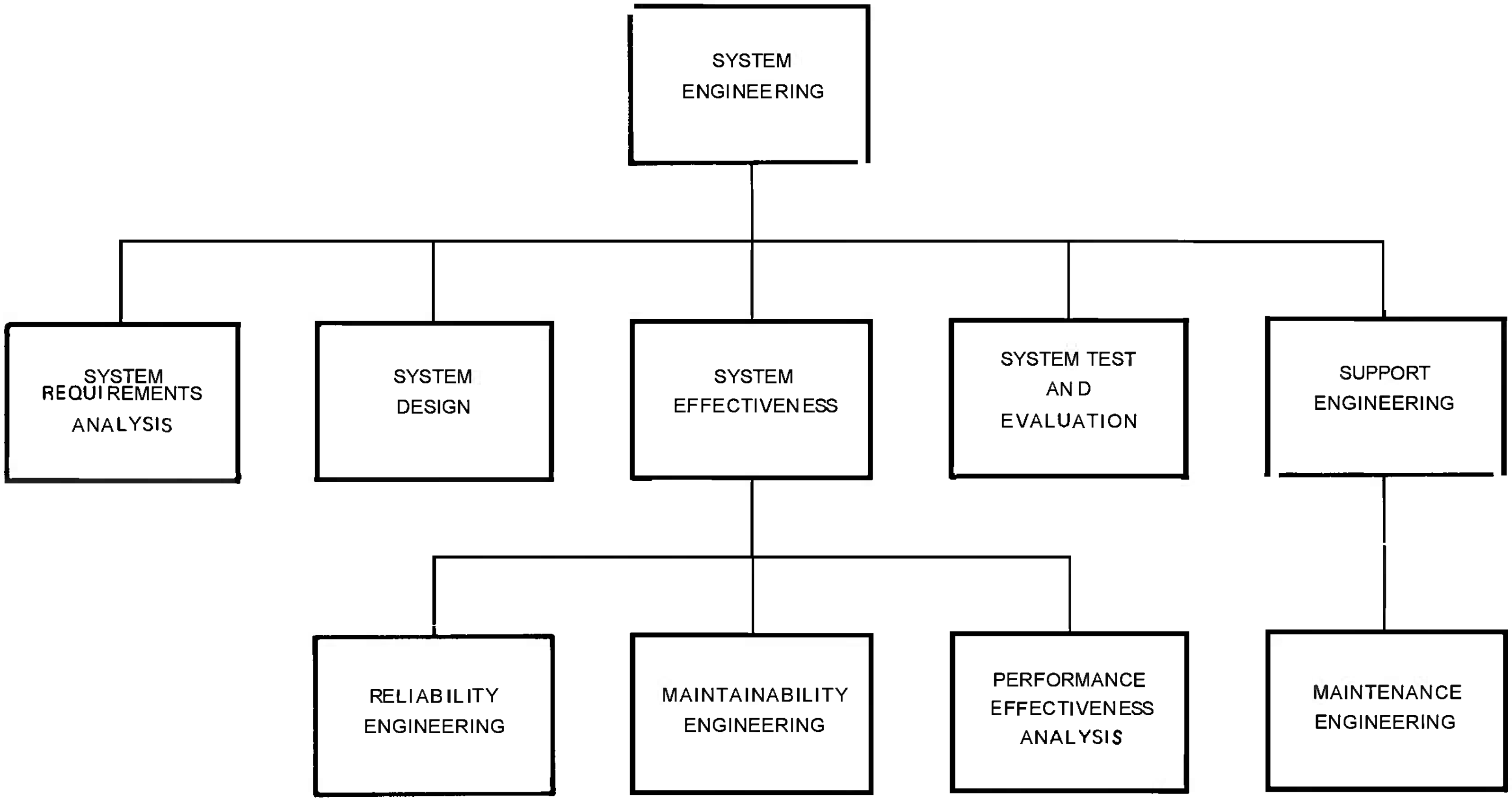


Figure 3-3. Maintainability Engineering as Part of System Engineering/System Effectiveness

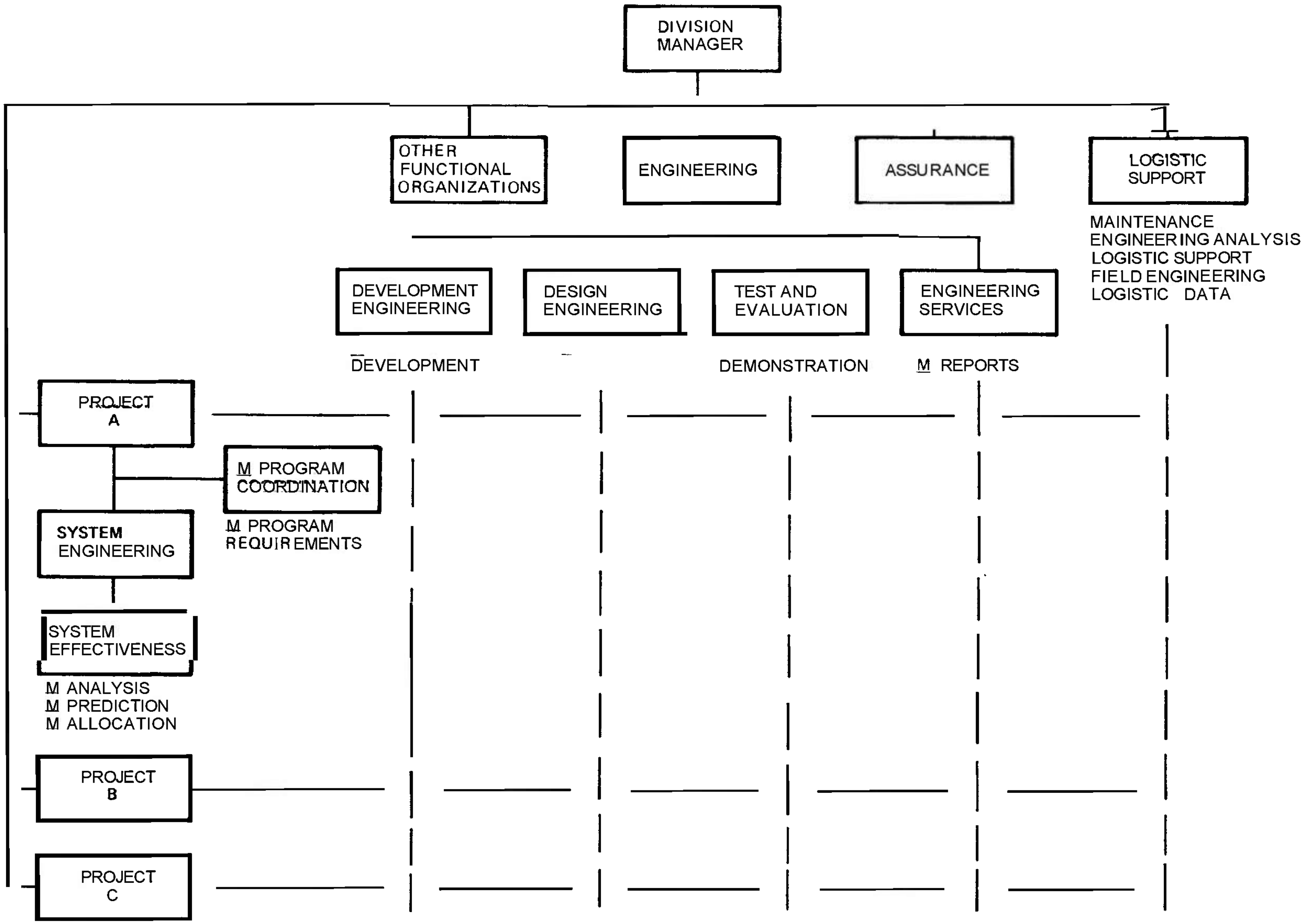


Figure 3-4. Example of Decentralized Maintainability Functions in a Matrix Organization

be remembered that maintainability (and reliability) are **engineering** disciplines and rightfully belong in the engineering organization. Because of their effect upon system design, these disciplines must be in a position to influence design.

There has been a tendency on the part of many contractors to restructure their internal organizations as they think the customer wants them to or to reflect the customer's organization. In many cases, this has resulted from the customer essentially specifying, directly or indirectly, how it wanted the organization to be. In others, such restructuring has been a direct result of the fact that the request for proposal and specification requirements placed undue emphasis on the organizational aspects of the project or the company's management structure. This reactive tendency on the part of contractors has created confusion within the internal company organizations, particularly in those aerospace companies which deal with different defense and space agencies. It has also led in some cases to efficient, well run company organizations becoming ineffective, with a consequent increase in both direct and indirect personnel, increased emphasis on paper design and unnecessary documentation efforts, and increased costs—even including large overruns.

It is incumbent on company and project management to organize and manage not only its maintainability effort, but its entire program in the manner which is most natural, effective, and efficient for the company. Such an organization should have little difficulty in convincing the customer that it can meet program objectives.

As an outgrowth of the recognition of the necessity of effective management to design and engineer mature systems, a strong trend has developed in the large aerospace companies to combine under one "hat" all supporting engineering disciplines that have a direct impact on design. This leads to such organizational structures as product effectiveness directorates that operate under the engineering organization and encompass reliability, maintainability, safety, human factors, value engineering, and cost-effectiveness as departments or branches.

3-2 EFFECTIVE MAINTAINABILITY MANAGEMENT

Effective maintainability management depends upon a number of factors. First is the recognition by top management that maintainability is an essential characteristic of system/equipment design. Proper attention will be paid to maintainability only in those organizations where top management is fully aware of this and has established policies for the effective application of maintainability engineering (Refs. 2, 3). Second is the establishment of maintainability engineering as a functional entity in the company organization, as described in par. 3-1, at an organizational level such that its relationships and functions with respect to other organizational entities can be effectively carried out (Ref. 2). Third, not only must maintainability be recognized and receive official sanction and status within the overall enterprise, it must also be accepted by all members of the organization as one of the technical disciplines, along with design engineering, reliability engineering, human factors engineering, safety engineering, test and evaluation, maintenance engineering, and integrated logistic support, with all of which it has strong interfaces. The physical location of the maintainability organization and the extent to which its functions are centralized or decentralized may have a significant impact on the effectiveness of the maintainability effort. Isolation or decentralization of the group creates problems of coordination and tends to de-emphasize the importance of the maintainability engineering effort.

In addition, effective maintainability management requires that the maintainability function be planned, organized, directed, budgeted, monitored, and controlled in the same manner as the other disciplines. Of particular significance is the establishment of maintainability policies and procedures as part of total engineering management and their inclusion in policy and procedure manuals. Establishment of program plans for carrying out the maintainability function is vital to effective maintainability management. These plans must be in accordance with life-cycle management from the recognition of the need, analysis of system requirements with respect to maintainability, establishment of maintenance concepts and features, and incorporation of maintainability as a significant design characteristic.

3-2.1 MANAGEMENT FUNCTIONS THROUGHOUT THE LIFE CYCLE

In order to achieve an effective and efficient system design, maintainability considerations must occur throughout the system life cycle. The life cycle for Army systems extends from the development of the concept of a new end item arising out of DoD and Army operational capability studies through validation, design and development, production, installation, support, and operation (Refs. **4-7).**

Effective implementation of the maintainability program includes all phases of the system life cycle by means of the user-producer relationship described in par. 1-4.1. A stated Army principle concerning life-cycle management of reliability and maintainability is:

"Effective and thorough managerial direction, planning, programming, and resource allocation will be provided throughout the life cycle of each item so as to enable the objectives of the reliability and maintainability program to be achieved for that item. Reliability and maintainability will be identified as principal characteristics of the item, and the status of these characteristics will be assessed throughout all phases of its life cycle" (Ref. 8).

3-2.1.1 The System Life Cycle

The system life cycle consists basically of the following phases:

1. Concept Development
2. Validation
3. Production
4. Operation.

These phases are illustrated in Fig. 3-5.

During concept development, an operational need or threat is transformed into a set of operational requirements, and high risk areas are identified. During validation, the concepts are verified, high risk areas resolved or minimized, and the operational requirements transformed into a set of system requirements. These requirements are then transformed into a system design, prototype for test and evaluation, and drawings and specifications to be used in the production phase. During production, the system is produced, accepted and installed in a ready-to-use condition for subsequent operation. During the operation phase, the operating system is used, logistically supported and modified when necessary. These phases are discussed in greater detail in the paragraphs which follow.

3-2.1.1.1 Concept Development

The objective of concept development (Fig. 3-6) is to develop and select the best materiel approach to satisfy an established operational need and to prove the feasibility of the approach from a technical, cost, and schedule standpoint. In addition to the preparation of materiel development objectives, some of the activities which characterize this phase include the preparation of a recommended approach, advanced development objectives, a system development plan, and other documentation associated with the prerequisites for system definition. At the conclusion of concept development, a review and system status evaluation of the system development plan is conducted to assure that the necessary preliminary work has been done. Threat and operational analyses, trade-offs, cost and mission effectiveness studies, and the state of development of components and technology provide a firm foundation for entering the validation phase.

During concept development, the primary maintainability concern is the derivation of system effectiveness requirements and criteria as discussed in Chapter 2, and the determination from operational and mission profiles of the maintenance and logistic support policies and the boundaries required to meet mission objectives. As a result of threat and mission analyses and a description of the operating environment, and consistent with Army doctrine and logistic support policies, the following must be accomplished in order to establish system maintainability requirements:

1. Description of mission and performance envelopes and system operating modes
2. Determination of mission time factors and system utilization rates
3. Determination of the duration of the system life cycle, including system deployment and out-of-service conditions
4. Elaboration of system effectiveness criteria expressed in mission-oriented terms
5. Description of the overall logistic support objectives and concepts, including maintenance concepts.

Army policy requires that:

"Suitable planning and consideration be given to reliability, maintainability, and availability during the concept development phase. In the conduct of feasibility studies and component development, consideration must be given to the reliability and maintainability potential of the equipment and its components. Reliability and maintainability predictions and information from similar systems should be assimilated to assist in selection of proper technical approaches, to identify areas of high technical risk, and to assist both in trade-off and cost-effectiveness studies and in final concept selection. Planning and requirement documents should be based on reliability and maintainability data from similar systems and from feasibility studies and should include sufficient detail to show how reliability and maintainability requirements are to be attained. Availability, reliability, and maintainability requirements must be stated in terms appropriate to the item considering its intended purpose, complexity, and quantity expected to be produced, and must be clear and capable of being measured, tested for, or otherwise verified" (Ref. 7).

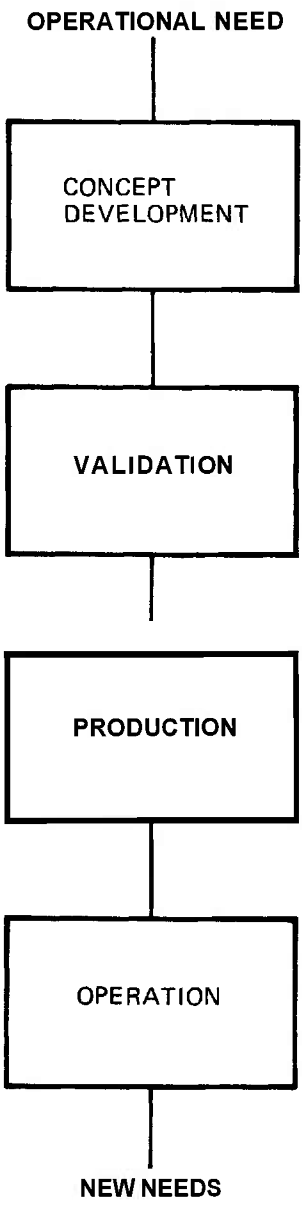


Figure 3-5. The System Life Cycle

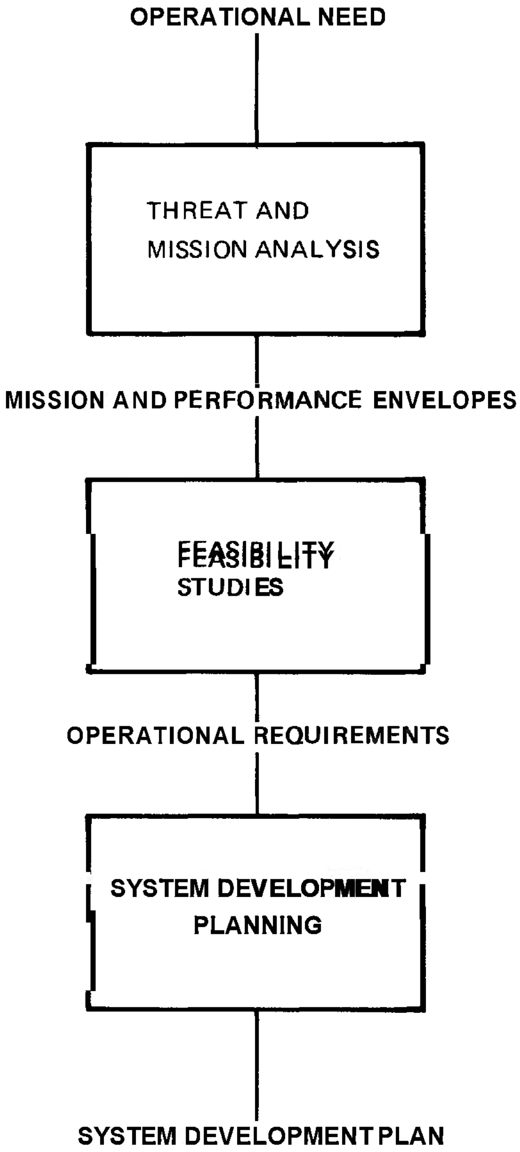


Figure 3-6. Stages in the Concept Development Phase

3-2.1.1.2 Validation Phase

Validation is the phase in which the operational requirements developed and formulated during concept development are further refined in terms of system design requirements. It is essentially a first step in system development which might be called preliminary system design. During this time period, major program characteristics are defined, and the high risk areas identified during concept development are resolved or minimized through extensive analysis and hardware development. This effort may be conducted by competitive contractors, a sole source contractor, or in-house.

Validation may be accomplished through the use of comprehensive design analysis and system definition studies, or through the use of hardware development and evaluation, especially in the identified high risk areas, or both. Parts of the system or a complete model may be developed to demonstrate that desired *performance* objectives can be achieved. This is sometimes called ***prototyping*** or ***parallel undocumented development*** when performed competitively by contractors.

The objective of validation is to assure that full-scale development is not started until costs and schedules, as well as performance and support objectives, have been carefully prepared and evaluated. This may include prototype construction, test, and evaluation in high risk areas, and should result in a high probability of successfully accomplishing the development of the system or end item. The ultimate goal, where full-scale development is to be performed by a contractor, is achievable performance and support specifications that are responsive to the operational requirements and are backed by a firm fixed-price or fully structured incentive-type contract.

Adequate and effective materiel support planning must be accomplished to insure inclusion of support requirements—including integrated logistic support goals and objectives, maintenance support planning, and maintainability requirements. Fig. 3-7 shows the major steps in the validation phase.

The Request for Proposal Work Statement for validation and the specimen work statement for engineering development must contain requirements for a reliability and maintainability program, including test and demonstration requirements. Guidance given to a contractor concerning incentives in this area should be reviewed for completeness and accuracy, and for schedule and cost implications. The description of evaluation criteria will make explicit the fact that the proposed reliability and maintainability program is a significant element of proposal evaluation (Ref. 7).

Maintainability management during a validation effort is concerned with the following tasks:

1. Preparation of maintainability program plan in accordance with contractual requirements.

2. Determination of specific reliability, maintainability, and system effectiveness requirements.

3. Preparation of maintainability policies and procedures applicable to the validation and follow-on full-scale engineering effort.

4. Assistance to maintenance engineering in the performance of maintenance engineering analyses, including the preparation of the overall maintenance concepts resulting from the analysis of mission profiles, logistic endurance factors, logistic policies, and system effectiveness requirements.

5. Participation in trade-off analyses and performance of maintainability predictions and allocations for subsystems/equipments and end items in connection with system effectiveness/reliability trade-offs.

6. Preparation of a maintainability demonstration plan.

7. Derivation of specific maintainability design guidelines for use by design engineers, resulting from maintenance engineering analyses.

8. Coordination and monitoring of the maintainability efforts of the entire organization.

9. Participation in project and design reviews with regard to maintainability.

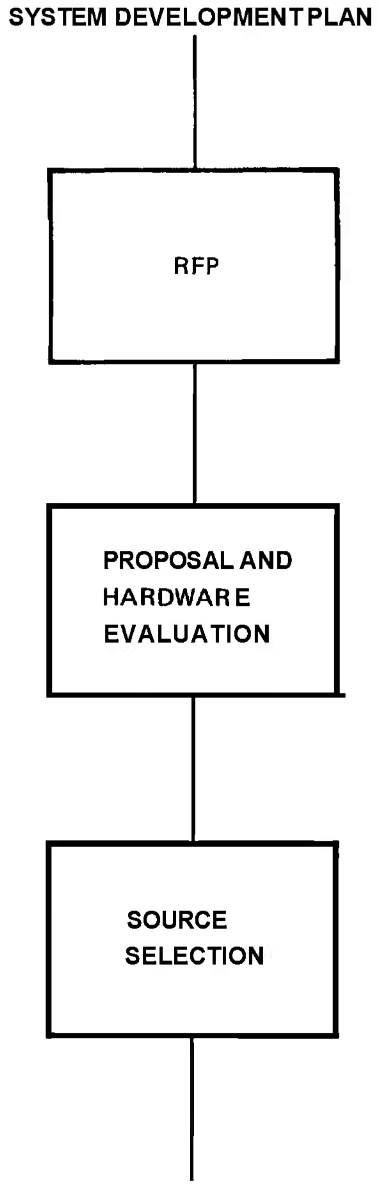
10. Preparation of a plan for data acquisition, collection, analysis, and evaluation.

11. Establishment of maintainability incentives or penalties.

The final output of the validation effort in the area of maintainability consists of the specific maintainability design requirements and guidelines to be used during the following phase. Fig. 3-8 is a flowchart showing the sequence of activities which occur during the concept development and validation phases from the input mission requirements to the output reliability and maintainability design requirements.

Army policy is that:

"Proposals for engineering and operational development will be evaluated for reliability and maintainability aspects to assure that the contractor understands and is responsive to the requirements, and that he has proposed an effective and realistic set of resources and management tools to assure timely attainment of the requirements and demonstration of that attainment.


SYSTEM DEVELOPMENT PLAN
RFP
PROPOSAL AND HARDWARE EVALUATION
SOURCE SELECTION

Specific detailed evaluation will be made of the reliability and maintainability program plans. The system description and development descriptions also will be reviewed for technical adequacy in those areas pertaining to reliability and maintainability characteristics. Subsequent to source selection, refinements will be made as required to assure a complete, technically acceptable package" (Ref. 7).

In some cases engineering and prototype development and pilot production are among the final objectives of the validation phase. There is sometimes an overlap with the production phase when advanced production engineering, production planning, and long lead-time item procurement or fabrication are required concurrently with development. Engineering and prototype testing is performed to demonstrate that a system or end item satisfies the military requirement.

The final product of validation is information which can be provided to a chosen contractor for use in producing the end item or system developed and in the logistic support of the fielded system. Fig. 3-9 illustrates the stages in the engineering development aspects of validation. Detailed descriptions of these stages are contained in Refs. 8, 9, and 10.

The bulk of the maintainability engineering effort occurs during concept development and validation. Maintainability management is particularly critical at this time. It includes the activities described in par. 3-1. More specifically, the following functions must be accomplished prior to production:

1. Updating of the maintainability program plan in accordance with final development contract specification requirements.

2. Preparation and issuance of detailed program schedules, milestones, budgets, work orders, and their periodic review and updating.

3. Monitoring and controlling of the maintainability engineering effort in accordance with the approved program plan and management policies and procedures.

4. Detailed prediction and allocation of quantitative maintainability requirements down to the lowest configuration end item.

5. Participation in system effectiveness and design trade-offs involving maintainability in order to meet predetermined maintainability predictions and allocations and overall system effectiveness requirements.

6. Assistance to maintenance engineering in the performance of detailed maintenance engineering analyses (Ref. 11).

7. Preparation of specific maintainability test and demonstration plans as part of equipment and system test and evaluation, including the collection and analysis of test data, initiation of corrective actions as a result of test and demonstration, and the preparation of maintainability demonstration reports.

8. Provision of consultation to design engineers with regard to specific maintainability design features and evaluation of the effects of maintainability design on overall maintainability and system effectiveness quantitative requirements.

9. Coordination and monitoring of subcontractor maintainability efforts.

10. Participation in detailed project and design reviews and drawing approval with regard to maintainability.

11. Assurance that the interfaces among other engineering disciplines such as reliability, human factors, safety, logistic support, test and support equipment design, and technical data are coordinated with respect to maintainability.

3-2.1.1.3 Production Phase

The purpose of the production phase is to manufacture, test, deliver, and in some cases install the specified system in accordance with the technical data package resulting from the previous life-cycle phases. The maintainability engineering design effort will be largely completed at this time. However, the continuing life-cycle management of maintainability should then be carried on as part of a sustaining engineering effort, during which the maintainability design is reviewed and updated as a result of initial field experience, engineering changes, and logistic support modifications. The maintainability effort during this phase includes:

1. Monitoring the production process

2. Evaluating production test trends to assure that there are no adverse effects on maintainability, maintenance concepts, provisioning plans, etc.

3. Assuring correction of all discrepancies having an adverse effect on maintainability

4. Reviewing and evaluating all change proposals for their impact on maintainability

5. Participating in the establishment of controls for process variations, errors (workmanship and design), and other fabrication and test discrepancies that could affect maintainability.

3-2.1.1.4 Operation Phase

The operation phase of the life cycle of Army

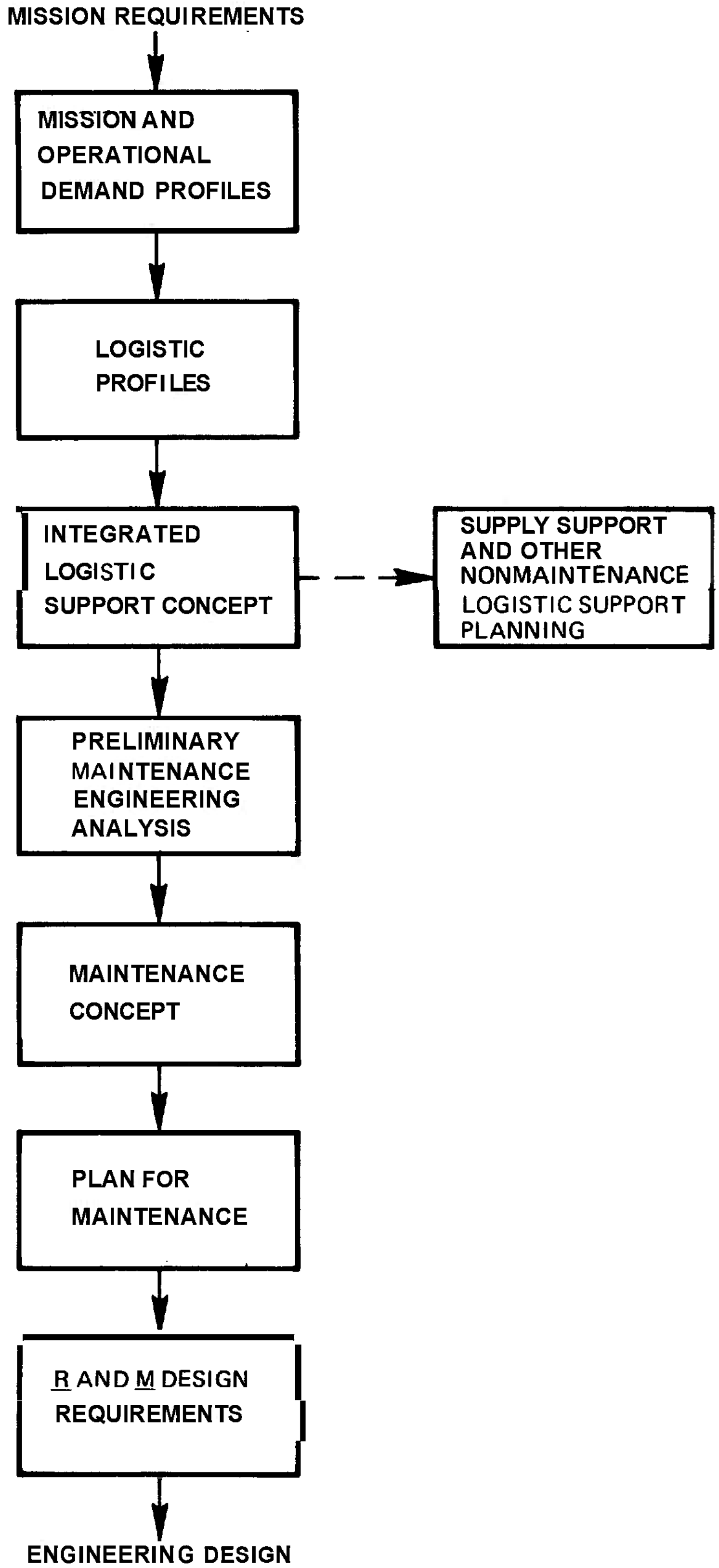


Figure 3-8. Maintenance Engineering Analysis During Early Program Stages

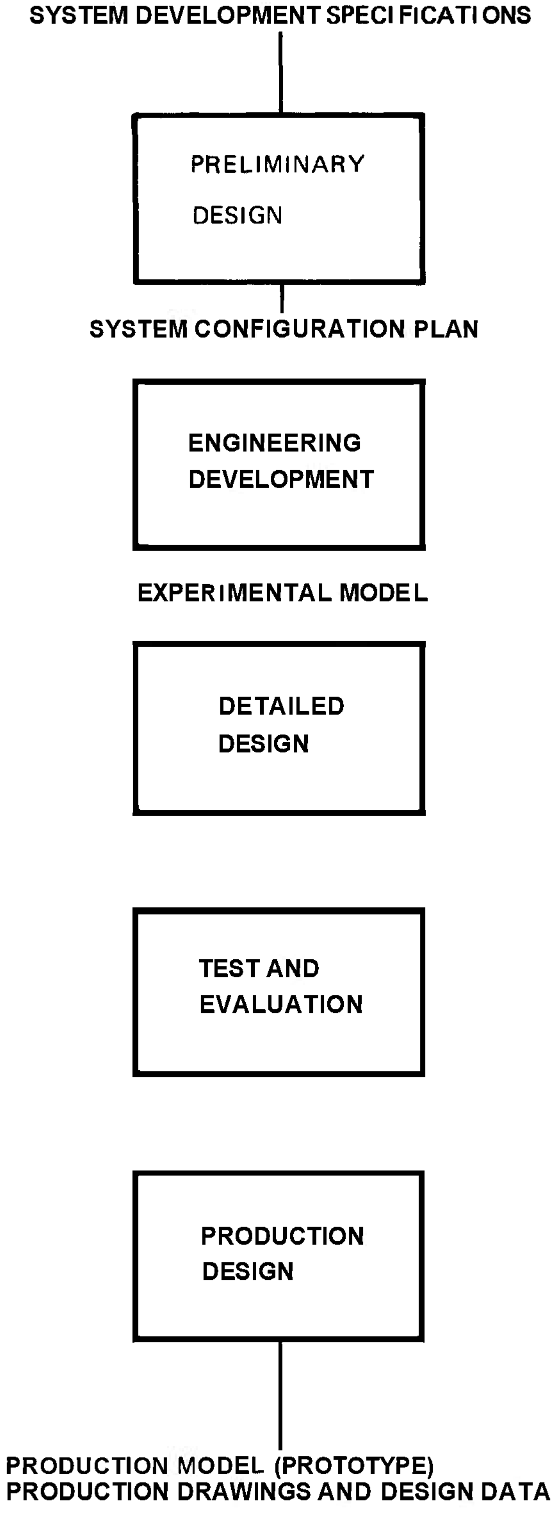


Figure 3-9. Stages in Engineering Development Aspects of Validation

materiel starts when the first military unit is equipped, and ends when the end item or system has been declared obsolete and has been removed from Army inventory. This phase is characterized by supply, training, maintenance, overhaul, and materiel readiness operations on the end item or system being used by operational units. This time period is most significant because it is here that the true cost-effectivenessof the system and its logistic support are demonstrated, and historical maintainability data are recorded for use on future products.

There are no specific maintainability engineering requirements during this phase. Feedback data from the field with regard to system effectiveness, reliability, actual field maintenance, and maintainability should be used as a basis for product improvement and the correction of deficiencies as a result of system operation and support.

Army policy states that the following elements, as appropriate, will be executed:

a. Effective collection, analysis, and follow-up of failure data in accordance with the selected data collection plan; timely identification and resolution of problems, including product improvement where required.

b. Effective controls over parts substitution during maintenance operations.

c. Periodic stockpile reliability evaluation of selected items (an integral part of the surveillance program).

d. Evaluation of the effects of repetitive maintenance.

e. Effective program to control application of approved modifications.

f. Continuous assessment of reliability and maintainability characteristics, based on operational data (Ref. **7**).

Disposal takes place when an end item or system has been declared obsolete and no longer suitable for use by Army units. The item is then removed from inventory and scrapped or salvaged.

Summarizing this chapter, to make management of maintainability most effective, maintainability engineering should be so placed organizationally that it can impact on design and also interface directly with other disciplines, such as safety, reliability, human factors, value engineering and system-cost effectiveness, with whom maintainability is inseparably interrelated. The benefits of such organizational structure are conservation of resources and specialists, experience retention, less duplication, and lower program costs (Ref. **12**).

REFERENCES

1. MIL-STD-470, ***Maintainability Program Requirements (for Systems and Equipments),*** 21 March **1966.**
2. **B.** S. Blanchard and E. E. Lowery, ***Maintainability,*** McGraw-Hill, New York, **1969.**
3. DA Pam **705-1**, ***Maintainability Engineering.***
4. **AR** 70-1, ***Army Research, Development, Acquisition.***
5. Deleted
6. DA Pam 1**1-25,** ***Life Cycle Management Model for Army Systems.***
7. **AR 702-3,** ***Army Materiel Reliability, Availability, Maintainability.***
8. M. B. Kline, ***Maintainability Considerations in System Design,*** Doctoral dissertation, Department of Engineering, UCLA, **1966,** University Microfilms, Inc., Ann Arbor, Michigan.
9. M. **B.** Kline and M. W. Lifson, "Systems Engineering", Chapter 2 in ***Cost-Effectiveness: Economic Evaluation of Engineered Systems,*** J. **M.** English, Ed., Wiley, New York, **1968.**
10. M. **B.** Kline and M. W. Lifson, ***Systems Engineering,*** Notes for UCLA short course, Engineering & Physical Sciences Extension, UCLA, Los Angeles, **1970.**
11. TM 38-703, ***Integrated Logistic Support,*** 5 Vols., Nov. **1969.**
12. S. L. Roush and J. L. Maybell, ***Effectiveness Engineering Organization in the Aerospace Industry,*** McDonnell Douglas corporation, Douglas Paper No. **3768,** October **1966.**

CHAPTER 4

MAINTAINABILITY ALLOCATION AND PREDICTION

SECTION I

INTRODUCTION

4-1 GENERAL

Maintainability allocation and prediction are tasks required by MIL-STD-470 (Ref. 1) within the framework of establishing and maintaining an effective maintainability program. In this Standard, twelve specific tasks are defined. We are here concerned with two of these tasks, namely, Tasks **5.2** and ***5.6.***

Task **5.2** requires the performance of a maintainability analysis of which maintainability allocation is an important part, stated in the following terms: "As a major task of the analysis, the contractor shall allocate quantitative maintainability requirements to all significant functional levels of the system/equipment".

Allocation is performed in the initial phases of a program when a system is defined and its overall maintainability objectives are established. Allocation apportions the overall system objectives to the functional block, thus providing the designers with maintainability requirements that they have to meet, possibly down to repairable items level. Reallocations may be necessary as the design features gradually are established.

Task ***5.6*** requires the prediction of maintainability, stating that such prediction "shall estimate quantitatively the maintainability system/equipment parameter values for the planned design configuration. The quantitative estimates shall be used to judge the adequacy of the proposed design to meet the maintainability quantitative requirements and identify design features requiring corrective action".

Prediction periodically assesses the maintainability characteristics of emerging designs to insure that the allocated maintainability requirements are being met or to identify qualitative design features that require corrective action so as to meet the overall system maintainability requirements. Prediction is performed periodically until the design configuration has the potential to meet the system maintainability goal and to pass the maintainability demonstration test, if such is required.

Allocation and prediction may use different techniques. However, they are based on two specific factors that are common to all maintainability assessment techniques and must be quantitatively determined in each case.

4-1.1 MAINTAINABILITY FACTORS

There are two specific maintainability factors that have been recognized as the basic ingredients of maintainability techniques. Their quantitative values are of decisive importance in meeting maintainability objectives. The two basic factors are:

1. The ***time*** required to restore a failed system or equipment to an operationally ready state by performing corrective (unscheduled) maintenance, or to sustain a desired performance and reliability level by performing preventive (scheduled) maintenance.

2. The ***frequency*** at which corrective and preventive maintenance actions occur at the system/equipment level.

These two factors jointly determine the quantitative maintainability characteristics of a design. Specifically, they determine the mean active corrective maintenance time M_c, the mean active preventive maintenance time M_p, the mean active corrective and preventive maintenance time $\bar{M}$, the maximum maintenance time M_{MAX}, and the equipment maintainability function $M(t)$, as already defined and discussed in Chapter 1, Section **11.**

As an illustration of the impact of the maintenance time and of the frequency of occurrence of maintenance on maintainability let us recall here some basic equations.

The mean active corrective maintenance time $\bar{M}_c$ of a system consisting of n replaceable or repairable items is given by

$$\bar{M}_c = \sum_{i=1}^{n} M_{ci} f_{ci} \Big/ \sum_{i=1}^{n} f_{ci} \tag{4-1}$$

where

M_{ci} = system active corrective or repair downtime when the ith item fails

f_{ci} = frequency of the ith item failures, usually expressed in terms of the failure rate λ_i of the ith item in units of "number of failures per one system operating hour".

The mean active preventive maintenance time $\bar{M}_p$ is given by

$$\bar{M}_p = \sum_{i=1}^{n} M_{pi} f_{pi} \Big/ \sum_{i=1}^{n} f_{pi} \tag{4-2}$$

where

M_{pi} = system active preventive maintenance downtime when the ith item is preventively replaced or otherwise preventively maintained

f_{pi} = frequency of such event per system operating hour.

The mean active corrective and preventive maintenance time $\bar{M}$ is given by

$$\bar{M} = (\bar{M}_c f_c + \bar{M}_p f_p)/(f_c + f_p) \tag{4-3}$$

where $\bar{M}_c$ and $\bar{M}_p$ are defined as before, and

$$f_c = \sum_{i=1}^{n} f_{ci} \tag{4-4}$$

$$f_p = \sum_{i=1}^{n} f_{pi} \tag{4-5}$$

are the frequencies at which the system is correctively or preventively maintained, both expressed in system operating hours, not calendar time. It is essential that in Eq. 4-3 the frequencies be expressed in the same units, i.e., number of system maintenance actions per system operating hour, and that the mean active corrective and mean preventive maintenance times also be expressed in the same time units, i.e., seconds, minutes, or hours.

By inspection of the given equations we see that the system mean downtime indices or measures depend only on system maintenance ***downtime*** and on the ***frequency*** at which system outages occur. At first glance, it would appear that these equations apply only to series systems. However, if systems contain redundant elements and the ith item is defined as a serial element which may or may not contain redundancy, the preceding equations become generally valid as long as we recognize that the ***time*** element M_i applies to system downtime and the ***frequency*** element f_i is the frequency at which the system goes into a down condition. In addition, when considering Eq. 4-3, one must realize that preventive maintenance may, and usually does, have an effect on the frequency f_c of corrective maintenance actions because preventive maintenance is applied to postpone the occurrence of failures. Thus, if the ith item is subject to periodic preventive maintenance (such as scheduled replacement), system failures on account of the ith item will no longer occur at the item's own failure rate but at a rate determined by the nature of the preventive maintenance policy (Ref. 2, Chapters 3 and 4).

4-1.2 METHODOLOGIES

As already stated, several methods exist to allocate and predict system or equipment maintainability. These are in MIL-HDBK-472 (Ref. 3) and the Air Force Design Handbook (Ref. 4, Section 3B). In general, these techniques utilize the time summation method of individual maintenance action, based on the frequency of occurrence of individual maintenance actions and their average duration, in order to determine the overall system mean maintenance time and related maintainability indices.

All methods use various "building block" type breakdown diagrams to establish the required respective maintenance actions. The mean maintenance time determinations are based on the equipment qualitative design features. The maintenance times are derived from statistical historical data, selected observation data, expertise judgments, simulation and synthesis modeling, design checklists, extrapolation, or matrix tabulation methods.

In the selection of specific prediction and allocation techniques, the maintainability analyst need not con-

.ine himself to a single method. He may utilize several methods in a single project. Ordinarily, there may be one major allocation and prediction method approach upon which principal reliance is placed. The selection of a specific method will affect the plans for data collection and program control. As discussed in Ref. 5, which includes some of the prediction methods of MIL-HDBK-472, certain practical considerations or factors may control the choice of a prediction technique, namely:

1. Environment consideration (maintenance level and type, maintenance concept—what, when, where, why, and how maintenance will be done, and the logistic support situation)
2. Similarity to other equipment
3. Scope of the prediction and allocation effort
4. Accuracy of estimate required
5. Degree of design guidance required
6. Point of application in the design cycle (early or late).

As examples of the application of these factors, consider the following situations:

1. Suppose that factor 3 is very limited because 'here is not much time or money available; then the ɔest one could do would be to (a) attempt rough extrapolation from maintenance history of similar equipment, or (b) conduct a brief judgment-type review by experts.

2. Now consider a new system which is "low" on the similarity factor 2, being quite different in concept and realization. In this case extrapolation is not indicated, and the checklist approach may have to be excluded if the tasks required of the maintenance technician are not well enough represented in the usual checklist. Probably some simulation modeling or mock-up determinations would be called for.

3. Suppose factor 4 is important because high accuracy is required. Then, extrapolation and expertise judgment would probably not be suitable; a time synthesis method appears most accurate, using a detailed qualitative checklist.

4. Suppose management places high value on factor 5. Then the sphere of interest not only would include the "gross" maintainability estimate, but also the specific "causes" or "specific design features" affecting a prediction. This would require an extensive qualitative-quantitative design checklist simulation and synthesis, and continuous feedback to the designers.

In any case, before a method is selected, one must develop a maintenance concept upon which to base a maintenance functional flow block diagram of the maintenance tasks to be performed which defines what, when, where, why, and how much maintenance. From this, a "building block" maintenance functional flow diagram can be developed for the proposed equipment design to be sure that all maintenance tasks are accounted for. Also of great importance are failure modes, effects, and criticality analysis requirements, in order to define the need for corrective maintenance and preventive maintenance actions. One must consider the field operational environment when developing maintenance time distributions used to determine the mean time and maximum time indices. Too often inherent indices are used, based on biased experimental values under laboratory environments, or controlled observations utilizing highly trained and biased technician skills not representative of field environments. Also, field operational environment degraded failure rates should be used instead of the inherent failure rates whenever operational field maintainability is of interest.

SECTION II

ALLOCATION

4-2 ALLOCATION FACTORS

4-2.1 BUILDING BLOCK THEORY OF ALLOCATION

In the systematic development of realistic allocation factors, the maintainability engineer must consider all of the following:

1. Two types of maintenance
2. Seven categories of maintenance action time elements
3. Three levels of maintenance
4. Five major steps of allocation
5. Nine functional breakdown levels of a system.

The two types of maintenance are:

1. Corrective or unscheduled maintenance (repair or restore equipment which has failed to meet the operational performance required; i.e., maintenance that cannot be scheduled due to randomness of failure).

2. Preventive or scheduled maintenance (maintain equipment to sustain the operational performance required; i.e., maintenance that can be scheduled on an operating time or calendar basis).

There are seven major categories of maintenance actions or time elements required to perform the two types of maintenance. Arranged in order of sequence of logical steps, one, all, or a combination of the following steps may be required for maintenance (corrective-unscheduled or preventive-scheduled) of a system.

1. Preparation. Inspection; obtaining support tools, equipment, repair parts and supplies; warm up and check out; verification of the system status

2. Diagnosis. Localization and determination of the cause of failure or condition; isolation or determination of the item location causing the failure or condition of item to be maintained

3. Replacement. Dissassembly and gaining access to the item; interchange of the item with a serviceable item; reassembly, including closing of accesses required to gain access for disassembly

4. Adjustment and/or alignment (may be part of the sequence of step 3)

5. Servicing. Performance of steps required to keep the item in an operating condition, such as cleaning, lubricating, fueling, and oiling.

6. Check out and inspection. Verification of the maintenance action to ascertain that the equipment is restored to its operational performance readiness

7. Item repair. Maintenance actions needed to restore a removed item if such item is not of the throw-away type; includes one, all, or a combination of the steps previously listed; may be performed at any level of maintenance as stated in the established maintenance concept.

The three levels of maintenance where the two types of maintenance and the seven maintenance steps can be performed (Refs. 6 and 7) are:

1. Organizational Maintenance Level—maintenance performed by the using organization on its own equipment. This maintenance consists of repairs of a first and second level-type within the capabilities of the authorized operator or organization maintenance technician and within repair parts, tools, and test support equipment available. Normally the skill level requires the lowest skills developed for maintenance work. Organization level personnel are generally occupied with the operation and use of the equipment and have minimum time available for detailed maintenance or diagnostic check-out. This is the level of maintenance where the minimum equipment downtime for maintenance must be achieved in order to obtain the highest equipment availability or to achieve operational readiness for war-time use of the equipment. Maintenance is usually restricted to periodic scheduled preventive maintenance checks, cleaning of equipment, front panel-type adjustments, and replacement of items on a gross accessibility level.

2. Intermediate Maintenance Level—maintenance performed by mobile, semi-mobile, and/or fixed specialized organizations and installations. For the Army, this is broken down into direct support and general

support functions, with highly trained specialists for specialized equipment.

The direct support units are often designated to provide close support of the combat organizational-level maintenance to facilitate tactical operations. The direct maintenance support usually is limited to the repair of end-item or unserviceable assemblies in support of the combat units on a return-to-user basis. A larger supply of repair assemblies and components of major modular types is usually authorized for direct support. The diagnostic equipment is usually installed in mobile vans. The direct support maintenance is also geared to provide the highest equipment availability with minimum equipment maintenance downtime. Rapid turn-around time is an essential criteribn in the maintenance time allocation indices.

General support maintenance is usually conducted at semi-mobile or temporary fixed installations in the battlefield area to support the tactical battlefield organizational and direct support units. The maintenance support is that which cannot be provided by direct support mobile units. General support units have high personnel skills, additional test support equipment, and better facilities. Equipment repair is generally the repair of those items replaced by direct support to a small module or piece part (throw-away) level. Rapid turn-around time is not as imperative at the general support as at the direct support and organizational levels of maintenance.

3. Depot Level—the highest level of maintenance; provides support for maintenance tasks beyond the capabilities provided at the lower levels. The locafion is generally removed from the theater of operations and may provide maintenance for several theaters of operation. In some areas subdepots may be used, in safe havens of the theater of operation or in countries adjacent to the theater of operation. The support equipment may be of extreme bulk and complexity. Usually major overhaul and rebuilding are performed at depots. The large number of support requirements lend themselves tu the effective use of assembly-line techniques that, in turn, permit the use of relatively unskilled labor for a greater part of the work-load, with concentration of highly skilled specialists in key positions. For newly procured equipment in the military inventory, the contractor who produced the equipment may be employed for depot functions until such time as the depot staffed with Army personnel has enough work and experience to accomplish the maintenance. The depots are usually called upon to provide the necessary standards and calibration maintenance functions.

Maintainability allocations must be weighted and balanced for the economic use of the three levels of maintenance. Manipulation of the allocations directly affects the system logistic support costs and availability. The five major steps involved in maintainability allocation are:

1. Identify maintenance function to be performed on the system at each level of maintenance (organizational, intermediate, and depot) required to restore the system to an operational status.
2. Identify the elements that constitute a system down to the replaceable throw-away part.
3. Determine the frequencies for (a) corrective, unscheduled maintenance, and (b) preventive, scheduled maintenance, for each item of the system down to the replaceable throw-away part.
4. Determine the task times (mean time, median times, and maximum maintenance times at respective percentiles) for each item of the system down to the throw-away part.
5. Compute the mean times, median times, and maximum times at given percentiles (as required) for both corrective and preventive maintenance of the entire system.

Finally, the nine functional breakdown levels for a system—recognized by military specifications and used in making allocation (see Ref. 3, pp. 2-9 and 2-10)—are:

1. System
2. Subsystem
3. Equipment
4. Group
5. Unit
6. Assembly
7. Subassembly
8. Stage
9. Part.

In order to achieve the maintainability objectives it is essential that the two types of maintenance, the seven major categories of maintenance actions, the three levels of maintenance, the five major steps for making allocations, and the nine functional breakdown levels of a system be brought to bear in proper perspective. In order to achieve this and to make certain that all aspects of maintenance are covered and justified, maintainability engineers must use the following analytical tools:

1. Maintenance functional flow block diagrams
2. System functional-level building block diagrams

3. Failure modes effects and criticality analysis.

The extent of the application of these tools depends on many factors including the system constraints and the phase of the system life cycle.

4-2.1.1 System Description

4-2.1.1.1 Maintenance Functional Flow Block Diagram

This is a systematic method of outlining, interfacing, and relating the two types of maintenance (corrective and preventive) at each level of maintenance from the time a system goes down for maintenance until the last maintenance function has been performed to restore an item to its operational readiness. Stated in terms of system breakdown, the functional flow breakdown starts when the equipment goes down for maintenance and continues until the replaced item either is repaired for use in the original equipment, placed in stock as a replacement part, or thrown away. The function is expressed in a noun-verb fashion:

1. Organizational Level Maintenance: For failed equipment, the first level function is "perform corrective maintenance"; the second level function is "perform inspection", "diagnose failure", "remove faulty item", "replace", and "checkout" until the equipment is restored; the third level function is concerned with the repair of the removed faulty item.

2. Intermediate Level Maintenance: For failed equipment which cannot be repaired at the organizational level, the first level function is again "perform corrective maintenance", followed by the second and third level functions.. This is continued until the faulty item has been broken down to its throwaway status in the functional flow block diagram. For illustration, Figs. 4-1 and 4-2 are typical formats for a top level maintenance functional flow block diagram, showing a typical breakdown, numbering system, and entries made of the allocations. For each functional breakdown, the mean time to repair and frequency allocations are entered as the analysis proceeds. The functional flow block diagram is nothing more than putting down the analyst's thinking in a chronological and systematic order to assure that all maintenance functions and actions have been covered in his analysis and allocations. It also is a visual aid in explaining to management that all the factors and indices necessary to justify the maintainability quantitative and qualitative design requirements have been included. If trade-offs become necessary or corrective actions are needed, the interrelationship effects can be determined and the effects justified (see Ref. 8).

4-2.1.1.2 System Functional-Level Building Block Diagram

This is a systematic method of showing and defining the maintenance features and task actions required for each of the system-to-part breakdowns. This type of visual display is essential in order to explain further the equipment details that comprise the maintenance functions shown in the maintenance functional flow block diagram. This also complements the established maintenance concept in that it shows the essential maintenance tasks for each item of the requirements breakdown. The breakdown does not show the "where" of a maintenance action but does show "what" is needed and the equipment level at which an action takes place.

Each branch of the equipment diagram should indicate a termination point indicated by a consistent, identifying symbol or code. Examples follow:

1. A ***circle*** enclosing the item: to indicate the level at which a replacement (throw-away type) completes the correction of a malfunctioning item

2. A ***rectangle*** enclosing an item: to indicate the equipment breakdown

3. A ***triangle enclosing an "L"*** inserted next to the rectangular block: to indicate the level that an item may be fault "localized" without employing accessory support equipment

4. A ***triangle enclosing an "I"*** inserted next to the rectangular block: to indicate the level to which an item may be fault "isolated" using built-in or accessory equipment

5. **A *triangle enclosing an "A"*:**
   a. Inserted next to a rectangular block: to indicate adjustment or alignment before removal of a replaceable item
   b. Inserted next to a circle: to indicate adjustment or alignment after replacement of a replaceable item

6. **A *triangle enclosing a "C"*** inserted next to the symbol: to verify operation by built-in self-test or other testing equipment.

Examples of the application of these symbols in combination to indicate degrees of access level and termination points follow:

1. Faulty, replaceable throw-away item requiring disassembly maintenance action of a higher order:
   a. Place a rectangular block above the item, enclosing (identifying) the next higher assembly to be broken down.
   b. Encircle the faulty item to indicate throw-away replacement.

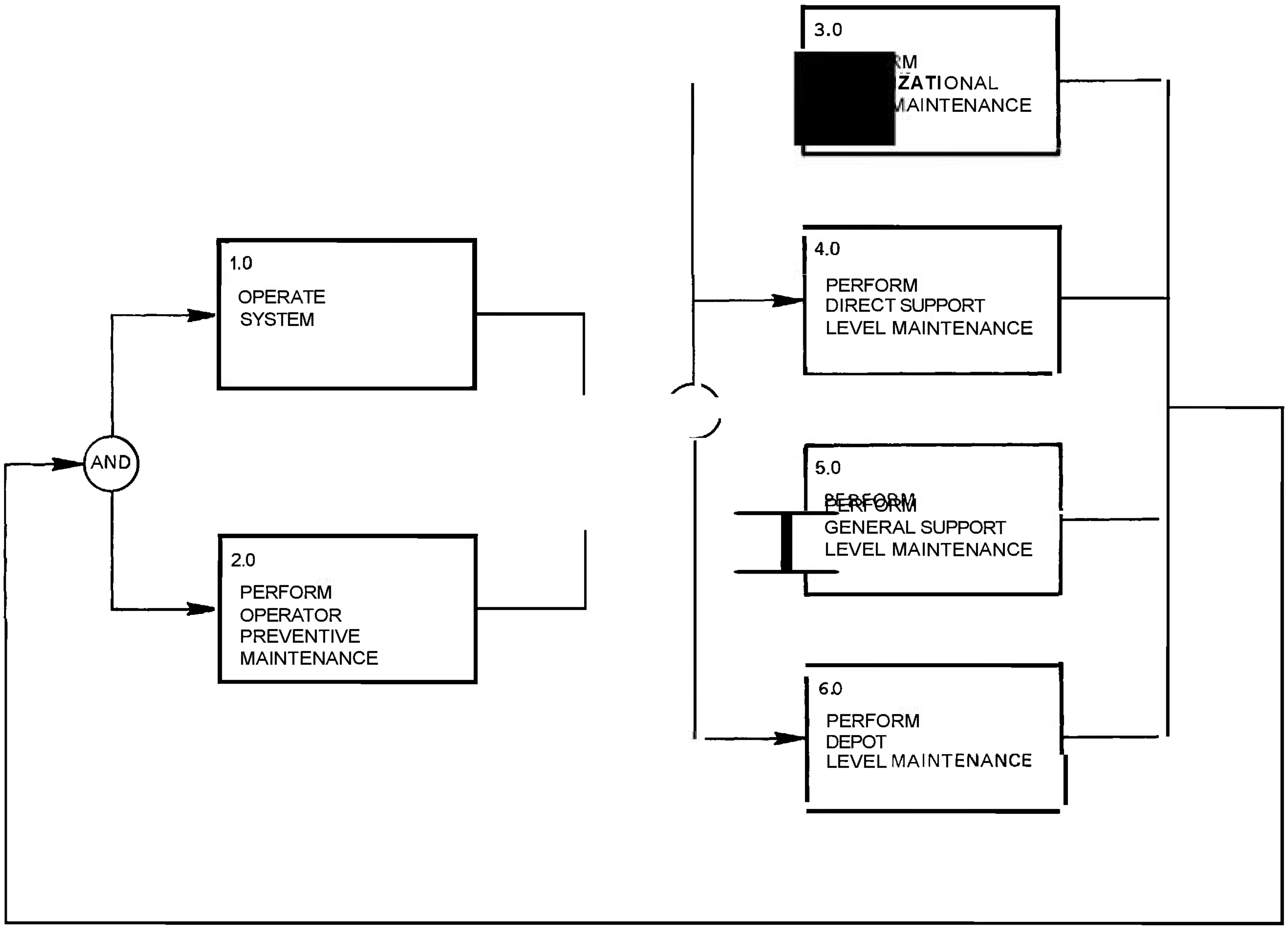


Figure 4-1. Typical Illustration of Maintenance Functional Flow Block Diagram (Top Level Flow)

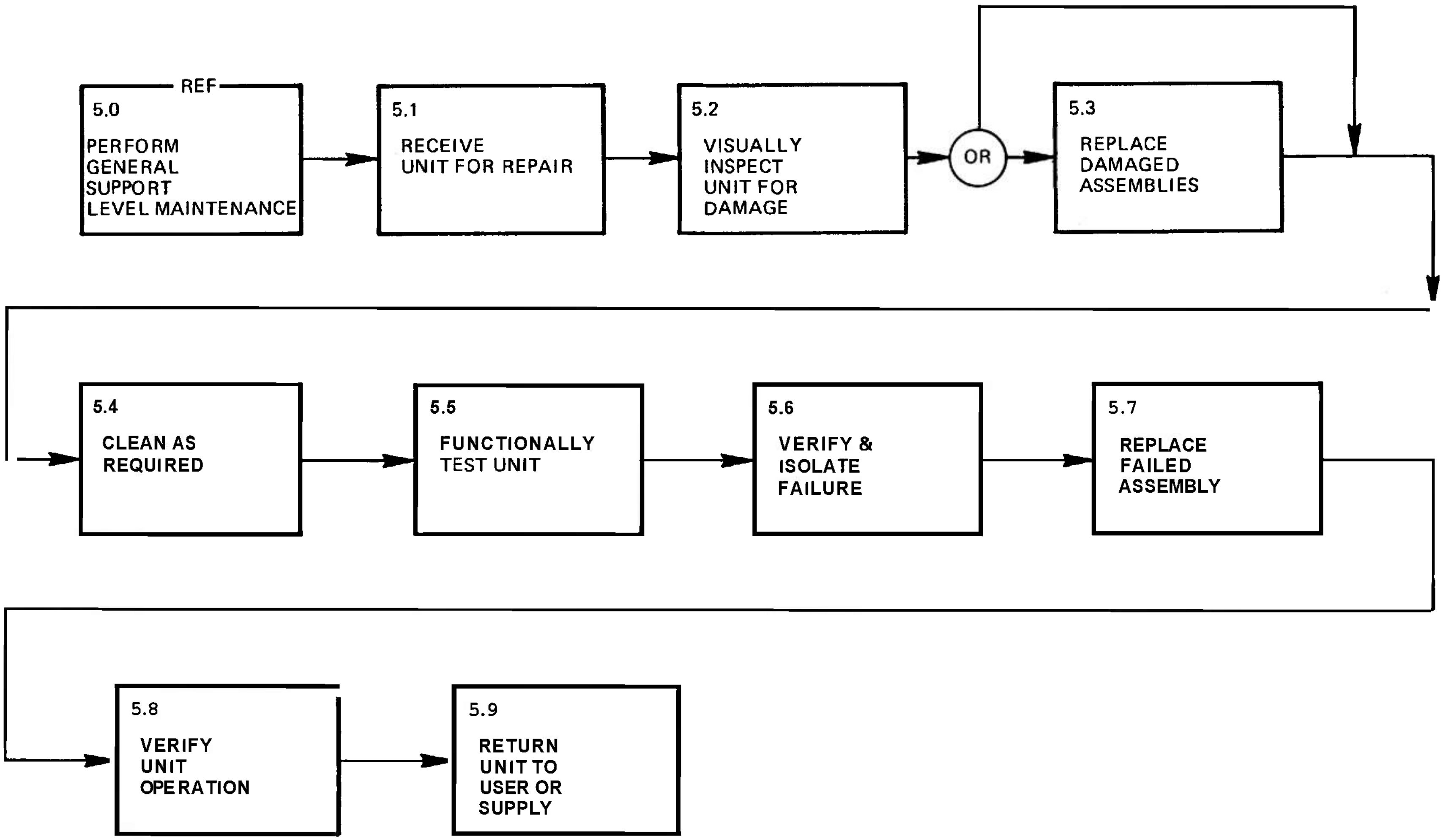


Figure 4-2. Functional Analysis of General Support Maintenance Level

2. Faulty, replaceable component that requires further maintenance and disassembly maintenance of a higher order:

   a. Place a rectangular block above the item, enclosing (identifying) the next higher assembly to be broken down.
   b. Place a circle above the component, enclosing (identifying) the item to be replaced for subsequent repair.
   c. Encircle the faulty component to indicate throw-away replacement.

The connecting lines on the diagram should indicate a physical relationship and not an electrical or mechanical conversion. The electrical or mechanical connection is obtained from an interface or detail drawing of the items, which—coupled with the failure modes-effects-criticality-and detection analysis—furnishes the effects on performance interface. Fig. 4-3 is a typical example of a functional level building block diagram (Ref. 7, p. 64).

When coupling the maintenance action time and the respective failure rates for corrective maintenance or frequency for preventive actions, the effects of the maintenance—as revealed by the building block diagrams—are realized and the optimization of the maintainability goals can be achieved or corrective actions and associated trade-off studies justified, such as changes in the maintainability qualitative design requirements.

As an additional visual aid, an allocation functional system block diagram of the system breakdown could be used to display the related quantitative allocations for each ipdependent level of equipment, where all values which have been extracted from the analysis can be entered in the respective blocks of the breakdown to show the allotments for each block and the summations up to the top system level.

4-2.1.1.3 Failure Modes, Effects, and Criticality Analysis

This may be used as a tool for determining the maintenance requirements. This type of analysis is the basis for determination of the frequency of maintenance and may be used for determination of the qualitative maintainability design features to be incorporated. It also will indicate the need for and the effectiveness of preventive maintenance. Early analysis in the concept and definition phases is essential, even if on a gross system basis, because it affects not only maintainability and reliability but also design characteristics and requirements for new system concepts and planning. A discussion of the failure modes, effects, and criticality analysis formats is not included in this text, since it is readily available in texts on reliability. It is important that the maintainability engineer be aware of these methods, however, and integrate his activity with the reliability engineer in order to receive the full benefits that this analytical tool offers.

In the use of these maintainability analytical tools and methods, the maintainability engineer should give adequate consideration to the time distributions inherent in maintenance actions. Human factors and system complexities have a bearing on a technician's skill and capabilities to handle the maintenance actions and thus on the time distributions. Specific attention should be given to the type, degree, and range of skills available at the various levels of maintenance.

4-2.1.2 Assignment of Maintainability Factors

In par. 4-2.1.1 three analytical tools are described. Once the extent of the utilization of these three tools has been established, the assignment of maintainability factors is made.

If a desired maintainability quantitative goal or constraint has been specified, allocations for the eight functional level system breakdowns are performed and summations made to ascertain that the specified maintainability goal is achievable. The extent of breakdown to the lowest level depends upon the phase of the equipment life cycle; e.g., at the contract development phase allocations may be only possible to the group assembly level for which end-item specifications will indicate the constraints, whereas in the validation phase the analysis would be to the throw-away part level. In cases where the summation indicates that the quantitative goals or constraints are not achievable, further analysis of the design concepts must be performed to effect design changes and provide justifications for changes by associated trade-off studies, as necessary.

In general, allocations are made to the lowest possible breakdown level for which reliable and realistically achievable projections can be made. The specific allocation values are maintenance mean times, median times, and maximum times at given percentiles, as required, and the failure rates for corrective maintenance and frequencies for preventive maintenance. In some instances, maintenance man-hours per operating hours, per miles traveled, or per round fired may be used in lieu of maintenance action times and frequencies. Whatever factors are used, they must be compatible throughout the analysis, not mixing corrective actions and preventive actions and their frequencies in the summations. For each functional level system break-

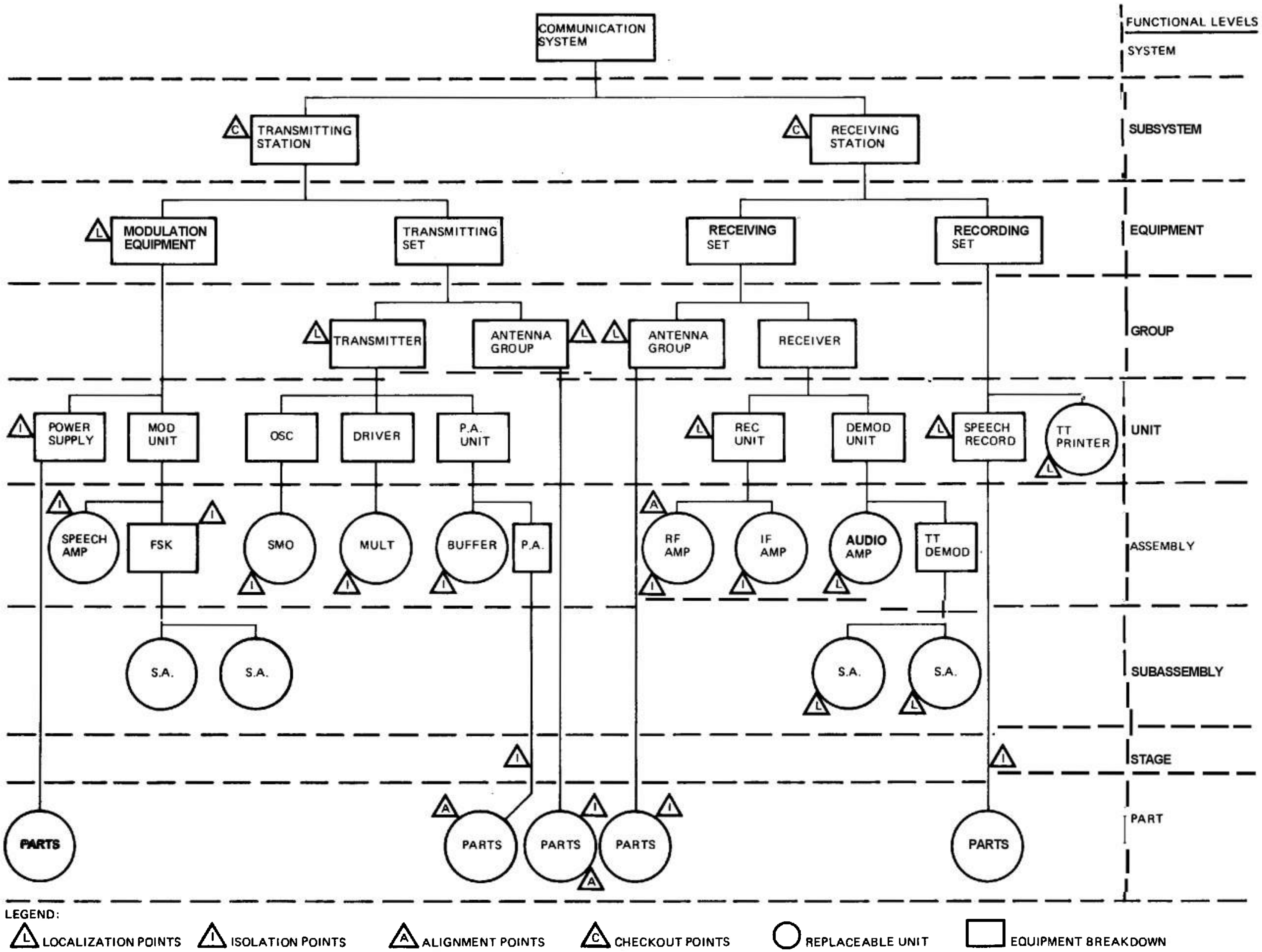


Figure 4-3. Functional-level Building Block Diagram

down, the factors may be listed on the maintenance functional flow block diagrams, or functional level breakdown charts and diagrams, or arranged in other suitable format for ease of the summation process. To perform the summation, Eqs. 4-1 through 4-5 are used, remembering, however, that when an item is subject to preventive replacements at a frequency f_{pi}, its corrective maintenance frequency f_{ci} becomes a function of the preventive maintenance policy (see Ref. 2, Chapters 3 and 4).

Two allocation methods for systems that apply to corrective maintenance only are shown in Ref. 4, Section 3B. The first allocation method is based on system maintainability synthesis; the second method applies to maintainability improvement allocation. These two methods are now discussed.

4-2.1.2.1 System Maintainability Synthesis

First, a building block diagram of the system is developed, with each block representing an independently maintained unit, and a failure rate estimate is assigned to each block, with λ_i being the failure rate of the ith block. The next step is to estimate the mean corrective maintenance downtime $\bar{M}_{cti}$ for the system when the ith block fails. The system estimated mean downtime per failure $\bar{M}_{ct}$ is then

$$\bar{M}_{ct} = \sum_{i=1}^{n} \lambda_i \bar{M}_{cti} \Big/ \sum_{i=1}^{n} \lambda_i \qquad (4\text{-}6)$$

which is the equivalent of our Eq. 4-1.

If the specification imposes a requirement on the system maintenance man-hours per system operating hour, the following allocation equation is used:

$$\bar{M}_h = \sum_{i=1}^{n} \lambda_i \bar{M}_i \qquad (4\text{-}7)$$

where

$\bar{M}_h$ = specified system mean maintenance man-hours per system operating hour (MMH/OH)

$\bar{M}_i$ = estimated mean maintenance man-hours required to repair the system if the ith block fails.

If the specification imposes a requirement on the system median maintenance time per failure, we use

$$\ln \tilde{M} = \left(\sum_{i=1}^{n} \lambda_i \ln \tilde{M}_i \right) \Big/ \sum \lambda_i \qquad (4\text{-}8)$$

where

$\tilde{M}$ = specified median maintenance downtime of the system

$\tilde{M}_i$ = estimated median maintenance downtime of the system if the ith block fails.

The allocations of $\bar{M}_{cti}$ and $\bar{M}_i$ and $\tilde{M}_i$ are made so that the system requirements are met. If the specified values of the system requirements are met, the allocation is considered complete. If the allocations at the block level do not result in meeting the system maintainability requirements, the second method is applied.

4-2.1.2.2 Maintainability Improvement Allocation

Assuming that the maintainability improvement which can be achieved in a given block is directly proportional to the originally allocated or predicted value of maintainability of the block and is independent of the failure frequency of the block we get

$$\bar{M}_{cti} = \bar{M}_{cti} \bar{M}_{ctg} / \bar{M}_{ct} \qquad (4\text{-}9)$$

where

$\bar{M}_{cti}$ = improved allocated mean maintenance downtime of the system when the ith block fails

$\bar{M}_{cti}$ = originally estimated mean maintenance downtime of the system when the ith block fails

$\bar{M}_{ctg}$ = specified system mean maintenance downtime

$\bar{M}_{ct}$ = as defined by Eq. 4-6

When, under the same assumptions, a system MMH/OH requirement is specified, we get

$$M_i = \bar{M}_i M_g / \bar{M}_h \qquad (4\text{-}10)$$

where

M_i = improved allocated MMH/OH when the ith block fails

M_g = system MMH/OH specification requirement

$\bar{M}_i$ = originally estimated system MMH/OH when the ith component fails

$\bar{M}_h$ = as defined by Eq. 4-7

4-2.2 COMPATIBILITY OF MAINTAINABILITY FACTORS

As has been previously stated, total system maintainability factors cannot be determined by summation of individual lower system breakdown factors unless the indices are consistent and the preventive maintenance and corrective maintenance determinations are not combined in an overall summation process.

For instance, maintenance man-hours per operating hour for one item cannot be added to the maintenance man-hours per mile for another item to obtain total man-hours; man-hours and mean time to repair cannot be added together or multiplied together by frequencies of occurrence; frequency for corrective maintenance cannot be added to frequency for preventive maintenance unless the rate base is the same—such as per operating hour, or per mile. Maximum times at given percentiles vary according to the task and are not additive. A critical area is the possible task-step frequency variant. Too often when explaining total time to perform a function, one adds times required to perform the function, without consideration of the different frequencies of the respective task steps. For example, if the adjustment step occurs only 20 percent of the time, then that step is weighted accordingly in the summation.

During demonstration tests there is a tendency to compute total man-hours and mean times by adding the times and dividing by the number of actions or failures. This is only valid when sufficient maintenance samples have resulted from total life test and all the reliability prediction frequency factors have been verified. This is why simulation of maintenance tasks is conducted and the times are multiplied by the expected failure rates to determine whether the maintainability design is accepted or rejected.

Another area of incompatibility is the use of a constant system derating factor for field failure rate usage versus inherent laboratory failure rates. All items do not fail during field use at the same derating factors; the same applies for varying environment such as airborne and ground environments.

On the maintenance engineering analysis sheets, provisions are made for differences in measurement indices—such as operating hours, miles, or rounds. Unless consistent indices are used, the summations are not valid. Another common error is mixing seconds, minutes, hours, and calendar times as individual units, instead of decimals of an established unit.

A critical incompatibility in the summations is the mixing of a task time based on a single sample time or a single expert judgment time, and maximum times with mean times; such summations result in erroneous conclusions which can cause very serious over- or under-design characteristics to be built into equipment.

To make certain of the compatibility of each assigned parameter the maintainability analyst should adhere to the following procedure:

1. Review the maintainability constraints, goals, and objectives and determine the basic indices desired.

2. Define the maintainability quantitative time parameters desired in compatible terms, such as mean times, man-hours, median times, or maximum times at a specific percentile. Select the base to be used in the analysis (mean time is the most commonly used and is easily handled for various types of distributions and summations).

3. Determine the time unit reference base, such as seconds, minutes, hours, calendar time, or operating period. Prepare a conversion table of the various times to the base selected.

4. Define the frequency baseline, such as mean time between actions (which is 1/frequency in hours, calendar time, operating hours, miles, rounds, etc.); failure rates per hour per 100 hr, per mile, per round; and scheduled maintenance rates per calendar time, operating hour or hours, rounds, miles. Prepare a conversion table for the base selected, making certain it is compatible with Item 3.

5. Define frequency derating factors for inherent, laboratory, peace-time, and/or war-time environments. Select the base in accordance with the quantitative constraints imposed. Note: a single derating factor is not compatible with various types of hardware at different operating environments.

6. In establishing the maintenance action mean times, etc., make certain that the data are based on a sufficient sample size to show the variance and confidence level desired.

7. When using mathematical formulas and associated statistics, review the definitions and terms to make certain the parameters are consistent and compatible with the baselines previously established. Apples and oranges cannot be added, subtracted, multiplied, divided, integrated, or analyzed as a common item.

4-2.3 STATISTICAL INTERDEPENDENCE

In developing maintainability quantitative parameters and associated maintainability design characteristics, the reliability quantitative and qualitative design features and characteristics must be kept in mind. Accomplishment of the maintainability objectives is solely dependent upon the frequency for a maintenance action, which is definitely determined by reliability, both for unscheduled random failure rates and scheduled rates for preventive maintenance actions. An old saying among reliability engineers is that if equipment doesn't fail, there is no need for maintainability; maintainability engineers answer "Show me the equipment that has never failed".

The reliability engineering failure modes, effects, and criticality analyses define the specific needs for maintainability design and for detecting failures. This is not only a real interdependence, but is a statistical relationship. This relationship exists in the distribution of times to failure and the wearout statistics for preventive maintenance replacements. It reflects distribution of maintenance action times due to variations in equipment complexity and human skill. Human factors in relation to equipment maintenance are of considerable concern in reliability. Therefore when (a) allocating and predicting the maintainability quantitative parameters of time to perform maintenance, (b) establishing the qualitative design features, and (c) the determining of technician skill requirements, repair parts, and other integrated logistic support needs, the statistical interdependence of reliability and maintainability must be considered.

For example, when the maintainability engineer makes his statistical predictions and allocations, he is dependent on failure rate statistics; if his analysis shows he has exceeded his constraint, he must first look to reliability to determine how the failure rate can be reduced by design trade-off. In the same fashion, when the failure rate is excessive, the reliability engineer looks to maintainability to determine how the downtime can be reduced by design trade-off. Both are vitally concerned with availability because the availability ratio depends upon "uptime" (reliability) and "downtime" (maintainability). Historically, newly developed complex systems have been designed with low *MTBF*'s, and logistic support costs have risen; this indicates that the reliability-maintainability interdependence has not been realistically considered and meshed. The new trend in the Army is to pay greater attention to this interdependence by combining the two disciplines in an RAM Integrated Program Plan (Reliability, Availability, Maintainability) from the first allocation of system requirements.

SECTION III

MAINTAINABILITY PREDICTION

4-3 GENERAL

While maintainability allocation is initially made to major functional building blocks of a proposed system in order to serve as a guideline for detailed design of the maintainability features of each block (such as subsystems) maintainability prediction is concerned with the quantitative estimation of the maintainability parameters of specific design configurations to determine whether or not such configurations have the potential to meet the maintainability specification requirements and, where necessary, to identify maintainability problem areas in the design which require changes.

Maintainability prediction evaluates designs as to their effect on system maintenance and repair, the associated downtimes, maintenance labor, repair parts, and, ultimately, maintenance costs. The prime purpose, however, is to predict the maintenance time parameters of the design from the qualitative design features, considering the time elements involved in performing maintenance actions—namely, preparation, diagnosis, replacement, adjustment, servicing, check-out, and failed item repair if the item is not of the throw-away type. Also, prediction is concerned with all applicable maintenance levels (organizational, direct support, general support, and depot) since it is not only the system downtime which is of importance for system availability, but also the maintenance man-hours and repair parts expended at the lower maintenance levels to keep the system operational. A detailed description of the maintenance task time elements and maintenance levels is presented in par. 4-2.1

To predict the maintainability quantitative parameters or figures of merit—such as $\overline{M}_c$, $\overline{M}$, M_{MAX}, MMH/OH—availability we must estimate the maintenance task time elements, synthesize these into estimates of maintenance action duration for each kind of failure or repair action (including preventive maintenance actions), obtain frequency of occurence estimates (such as failure rates), and then through mathematical models or other techniques (such as graphic methods) evaluate the quantitative figures of merit mentioned. Not all of these need to be evaluated on each occasion.

Depending on the choice of these figures of merit and on the developmental phase of the system, one may use different prediction techniques. Distinct prediction techniques can emerge for phases of system life throughout the life cycle.

Predictions from the operational phase are usually most accurate if the data base of the operational records is statistically reliable. One may question the usefulness of predicting during the operational phase; however, such prediction will provide estimates for next year's performance at required confidence levels and is also very useful in comparing the operational results with predictions made in the previous phases of the system life cycle, just to evaluate their accuracies and to gain the experience.

The usefulness of prediction in the early life-cycle phases is obvious, since it is at such times that the main features of maintainability should be incorporated in the basic design to avoid costly redesigns, schedule slippages, and even big flops. The maintainability, along with the reliability, designed into a system has an immense impact on the operational availability and life-cycle costs of the hardware to be built. Because of this impact, maintainability prediction must be applied as an iterative process to all phases of the system life cycle to detect any shortcomings and to perform corrective actions at the earliest possible time. This is the most economical approach.

4-4 DEVELOPMENT OF A MAINTAINABILITY CRITERION

4-4.1 BASIC ASSUMPTIONS

An often made basic assumption is that recorded reliability and maintainability data previously obtained from comparable systems and components operating under similar conditions are "transferable" and can be used to predict the maintainability of new designs of comparable systems (Ref. 3, p. 2). This assumption, along with other assumptions that follow, requires a critical and very careful evaluation in each instance.

Before the numerical maintainability prediction starts, there should be a "gathering of facts" phase to substantiate or weigh the assumptions as to their acceptability, justification, and applicability to a new program or to disallow them. Some of the assumptions to be evaluated as to their applicability are:

1. Historical data from existing similar equipment
2. Maintenance levels at which repairs are performed
3. Established maintenance task time elements tables
4. Human factor tables and standards for technician skills
5. Traditionally used statistical distributions
6. State-of-art tools and test equipment
7. Existing logistic support system
8. Cost estimating procedures and cost rates
9. Personnel skill populations
10. Mathematical models
11. Time-to-repair indices and characteristics
12. Failure rates
13. Maintenance task sequences
14. System operational profiles
15. Prediction methods.

Many of these and other assumptions may have been defined during the concept formulation stage, many may need expansion, and some may not exist. In all cases, in the "data gathering stage", the validity must be justified and/or analyzed to show justification for the assumption. The key point is that assumptions must be stated in order to show the baselines for the predictions. There is no limitation imposed upon assumptions, but if not stated, the results and validity of predictions are left open to question. Assumptions that are not validated are then used to define risk areas that may need further definition and acknowledgment if they are critical to the relization of the maintainability objectives and goals.

4-4.1.1 Time-to-repair as an Index

As an example of an assumption, the time to repair is an index of the maintainability quantitative criterion upon which the maintainability qualitative design feature requirements are established. Conversely, the qualitative design features built into equipment establish both the time-to-repair quantitative criteria and the associated technician skills and integrated logistic life-cycle support requirements. Therefore, the quantitative time to repair is the prediction index upon which the maintainability objectives are based. Prediction studies in complex equipment always encounter the problem of time-to-repair criteria. In fact, this term refers to the collection of logical and empirical issues involved in establishing some class of data as a standard for the performance of the maintenance function and its associated maintenance tasks.

The principal feature of the time-to-repair criterion is its relevance. Logical relevance is established through a network or a maintenance functional flow diagram and equipment functional level process. In forming and applying this logical sequence, it is desirable that each step in the chain back to the system objectives be clearly traceable so that separate steps can be explicated and extended. The explication can be tested for logical relevance and thereby enhance (or detract from) the acceptance of the functional flow as defining the criterion series. Practical considerations sometimes result in biased allocation, imperfect discrimination between absolute criteria and noncriteria facts, and less than univocal scoring weights. Recognition of such facts must be stated in the analysis assumptions, but they do not reduce the requirement for logical relevance.

Reliability of the variable criteria must be established. There must be some regularity in the criteria or else the series will consist of random numbers and would be unpredictable in principle. Reliability is defined in terms of prevailing situations; there may be no fixed number, although numerical estimates over a wide range of maintenance functions and associated tasks may be consistent to encourage generalization. The most meaningful reliability is one associated with measurement of times and environments by selected data observation, collection and evaluation—ordinarily, a simple correlation between one observed system and another during the same period of time will provide a satisfactory estimate of the reliability of the prediction—or some elaborate statistical designs can be employed to define the "error of variance" or unreliability due to the several sources of variation.

A subproblem in establishing the time-to-repair criteria involves the "statistical" distribution of the indices representing the functional flow series. If the criteria consist of discrete states, then some fair portion of the total frequency must be registered for each "state category". If the criterion variable is continuous, then there should be a satisfactory spread of values. There is not much use in predicting mean repair times of different equipment configurations if the means are all about equal. The important principle is that the criterion distribution be regular enough to be specifically meaningful and manageable.

Closely associated with distribution characteristics is the "sampling" problem. Any criterion variable might be incomplete, but it is often possible to show that it is unbiased, in that the events included in the tabulation are sufficiently representative of the total collection of events. The sampling problem is seldom mentioned in the reporting of maintainability prediction methodology, but this does not reduce its significance.

The understanding of the variable criterion also helps to explicate the reliability of a specific prediction. The important concern is to show that the prediction is of demonstrable significance. When the measurable time-to-repair is implemented by competent analysis, it should lead to practical and "real life" changes where the effects are generally positive, and the criterion should gradually assume a more important role in the determination of the system quantitative parameters. It is important that "intrinsic" maintainability criteria be immersed in realistic field operational environment performance. This invariably introduces large variances in the maintenance performance function which must be recognized, analyzed, and justified. The adequate state-capability assumption and estimation must be included, especially in the early formulation of the system design, because it cannot be satisfactorily introduced later.

To many maintainability analysts, the preceding remarks about the repair time criterion attributes may seem unnecessary. The military field commander wants some clear assurance that a "Repair or Maintenance" function to restore an equipment to operational status will be completed within the downtime constraint and will assure the system availability. Therefore, the time-to-repair criterion is intuitively correct and has been accepted by the military as "the design for maintainability index" which is formalized in military specifications. The index is consonant in implied viewpoint and structure with other modeling for system operational effectiveness modeling and for life cycle costs. It also has a direct relation to the qualitative characteristics of system design for maintainability. Thus, the criteria are essential attributes to be exercised by the maintainability analyst in the establishment of his predictions and allocations (Ref. 5).

4-4.1.2 Time-to-repair Characteristics

The prediction assumptions concerning time-to-repair characteristics are defined as:

1. Those design features of the system that cause or enable a maintenance technician to perform the tasks of a maintenance function needed to restore an equipment to "operational ready for use" status

2. Those human factors characteristics of a maintenance technician which enable him to perform the tasks of a maintenance function

3. Those design and operational characteristics of maintenance support equipment (built-in or auxiliary) and the associated integrated logistic support elements which assist the technician to perform the tasks of a maintenance function.

In each of these assumptions there are statistical attributes with associated relevance, reliability, and certainty factors of the time-to-repair distribution. The mean-time-to-repair parameter for the "maintenance function" of a system "functional level breakdown" is derived from these assumptions, in order to incorporate these in the determination of the next higher functional breakdown time to repair. The relevance, reliability, statistical attributes, demonstrable significance, and certainty are discussed briefly in the previous paragraph and in detail in Ref. 5. In applying the characteristic assumptions to prediction, the analyst must be aware of the significant interrelationship of these assumptions (see par. 4-2). Also, the relationship of repair times and frequency of repairs must be evaluated. The characteristic assumptions must be recorded at each step in the analysis so that the relevance of the logical analyses can be justified and substantiated. Of special importance are the design features of the system that affect the time decision. The extrapolation of these characteristics from "inherent" state to the "field operation environment" must be considered and justified. The analyst must be aware of the interface of maintainability objectives with other design performance objectives and should utilize trade-off procedures to resolve areas of conflict which serve as the basis for predictions. Also, his predictions may result in additional maintainability features that are designed to help achieve the established prediction goals which will involve additional trade-off study. When using design feature checklists, the attribute relevance criterion for scoring must be exercised in the same fashion as in methods of equipment sampling.

The statistical principles pertaining to maintainability are discussed in Chapter 8, "Statistical Maintainability".

4-4.2 PREDICTION ELEMENTS

There are two prime elements in prediction:

1. The combination of failure rates (corrective maintenance functions) and scheduled rates (preventive maintenance functions)

2. Repair times (corrective and preventive maintenance functions).

The prediction analyst must analyze the two maintenance functions separately; they may not be mixed.

The constant failure rate (random failures used in making corrective maintenance frequency computations) lies between early mortality failures (burn-in failures for electronic-electrical items and/or wear-in failures for electrical-mechanical and mechanical items) and the increase in failures due to wear-out. The scheduled frequency of preventive maintenance is that period between maintenance functions when maintenance must be performed in order to avoid an increase of wear-out and resulting failure. The determination of early mortality rate is a quality control function to insure that the items are beyond the possibility of early mortality before being installed in systems; e.g., wearing-in of an assembled engine or the burning-in of electronic items. For purposes of predicting, such early mortalities are assumed to have been eliminated. This assumption must be justified by the analysts by such techniques as reviewing the quality assurance specifications of the items being analyzed (including those items that will be used to replace failed items in the field environments). In establishing the rates, the analysts must consider the rates as a function of the use and the environment, correlating them per unit of time. The unit of time must be constant in the summation process used. The rates can also be utilized in applicable regression equations to calculate maintenance action times. In addition, rates are used in all predictions to weight the repair times for the various categories of repair activity, thereby providing an estimate of its contribution to the total maintenance time.

The repair times are broken up into the basic "maintenance action" tasks whose times are summed to obtain the total time for a repair action. In most cases, the task times are summed, without regard to frequency, in a single repair time function, because the reason and need for maintenance repair are constant for a single logical repair action. The analyst must use caution in summing the task steps, because in some instances the steps may need to be repeated with a certain probability; for example, a repeat of the fault diagnosis and/or check-out for certain types of malfunctions may be necessary for a certain portion of the events. Also, during the sequence tasks the sequence may need repeating when two or more items may have caused the malfunction. This also applies to maintenance-induced faults. Where different "repair actions" of varying frequencies are involved, the mean repair times are summed using the frequency of contributions of the individual maintenance repair actions at the respective levels of maintenance.

In the use of the various prediction methods, the "principle of transferability" can be justified when the degree of commonality between systems can be established. Usually during the early concept and design phase, commonality can only be inferred on a broad basis depending upon the relevance, reliability, and accuracy of the historical data (extrapolation of total population use and/or specific observation of usage). However, as the design becomes refined during later phases of the life cycle, commonality is extended if a high positive correlation is established with regard to equipment functions, to maintenance task times, maintenance frequencies, and to levels of maintenance (maintenance concept). When using the principle of transferability to establish correlation and commonality, one must always consider the statistical parameters of the maintenance functions, the repair time distribution, and the frequency of occurrence (mean, median, maximum at given percentiles, standard deviation, and confidence limits).

4-4.2.1 Failure Rates—Scheduled Maintenance Rates

Measures of corrective maintenance rates are in terms of mean time between failures ($MTBF = 1/$failure rate, expressed in terms of failures per hour, mile, round, etc.). Time and other basic constants vary among many historical records and data collection processes, as well as the assumptions of the various operating conditions, environments, and rating factors. Therefore, the analyst must evaluate the basis for the rate and be consistent in its use during the analysis. The simplest and most widely used rate in a maintainability analysis is the reliability term A or f_c (failures per hour, mile, round, etc.). In the early 1950's the reliability term was λ expressed in failure per 100 hr.

Preventive maintenance frequency measures are in terms of mean time between preventive (scheduled) maintenance actions ($MTBPM$) expressed in terms of hours, rounds, miles, etc. The times vary among historical records and reliability analyses, and there are differences in the assumptions made concerning various operating conditions, environments, and rating factors. There has been ongoing reference to calendar times, which seems to infer that maintenance must be performed whether the need exists or not. The analyst must be aware that the need for preventive maintenance should be based on the objective, namely, the prevention of wear-out and/or resulting failures. Examples are lubrication, cleaning, and calibration under operating conditions. Should oil be changed according to calendar times or miles of use whether or not it needs it? If an oil sampling shows no degradation or contamination, why should it be changed? The Army is presently considering using a simple scheduled sampling and analysis to be performed by the organizational technician. Therefore, the maintainability analyst must justify the scheduled rates for preventive maintenance in terms of actual usage rates. The simplest index for rates is ***MTBPM*** in hours of use stated in terms of hours, miles, rounds, etc. The rate should use the same basic units as corrective maintenance; this allows a simple correlation with the same time base that can be mathematically handled to determine the overall system availability at any one level of maintenance for the combination of the two maintenance functions.

4-4.2.2 Repair Time

MIL-STD-721B (Ref. 9, Fig. 1), shows the relationship of the various maintenance function times and their respective references to active uptimes and downtimes. Maintenance time and repair time are synonymous, being a relation of the two maintenance functions (corrective and preventive):

$$\bar{M} = (\bar{M}_c f_c + \bar{M}_p f_p)/(f_c + f_p) \qquad (4\text{-}11)$$

where

$\bar{M}$ = mean time for both preventive and corrective maintenance
f_c = corrective frequency
f_p = preventive frequency

$$\bar{M}_c = \sum_{i=1}^{n} M_{ci} f_{ci} \Big/ \sum_{i=1}^{n} f_{ci} \text{ and } \bar{M}_p = M_{pi} f_{pi} \Big/ \sum_{i=1}^{n} f_{pi} \qquad (4\text{-}11a).$$

where

M_{ci} = corrective task time.
M_{pi} = preventive task time
f_{ci} and f_{pi} = associated frequencies (f = $1/MTBF$ or $1/MTBPM$, respectively)

In the use of terms, the analysts must be alert to the use of synonyms and labels by the maintainability profession, and the use of letters by the mathematicians. For instance, in mathematical communication, ***MTTR, MTBF, MTBPM, MTBM*** symbolize products of terms *M, T, R, P, F*, etc., whereas to the maintainability engineer they symbolize specific mean times.

In any maintainability prediction method the repair times are essential elements in developing the prediction analysis and are expressed in seconds, minutes, or hours. The simplest base is in terms of hours and decimals of hours. The repair times are determined for each of the task steps based on the complexities of the maintainability design features and characteristics, the associated complexity of human factors, and the interrelation of the logistic support equipment.

4-5 PREDICTION METHODS

There are many maintainability prediction techniques presently in use by industry and Government organizations. The procedures vary according to the specific reason for measurement, imposed requirements, peculiarities or similarities of system being evaluated, and the individual preferences of the agencies involved. It is not the purpose of this handbook to discuss the details of the basis for these predictions but rather to describe the overall pattern and interpretation, leaving the details to the cited references.

One of the prime considerations in choosing a specific method is to recognize the limitations of the different methods and the constraints imposed by the type of equipment and its use. Each maintainability group in a particular industry or military activity should develop a usable methodology based upon the historical data for the equipment of interest. Coupled with the operational use environment and the maintenance concept plan, a statistical background can be established upon which the essential time and frequency factors can be used in the respective prediction methods. Some of the more important prediction methods are discussed in the paragraphs that follow.

4-5.1 EXTRAPOLATION METHODS

Extrapolation is the process of inferring or predicting beyond known information to an area that is, to some degree, unknown. As applied to the field of maintainability, extrapolation is concerned with predicting maintenance characteristics of new equipment from its design features and from observed relationships between design features and maintenance characteristics of existing, similar equipment. The amount of uncertainty inherent in the extrapolation depends on the degree to which the new equipment differs from existing equipment, and on the precision with which the relation between design features and maintenance characteristics for the existing equipment is known.

4-5.1.1 Prediction By Smoothing

The term "smoothing" as applied to a sequence of data means to remove irregularities by fitting a smooth curve through the data points. This curve presumably averages out the "noise" or random disturbance present in the data and represents the underlying process or trend inherent in the data.

Once the curve is available it can be used predictively. For example, if the data were measurements made at time 0, 1, 2,. . .,t, a curve fitted to these data could be used to predict what the measurement would be at time $t + 1$, $t + 2$, etc.

It must be emphasized that such extrapolation is a logically hazardous process. Even if the smoothing function fits the empirical data very well, there is no guarantee that the pattern exhibited by these data will carry on into the future. The further beyond the data one seeks to predict, the more prone to error he is likely to be. The extrapolation is, of course, all the more risky if the fit of the curve to the data is poor. A classical, comprehensive source on curve fitting is Milne, Ref. 10.

A related smoothing technique is that of exponential smoothing. (See Brown and Meyer, Ref. 11.) This is an iterative technique in which the smoothed value at time t is formed by taking a weighted average of the actual observation at time t and the smoothed estimate at time $(t - 1)$. Algebraically, letting x_a denote the actual observation at time a, and $\bar{x}_a$ the smoothed estimate at time a, the method of exponential smoothing states

$$\bar{x}_t = \alpha x_t + (1 - \alpha)\bar{x}_{t-1} \tag{4-12}$$

where α, the "smoothing constant", is a fixed number such that $0 < \alpha < 1$. By substituting Eq. 4-12 into itself repeatedly, one can express $\bar{x}_t$ in terms of $x_t, x_{t-1}, \ldots, x_0$. The first step is

$$\begin{aligned}\bar{x}_t &= \alpha x_t + (1 - \alpha)[\alpha x_{t-1} + (1 - \alpha)\bar{x}_{t-2}] \\ &= \alpha x_t + \alpha(1 - \alpha)x_{t-1} + (1 - \alpha)^2\bar{x}_{t-2}\end{aligned} \tag{4-12a}$$

Repeating this process yields

$$\bar{x}_t = \alpha\sum_{j=0}^{t-1}(1 - \alpha^j)x_{t-j} + (1 - \alpha)^t x_0 \tag{4-13}$$

where t is assumed here to be a positive integer. (This corrects Eq. 3 on p. 675 of Brown and Meyer, Ref. 11.) Eq. 4-13 shows where the adjective "exponential" originates. It also indicates a property of exponential smoothing, namely, that the immediate past plays a larger role in the prediction than does the more remote past. This follows from the fact that $0 < \alpha < 1$, also $0 < 1 - \alpha < 1$, and successively higher powers of a number between 0 and 1 decrease to zero.

The paper by Brown and Meyer discusses "higher order" smoothing (that given by Eq. 4-12 being of first order) and the choice of the smoothing constant α, among other matters. It also provides an example of triple (i.e., third-order) smoothing.

A brief discussion of higher order smoothing is appropriate. Eq. 4-12 can be written in the form

$$S_t(x) = \alpha x_t + (1 - \alpha)S_{t-1}(x) \tag{4-14}$$

The smoothing operator of order n is defined by

$$S_t^n(x) = S[S_t^{n-1}(x)] = \alpha S_t^{n-1}(x) + (1 - \alpha)S_{t-1}^n(x) \tag{4-15}$$

with zero-order smoothing defined by

$$S_t^0(x) = x_t \tag{4-16}$$

This convention makes Eq. 4-15 consistent with Eq. 4-14.

Finally, the fundamental theorem referred to in the title of Brown and Meyer's paper should be mentioned. It states that if one predicts $x_{t+\tau}$ by a polynomial degree N in τ, then the coefficients $a_0, a_1, \ldots, a_N$ of this polynomial can be estimated as linear combinations of values obtained from the first N + 1 degrees of smoothing applied to a time series $\{x_0, x_1, \ldots, x_t\}$ with observations equally spaced in time. (Brown and Meyer give explicit formulas for these coefficients for N = 0, 1, 2.) In sum, using equally spaced data through time t and coefficients $a_0, a_1, \ldots, a_N$ based on N + 1 exponential smoothings of these data, one can predict $\hat{x}_{t+\tau}$, $\tau = 1, 2, \ldots$, by a polynomial of degree N:

$$\hat{x}_{t+\tau} = a_0 + a_1\tau + \cdots + a_N\tau^N \tag{4-17}$$

the tilde ($\hat{}$) designates a predicted value.

4-5.1.2 Prediction By Assuming Distribution Characteristics

If one knew the precise form of the distribution of the time to repair an item, he could obtain, with no uncertainty, any desired property of this distribution—e.g., the mean, variance, and other moments of repair time; the percentiles of the repair time distribution; or the probability that repair will exceed, or be concluded within, a given time. However, this knowledge is rarely, if ever, available. Instead one must assume that the repair time distribution belongs to some family of distributions (e.g., exponential, lognormal, gamma), and then estimate its parameters, percentiles, etc. on the basis of a sample of values.

On what basis can one assume that the time to repair a given item is distributed according to some distributional family? First, if one has certain data that he thinks come from some class of distribution function, he can test these data for fit to this class. There is a wealth of material available giving such tests. Some references are: testing for exponentiality, Epstein (Ref. 12) and Lilliefors (Ref. 13); testing for lognormality, Aitchison and Brown, (Ref. 14); testing for gamma, Mickey, et al. (Ref. 15).

Second, the nature of the repair activity may yield a clue as to the distribution of repair time. For example, Goldman and Slattery (Ref. 16, p. 46) state that the distribution of downtime (i.e., active maintenance time) tends to be exponential for "equipment that requires relatively frequent adjustments of very short durations or which may be put back into service via a quick remove and replace operation. Occasionally, much longer times may be required for major repair or spares". They go on to say "The lognormal distribution describes the downtime for a wide variety of reasonably complex equipments. This distribution is useful in describing the situation where there are few downtimes of short duration, a large number of observations closely grouped about some modal value, and a not insignificant number of long downtimes". They further state (p. 45) that the gamma distribution "is receiving increasing attention as a substitute for the lognormal because of algebraic simplicity".

4-5.1.3 Nonparametric Statistics

In the preceding subparagraph we touched on analyses requiring the assumption of a particular class of distribution function or the repair time random variable. In this subparagraph, we discuss methods appropriate when no specific assumptions (other than the continuity of the random variable) concerning this distribution function are made. Such methods are termed "distribution-free" or "nonparametric". The latter term comes from the fact that, for a given family of distributions, specifying a set of parameters will uniquely identify a distribution ***within that family***, but not generally. Thus, the term "distribution free" may be more descriptive of what is discussed in this subparagraph than the term "nonparametric".

Consider a sample of size n of times to repair some item. No assumption will be made concerning the distributional family to which the repair time random variable belongs. We will show how to estimate the distribution function of repair time and percentiles of this distribution function.

We remark that in this situation it is not appropriate to estimate moments, e.g., the mean and variance, for there is no guarantee that they exist. Instead of the mean, one estimates the median (the 50th percentile of the distribution function). Instead of the variance, one estimates some other measure of variability such as the interquartile distance (the difference between the 75th and 25th percentiles).

4-5.1.4 Estimation of the Distribution Function

Suppose that a sample of size n is taken, the observations are ordered from smallest to largest, and denoted: $X_1 < X_2 < \ldots < X_n$. (These ordered observations are called the ***order statistics.***) An estimate of the underlying cumulative distribution functions $F(x)$ is given by the ***empirical cumulative distribution function*** $F_n(x)$ defined by

$$F_n(x) = [\text{number of } (X_1, X_2, \ldots, X_n \leq x] / n \tag{4-18}$$

The empirical cumulative distribution function has the value zero for x less than the smallest observations, $1/n$ at the smallest observation and up to (but not including) the second ordered observation, $2/n$ at the second smallest observation and up to (but not including) the third smallest observation, etc. At and beyond the largest observation the empirical distribution function has the value 1.

A small example will illustrate the construction of the empirical cumulative distribution function. Suppose a sample of size 5 is taken and the ordered values are 1, 3, 4, 7, 12. The empirical cumulative distribution in this instance is

$$F_5(x) = \begin{cases} 0 & x < 1 \\ 1/5 & 1 \leqslant x < 3 \\ 2/5 & 3 \leqslant x < 4 \\ 3/5 & 4 \leqslant x < 7 \\ 4/5 & 7 \leqslant x < 12 \\ 1 & x \geqslant 12 \end{cases} \quad (4\text{-}19)$$

Its graph is shown in Fig. 4-4. The dots on the graph at the jump points indicate that value of the function at these points.

Generally, for a sample of size n, the empirical cumulative distribution function can be expressed as

$$F_n(x) = \begin{cases} 0 & x < X_1 \\ k/n & X_k \leqslant x < X_{k+1}, \quad k = 1, \ldots, n-1 \\ 1 & x \geqslant X_n \end{cases} \quad (4\text{-}20)$$

A distribution-free confidence contour can be given for the "true", underlying distribution function $F(x)$ based on the empirical cumulative distribution function $F_n(x)$ and some constants $\epsilon(n,\alpha)$ related to the "Kolmogorov-Smirnov" statistic. With a confidence of $1 - \alpha$, $F(x)$ will lie within the band $F_n(x) \pm \epsilon(n,\alpha)$. A table of $\epsilon(n,\alpha)$ is given in Table 4-1.
Also, using this table, an 80% confidence contour is constructed for the data of the previous example.

The confidence band is obtained as $F_5(x) \pm \epsilon(5,0.20) = F_5(x) \pm 0.446$, and taking account of the fact that a cumulative distribution function must be non-negative and cannot exceed one. Thus, the upper part of the confidence contour $U_5(x)$ is given by

$$U_5(x) = \begin{cases} 0 + 0.446 = 0.446, & x < 1 \\ 0.2 + 0.446 = 0.646, & 1 \leqslant x < 3 \\ 0.2 + 0.646 = 0.846, & 3 \leqslant x < 4 \\ 0.2 + 0.846 = 1.046^*, & x \geqslant 4 \end{cases}$$

* cannot exceed 1

(4-21)

and the lower part of the confidence contour $L_5(x)$ is given by

$$L_5(x) = \begin{cases} 0, & x < 4 \\ 0 + 0.154 = 0.154, & 4 \leqslant x < 7 \\ 0.2 + 0.154 = 0.354, & 7 \leqslant x < 12 \\ 0.2 + 0.354 = 0.554, & x \geqslant 12 \end{cases} \quad (4\text{-}22)$$

A graph of $U_5(x)$ and $L_5(x)$ is shown in Fig. 4-5.

The reader may consider this confidence contour to be quite broad—as, indeed, it is. However, it must be remembered that this band is based on a sample of only 5 observations. Also, no assumption concerning the form of the distribution being estimated has been made. The less one assumes, the "fuzzier" his estimates will be. Note in Table 4-1 that, for each value of α, the values $\epsilon(n,\alpha)$ decrease with with n. In the limit $\epsilon(n,\alpha)$ is inversely proportional to the square root of n.

4-5.1.5 Estimation of the Population Median

The population median ν is that value of the random variable which divides the range of the cumulative distribution function into two equal parts. Let X denote the random variable of interest and suppose that its cumulative distribution function $F(x)$ is continuous and monotonically increasing. Then ν has the property:

$$P(X \leqslant \nu) = P(X > \nu) = 1/2 \quad (4\text{-}23)$$

A point estimate for ν is the middle order statistic if the sample size is odd, or the average of the two middle order statistics if the sample size is even. That is, letting $\hat{\nu}$ denote an estimate of ν,

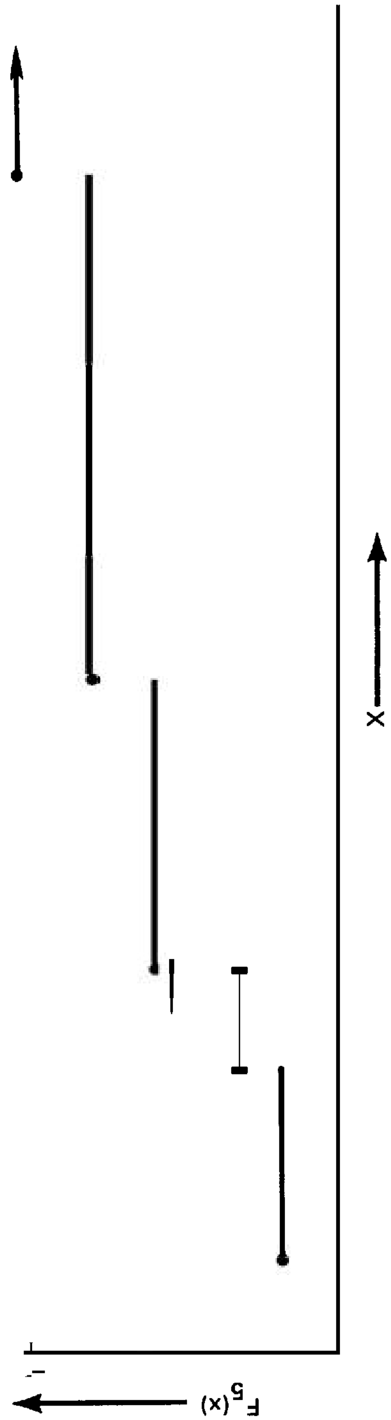


Figure 4-4. Empirical Cumulative Distribution Function

TABLE 4-1.
VALUES OF $\epsilon(n,\alpha)$

Sample size n	.20	.15	.10	.05	.01
1	.900	.925	.950	.975	.995
2	.684	.726	.776	.842	.929
3	.565	.597	.642	.708	.829
4	.494	.525	.564	.624	.734
5	.446	.474	.510	.563	.669
6	.410	.436	.470	.521	.618
7	.381	.405	.438	.486	.577
8	.358	.381	.411	.457	.543
9	.339	.360	.388	.432	.514
10	.322	.342	.368	.409	.486
11	.307	.326	.352	.391	.468
12	.295	.313	.338	.375	.450
13	.284	.302	.325	.361	.433
14	.274	.292	.314	.349	.418
15	.266	.283	.304	.338	.404
16	.258	.274	.295	.328	.391
17	.250	.266	.286	.318	.380
18	.244	.259	.278	.309	.370
19	.237	.252	.272	.301	.361
20	.231	.246	.264	.294	.352
25	.21	.22	.24	.264	.32
30	.19	.20	.22	.242	.29
35	.18	.19	.21	.23	.27
40				.21	.25
50				.19	.23
60				.17	.21
70				.16	.19
80				.15	.18
90				.14	
100				.14	
Asymptotic Formula:	$\frac{1.07}{\sqrt{n}}$	$\frac{1.14}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

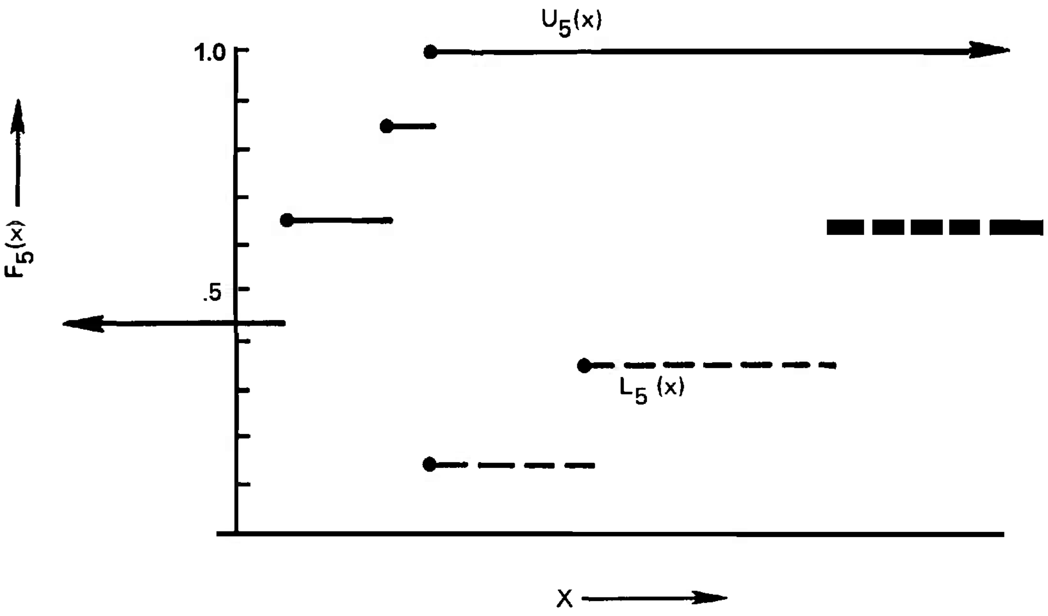


Figure 4-5. Confidence Band for True, Unknown Cumulative Distribution Function

$$\hat{\nu} = \begin{cases} X_{(n+1)/2}, & \text{if } n \text{ is EVEN} \\ [(X_{n/2} + X_{n/2+1})]/2, & \text{if } n \text{ is ODD} \end{cases} \tag{4-24}$$

An interval estimate for ν is also available. It makes use of two order statistics as lower and upper end points of the confidence interval, X_r and X_s, with $r < s$, respectively. We have (see, e.g., Lindgren, Ref. 17, p. 413)

$$P(X_r < \nu < X_s) = \sum_{k=r}^{s-1} \binom{n}{k} (1/2)^n \tag{4-25}$$

so that the confidence coefficient associated with the interval (X_r, X_s) as a confidence interval for ν can be evaluated by the sum of binomial probabilities, Eq. 4-25.* Often r and s are chosen to be equally spaced from the "bottom" and "top" of the sample, i.e., $s = n - r + 1$. In this case

$$P(X_r < \nu < X_{n-r+1}) = 1 - 2\sum_{k=0}^{r-1} \binom{n}{k} (1/2)^n \tag{4-26}$$

The population median is the 50th percentile of the distribution function. More generally ξ_p the 100 pth percentile of the distribution function, is defined by

$$P(X \leq \xi_p) = p \tag{4-27}$$

That is, 100 percent of the probability mass lies to the left of ξ_p.

A point estimate for ξ_p can be given. Denoting by $[np]$ the greatest integer less than or equal to np, an estimate of ξ_p is given by the $[np] + 1$ order statistic, provided np is not an integer; if np is an integer, any value between the np and $np + 1$ order statistics can be used, i.e.,

$$\xi_p = \begin{cases} X_{[np]+1}, & \text{if } np \text{ is not an integer} \\ \text{any value between} & \\ X_{np} \text{ and } X_{np+1}, & \text{if } np \text{ is an integer} \end{cases} \tag{4-28}$$

Analogous to the Confidence interval for the population median, a confidence interval based on order statistics can be given for ξ_p:

$$P(X_r < \xi_p < X_s) = \sum_{k=r}^{s-1} \binom{n}{k} p^k (1-p)^{n-k} \tag{4-29}$$

Again, the confidence coefficient can be evaluated from tables of the binomial distribution.

4-5.2 TIME SUMMATION SCHEMES

A time summation scheme is a maintainability prediction method by synthesis of elemental task times to arrive at total system maintenance time distribution. It consists of:

1. Considering from a maintenance technician's behavioral viewpoint, all maintenance task steps required to perform a maintenance function
2. Analyzing the maintenance action tasks in light of:
   a. Probability of successful completion
   b. Time to perform (over a distribution spread)
   c. Susceptibility to individual differences
   d. Associated frequencies of occurrence.
3. Summing the resultant maintenance burden load of the maintenance actions to obtain the expected maintenance load at each level of maintenance.

The summation scheme is appealing because of its simplicity and its long academic and industrial application. For example, Ref. 18 (p. 302), breaks down the total time for a complex decision into "sensations", "discrimination", "choices", and other acts. Once the times for the acts could be determined, they were put together and the total time for a behavior was predicted for a new behavior from the elements comprising the function. The effective application of micro-element synthesis to thousands of industrial jobs has demon-

* See AMCP 706-109, Engineering Design Handbook, *Tables of the Cumulative Binomial Probabilities.*

strated the practical power of time synthesis modeling (Ref. 5, p. 35). This time summation synthesis has been applied to military maintenance behavioral tasks, as shown, for example, in the four developed techniques shown in MIL-HDBK-472 (Ref. 3).

The following basic maintainability engineering tools must be utilized in summation synthesis:

1. Functional flow block diagrams
2. Functional level system breakdown charts
3. Reliability failure modes, effects, criticality, and detection analysis
4. Maintenance concept-plan maintenance function breakdown (what, where, how, why, and when factors for maintenance).

When these fundamental factors have been correlated with the requirements for the system operational performance (mission—primary and secondary) the use of the time a summation synthesis methodology begins.

Using the summation process to determine the maintenance task load, one can start either with the system operational maintainability goals/objectives and apply the time summation process in reverse, i.e., from the system down to the lowest maintenance level equipment breakdown allocation (throw-away part or some other convenient level) or at the lowest level of equipment breakdown desired and sum back up to the system level.

The accepted key factors used to define the maintenance burden load are:

1. Time to perform a maintenance task
2. Frequency of the task
3. Probability of completing the task
4. Standard deviation (variance) and confidence level of the task performance
5. Effects of the related maintainability qualitative design characteristics and associated qualitative-quantitative behavioral characteristics of maintenance technicians.

The basic summation elements are expressed:

$$\overline{M}_t = \sum_{i=1}^{n} M_{ti} f_i \Big/ \sum_{i=1}^{n} f_i \quad (4\text{-}30)$$

where

$\overline{M}_t$ = mean time of a higher level breakdown task function

M_{ti} = mean time of a sub-element maintenance task

f_i = frequency of the sub-element task.

This basic formula presently is used in all prediction summation processes (not necessarily using the same symbols as shown). There is no single assigned time to perform a maintenance task; the times vary in proportion to the complexity of the tasks. Therefore, in the summation the time used is an average or mean of a distribution of times; a simple, routine or automatic task involves a small deviation and can be assumed to follow a normal distribution, while a complex task (diagnosis) may follow a skewed distribution, such as lognormal. For example, there is no "common maintenance man"; the common man is an average man from the distribution of a given population of technicians. Time to perform a maintenance step, time to perform a series of steps which comprise a maintenance function, and time to perform a series of maintenance functions which comprise system time are summations of the mean times of the various distributions involved, where the distribution types, confidence levels, risks, variances, standard deviation, median times at the 50th percentile, and maximum times at given percentiles can be determined at each level of a series. These latter determinations are made using the principles cited in Chapter 8 and the nonparametric methods cited in par. 4-5.1.3. The same principles apply in the development of the regression analyses used to relate maintenance action times to equipment and human characteristics.

In all cases the time factors are associated with and based upon the characteristics of design, support equipment, and human factors. In some cases, regression analysis and design checklists are correlated to arrive at the time summation factors. In other cases, the regression analysis is combined with functional analysis of equipment maintenance tasks to define the time-synthesis modeling. In all cases, historical data and/or selected observations of maintenance function activity on existing equipment are used to define the time and frequency data for the synthesis summations of newly developed equipment or to improve the maintainability qualitative characteristics and the quantitative time and frequency factors. In most cases, the developed synthesis is used to justify decisions.

MIL-HDBK-472 (Ref. 3) describes four prediction methods in use: the procedural steps, application, and limitation of the ARINC approach (see Ref. 19, for details upon which Procedure I is based); the Federal Electric scheme (see Ref. 20 for details of the study upon which Procedure II is based); the RCA scheme (see Ref. 21, for details on which Procedure III is based); The Republic Aviation/Fairchild Hiller development (see Ref. 22, for details upon which Procedure

IV is based). These methods are discussed in the literature — Ref. 23 discusses in summary form the various prediction techniques used in MIL-HDBK-472; Ref. 5 discusses the various maintainability prediction methods and results of the MIL-HDBK-472 evaluation as well as the extrapolation, time summation, checklist, simulation, expert judgment, and the matrix tabulation schemes; Ref. 24 discusses,the four MIL-HDBK-472 prediction methods as well as interim approaches and follow-on studies. Further, Ref. 25 gives details of the corrective maintenance border procedure, Ref. 26 describes the Munger-Willis checklist scheme for predicting maintainability feature characteristics, and Ref. 27 describes the TEAM technique for evaluation and analysis of maintainability.

The purpose of this subparagraph is to outline the advantages, disadvantages, limitations, and assumptions of some important methods of predicting maintainability and to orient the analyst in their use. It is then up to the analyst to develop a methodology to suit the equipment of particular‘concern. As brought out in Refs. 3, 23, and 24, there is no mandatory method, and each approach chosen by an analyst need not be confined to a single scheme. The underlying principle is that the analyst must know and understand the type of equipment and its associated characteristics, the mission, the support objectives and the environment. He must develop a technique that is based on the historical data available for the type of equipment of concern. These backup data provide justification for the time and frequency factors used in the prediction methodology, the mathematical models, and the associated qualitative factors and regression analyses that best fit his adaptations. In all cases, the selected method must relate to the maintainability characteristics designed into the equipment, the characteristics of the integrated logistic support equipment and associated maintenance concept-plan and the maintenance and supervisory personnel skill levels available. A discussion of the important methods follows.

1. Federal Electric Scheme (Procedure 11, MIL-HDBK-472). The two methods of predicting maintainability given are for ship-based and shore-based electric systems. Method A predicts mean time to repair ***(MTTR)*** for corrective maintenance only, using tabulated maintenance task times based on 300 observations of maintenance activity in the US Fleet; median of individual repairs is expressed by specific maintenance formulations for various types of distributions (namely, normal, exponential, and lognormal). Method B predicts mean time in terms of man-hours required to perform the maintenance tasks, allowing for time estimates based on known characteristics of the system being developed. Method B includes mathematical formulation for both corrective and preventive maintenance summations of the mean times. Method A does not give adequate recognition of what effect the newly designed features may have on mean time to repair. Both methods use the seven maintenance action summation steps, and system functional level breakdown in developing the tasks which determine the next higher level maintenance time. In order to use Method A intelligently, one must verify the times by comparison with existing, or similar, equipment. The diagnostic task variables are of prime importance in use of times for determining a new equipment maintainability; such analysis should encompass the many complications involved in justifying modifications and exceptions. The analyst, in using this method, should develop tables with variances for the equipment of concern (Ref. 5, p. 36; Ref. 23, p. 21; and Ref. 24, p. 14).

2. Corrective Maintenance Burden Prediction Technique. This method (Ref. 25) utilizes the techniques of Method A of the Federal Electric Scheme with its limitations, advantages, and disadvantages, with the exception that the original seven maintenance category steps have been extended to thirteen by expanding the fault isolation and localization steps. In addition, consideration of the skill and knowledge needed for the thirteen steps is included. Maintenance requirements are correlated with the formal categories of available technician qualifications, training requirements, and associated times. Thus one achieves ultimate trade-offs of training need and equipment complexity. Unfortunately, the procedures for relating task difficulty to technician proficiency are not penetrating enough to be valid. The technique depends upon the use of the fundamental maintainability analytical tools—such as functional flow block diagrams, functional level equipment breakdown, maintenance concept, the failure modes, effects, criticality analysis, and associated fault tree logic networks (Ref. 5, p. 45).

3. ARINC Scheme (Procedure I, MIL-HDBK-472). This prediction method concerns itself with predicting system downtime resulting from unscheduled (corrective) flight-line maintenance of airborne electronic and electromechanical systems involving modular type replacement. Flight-line maintenance is divided into six “Maintenance Categories”. The building block method is used (Elemental Activity) from which other measures of downtime are developed through synthesis of time distribution. For each of the elemental activities the mean and standard deviation and probabilities are calculated. A provision for estimating

logistic delays is given. The synthesis formulation resulted from observing many trials in order to calculate the mean time and standard deviation. The elemental activity is directly related to the six maintenance categories. The model is statistically sophisticated, and considers the distribution characteristics of the various related phenomena; the chain of inferences from assumption through sampling to final probability statements appears clear. The philosophy of prediction techniques is based upon the principles of synthesis and transferability.

Although the method is applicable to flight-line (organizational level) maintenance, it may also be used for intermediate and depot levels by extending the formulas to include other elemental units. The methodology lends itself to computer modeling simulation. As presently formulated, however, it does not lend itself to taking major improved maintainability features into account for new systems/equipment unless the elemental activities for a new system synthesize the systems upon which the formulas were developed. The principal advantage of this method is the fact that maintenance time determinations are based on time distribution and not a point estimate (Ref. 5, p. 48; Ref. 23, p. 21; and Ref. 24, p. 11).

4. Munger-Willis Checklist Scheme (Ref. 26). This method was a pioneering study in 1959 which selected 241 design features with potential maintainability significance for Signal Corps equipment. An elaborate scoring system was used which showed the spread of scores and their related standard deviation. The scoring considered the specific consequences of a design feature. The method provides sensitivity to the differing importance of the listed features. From this summation of scoring, a "maintainability index", or checklist, is developed. A fundamental problem of this checklist, as well as any other type of checklist, is the weighting factor of an individual item. It is very difficult to allot quantitative weights to the attributes and provide convincing justification. In favor of the checklist is the fact that it provides a basis for time estimates once the complex qualitative relationships of the design features and the technical skill requirements are known (Ref. 5, p. 63).

5. RCA Checklist and Prediction Scheme (Procedure 111, MIL-HDBK-472). This prediction method is one of the best procedures developed. It relates

a. The qualitative design dictate characteristics of system

b. The associated qualitative design dictates—-facilities, assistance from other personnel, external support equipment, etc.

c. Maintenance skill time factors through the use of a regression analysis (mathematical model) of the effects of observed maintenance actions on the qualitative factors.

The research utilized a multiple correlation approach, where maintenance time (appropriately delineated) is the criterion that relates the three qualitative parameters. The data were based on corrective maintenance actions observed on ground electronics equipment at three selected Air Force bases. A total of more than one hundred events were monitored over an extended period,, Although the regression model developed was a result of gathering data on a specific type of equipment at selected sites of the total population, it does clearly demonstrate the usefulness of such a prediction technique. If new equipment is being developed to replace similar deployed equipment, the technique is invaluable in developing a justifiable basis for the times and certain qualitative features to be designed into the equipment. When coupled with the tools of maintainability analysis, the analyst has a means of basing predictions on the design characteristics, support factors, and skill level requirements. Once the need for maintenance and the maintenance tasks are defined in sufficient detail, the tasks are scored and summations made and inserted into the regression equation or related nomograph to determine the quantitative mean times. If desired, the analysis can be carried further and the central tendency and dispersion indices determined. Checklist scoring must be done objectively to eliminate optimistic assumptions regarding control of the qualitative factors during design and production. If the population of problems can be defined with confidence, the checklists method produces excellent forecasts. The method is based on the "repair by replacement" principle and the associated steps to determine the replaceable item and checkout repair; therefore, the analysis can be carried out at any level of maintenance in a functional level item breakdown of the system (Ref. 5, p. 65; Ref. 23, p. 21; and Ref. 24, p. 17).

6. Republic Aviation/Fairchild Hiller Scheme (Procedure IV, MIL-HDBK-472). This prediction method is based on historical experience, subjective evaluation, expert judgment, and selective measurement to predict downtime of a system. It was developed in the late 1950's using data from an aircraft system as the basis for prediction. It was one of the first attempts to tie the qualitative features of maintainability into the quantitative determinations. The technique is based on an orderly combination of maintenance task times,

through summation, by integrating the needs for maintenance on the various operational modes and functions involved. The methodology assumes that the time determinations are made by an analyst working closely with an equipment designer to assure that task times are practical, realistic, and applicable to the maintenance functions required to support the operational modes of the equipment. The method uses only single elapsed times for each maintenance task which are equal to the mean times to perform the task under the range of operational environment factors.

The method takes into account the inherent maintainability of the system, since administrative and other delays are not normally definable during the design of the equipment. However, the mathematical formulas developed can be extended to include administrative and logistic times, especially where such factors are known from historical data on similar equipment and logistic environments. The times will vary as a function of the conceptual and physical constraints such as design features, physical resource support, and operational and maintenance concepts. Such applicable constraints must be documented to justify the task time predictions. The methodology combines corrective and preventive maintenance at various summation levels (each task being analyzed separately). An innovation in this method is the consideration that corrective action may occur during preventive actions and thereby be evaluated as part of the total preventive task times. For example, maintenance-induced faults or malfunction may be deferred until preventive action is performed. Normally, the frequency of corrective actions is considered in the computed action analysis and the failure rates apportioned accordingly. The use of the maintainability analytical tools mentioned previously is essential for this method (Ref. 23).

In summary, the basic technique. used by all methodologies is the summation synthesis of time (mean /average) to perform maintenance task, multiplied by the frequency of maintenance action (need for maintenance). Such summations are based upon the steps to be performed for each action for a maintenance function at each level of maintenance activity for the respective functional level breakdown of a system. The methods vary with regard to the relation of the maintenance task to the system functional level breakdown, and the equipment design characteristics, and behavioral factors, and the factors of the integrated logistic life cycle. Most methods utilize regression analysis to relate the observed data, the extrapolations, and experiencejudgment factors on scoring of the design checklists. All methods suggest or infer the type of distribution to be used. Some relate the delay times due to operational readiness and standby, administrative, and/or logistic times. All methods use or imply the use of the basic tools of maintainability engineering analysis—i.e., functional flow block diagrams, functional level breakdowns, maintenance concept, and the reliability failure modes, effects, criticality, and'detection analysis. All methods are based on trends observed in a limited type of equipment. Most methods reveal that the summation of mean times follows the lognormal distribution at the higher level summations.

It is evident that the maintainability analyst need not confine himself to a single method, with the choice based on the usefulness of a given method under the particular circumstances of the equipment of interest. A wise choice of method will eliminate the uncertainties involved in the present state of the maintainability prediction art, especially with regard to behavioral factors of the maintenance technicians. The approach selected by the analyst will affect the plans for data collection, evaluation, analysis, and control which are essential in justification of the predictions made. Careful selection is therefore of the utmost importance.

4-5.3 SIMULATION METHODS

By the term "simulation" is meant the use of a model to capture some aspects of a situation without experiencing the situation itself. Simulation is an important tool of designers and decision makers. For example, wind tunnel tests of scale models of different fuselage and airfoil configurations are used to simulate actual flights of corresponding full-scale aircraft. Changes in the configuration of the scale model are easy and inexpensive to make (if not absolutely, then certainly relative to making similar changes in an actual aircraft), and thus a number of alternative configurations can be considered and the best design from among the available choices can be identified.

The example that follows contains the essentials of a simulation. Given a model of some real situation, the designer or decision maker has under his control certain factors or inputs. Other factors or inputs are beyond his control. For each set of inputs (controllable and uncontrollable), there results a response or output. This output is translated into a measure of performance. Thus, it is possible to relate values of the controllable inputs to values of some measure of performance. Having this relationship, one can determine the optimum inputs, i.e., those inputs from among the set of allowable inputs which yield the best measure of performance.

Much of the discussion that follows concerning the merits of a simulation approach in comparison with purely mathematical methods is either taken directly or paraphrased from Hillier and Lieberman, Ref. 28, Chap. 14.

If it is possible to construct a mathematical model which abstracts the essence of a real situation, reveals its underlying structure, provides insights into cause-and-effect relationships, and is amenable to solution, then the analytical approach is usually superior to simulation. However, many problems are so complex that they cannot be solved analytically, and simulation is the only practical approach.

While the wind tunnel example is a physical simulation, most operations research-type simulations (in which area many maintainability analyses could be said to fall) are mathematical in nature. Such simulation models describe the operation of systems in terms of individual events that their components experience. In a simulation model the system under study is partitioned into elements or subsystems whose behavior can be predicted (at least probabilistically) for each possible state of the system and its inputs. The interactions and interrelations among these elements are also built into the model. Thus, simulation provides a way of dividing the model-building job into smaller units and then combining them appropriately so that interactions are properly represented.

When a simulation model has been constructed, it is activated ("run") and the actual operation of the prototype system is simulated. This requires that input data be supplied to the model, and that the output of the model be recorded. In some instances there may be interest in certain of the inner workings of the model so that system states intermediate to input and output must be recorded as well. By repeating this for the various system configurations under consideration and comparing their performance, the optimum configuration can be approximated. Because of statistical error, one cannot guarantee that the configuration yielding the optimum simulated performance has actually been found. But, if the simulation was properly run, the result ought to be near optimum.

Thus, simulation is essentially a form of random sampling experiment on the model of a system. The experimentation is done on the model rather than the system itself, because the latter would be too inconvenient,, expensive, time-consuming, or unsafe. (In some instances, the system being modeled does not yet exist; the simulation is performed as an aid to proper system design.) Most often simulation experiments are performed on a computer. This is not for any inherent superiority of computer techniques but rather because of the large amount of computing ordinarily required in a simulation.

4-5.3.1 Random Variates for Simulations

Random variates arise as inputs in maintainability simulations in modeling time-to-failure and repair time. More complicated models may also entail random variates involved with delays due to limited resources of manpower, repair parts, maintenance equipment, facilities, etc. These random variates must be selected from appropriate, completely specified distribution functions.

We will presently give a brief discussion of how random variates from prescribed distributions can be generated on a computer. First, we refer the reader back to par. 4-5.1.2 in which we gave a brief discussion of how one might select a particular distribution to describe some random variates. This, of course, is not at all definitive. The choice of a distribution function for a random variable is an area in which judgment, expertise, experience, and statistical methodology can all be brought to bear. As these kinds of choices determine some of the inputs to a simulation, they also influence the outcome of the simulation experiment. One cannot say *a priori* how sensitive the results are to the input. In some cases the influence of the input on the output will be substantial; in other cases there will be little change in the output over a considerable range of variation of the input. In any given simulation, one ought to make an analysis of the sensitivity of the experimental results (i.e., the output) to variations in the inputs. This will reveal how crucial the inputs are to the outputs. If a particular input can be considerably varied without changing the output more than a limited amount (expressed either as an absolute or a relative change), then perhaps less attention needs to be paid to the specification of that input than to inputs which affect the output strongly.

4-5.3.2 Computer Generation of Random Variates

All computer generation of random variables begins with uniform random variables from the unit interval. Such a random variate u has probability density function

$$f(u) = \begin{cases} 1, & 0 \leq u \leq 1 \\ 0, & \text{elsewhere} \end{cases} \tag{4-31}$$

and are themselves produced by a computer process. Every computer center will have a routine which provides such uniform random variates. These routines are typically called random number generators. The casual user has no need to know how these work, so we content ourselves with citing the paper by MacLaren and Marsaglia (Ref. **29**) and Chap. **3** of Naylor, et al. (Ref. 30) for those who want to delve into their inner workings.

In addition to computer schemes for producing random numbers, there are available tables of such random numbers. The most famous and comprehensive of these was produced by the RAND Corporation (Ref. **31**). The introduction to the RAND tabulation describes how the random numbers were produced and gives instructions concerning the use of the tables.

With a uniform random variate u, one can obtain a random variate x, having a strictly increasing cumulative distribution function $F(x)$ by solving the equation

$$\left.\begin{aligned} u &= F(x) \\ x &= F^{-1}(u) \end{aligned}\right\} \qquad (4\text{-}32)$$

This follows from the fact (see, e.g., Lindgren, Ref. **17,** p. 408 or Naylor, et al., Ref. **30,** p. **70**), that if x is a random variable with cumulative distribution function F, then $F(x)$ has the uniform distribution on the unit interval.

The procedure inherent in solving Eq. **4-32** for x in terms of u will be illustrated first graphically and then algebraically. Graphically, as shown in Fig. **4-6,** one enters the graph at the point u on the vertical axis, proceeds horizontally to the cumulative distribution function curve, drops vertically down to the horizontal axis, and then reads off the value x. The figure shows a distribution function for a random variable which can take on negative as well as positive values. Most random variables encountered in maintainability analysis are non-negative, typically, being times measured from a reference point until an event occurs (e.g., failure, repairs).

Now to an algebraic example. Suppose one wants to generate random variates from an exponential distribution with failure rate λ. The cumulative distribution function is

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - \exp(-\lambda x), & x \geq 0 \end{cases} \qquad (4\text{-}33)$$

Solving Eq. **4-33** for $u = F(x)$ for x, one obtains

$$x = -(1/\lambda)\ln(1 - u) \qquad (4\text{-}34)$$

Random variables taking on only discrete values, instead of a continuum of values, are obtained in a similar manner. A slight modification must be made, however, because the cumulative distribution function for discrete random variables is not strictly increasing, but consists of a series of horizontal segments (often called a "step" function). Fig. **4-7** is a representation of a typical cumulative distribution function for **a** discrete random variable. When one enters on the vertical axis at a random level u, he is unlikely to intersect any of the horizontal segments. (Indeed, such an intersection occurs with probability zero.) Thus, instead of solving the equation $x = F^{-1}(u)$, one selects the smallest x for which $F(x) \geq u$.

This procedure is illustrated in Fig. **4-8.** There u is the (uniform) random number that was selected and, for that value of u, the corresponding value of x is x_3.

The method of generating random variates from specific probability distributions via the inverse of the cumulative distribution function (the method just described) is completely general. It will always "work". However, this does not mean that it is always the most efficient method. The generation of random variates from some distributions can be accomplished more cheaply (in the sense of amount of computer time per random variate) by taking account of the special characteristics of the distribution at hand. We will not pursue this here, but will refer the reader to Chap. **4** of Naylor, et al., Ref. 30 (and the bibliography thereto) for a very good discussion.

4-5.3.3 Example

We will close this paragraph with an example of a simulation used to evaluate some alternative maintenance policies for a series of high-pressure injection pumps. (This example is taken from Bowman and Fetter, Ref. **32,** pp. 426-431).*

A chemical company has a series of high-pressure injection pumps operating under similar conditions and wishes to determine a proper maintenance policy. The pump valves are subject to failure, and their routine maintenance costs about 9,500 man-hours per year. Each pump has three intake valves and three exhaust

---

* **Adapted with permission from Bowman and Fetter, *Analysis for Production and Operations Management*, 3rd ed.; Homewood, Ill., Richard D. Irwin, Inc., pp. 426-31.**

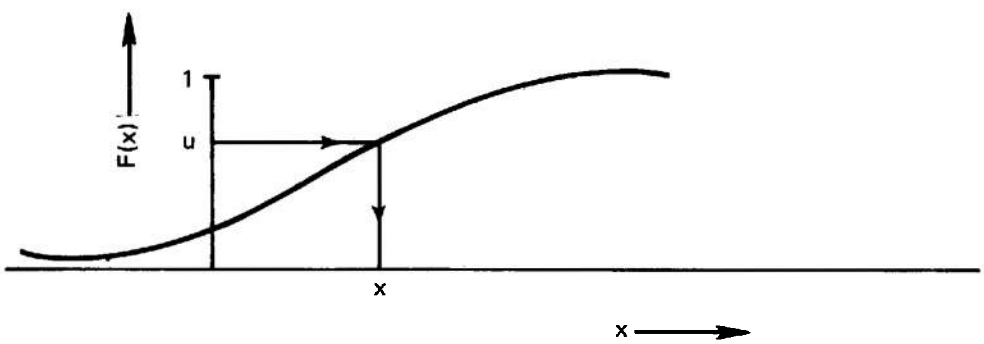


Figure 4-6. Solving $x = F^{-1}(u)$ Graphically

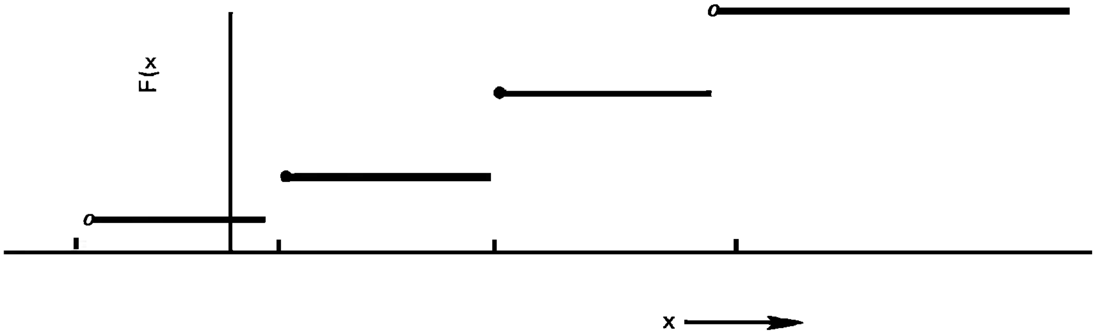


Figure 4-7. Cumulative Distribution Function for a Discrete Random Variable

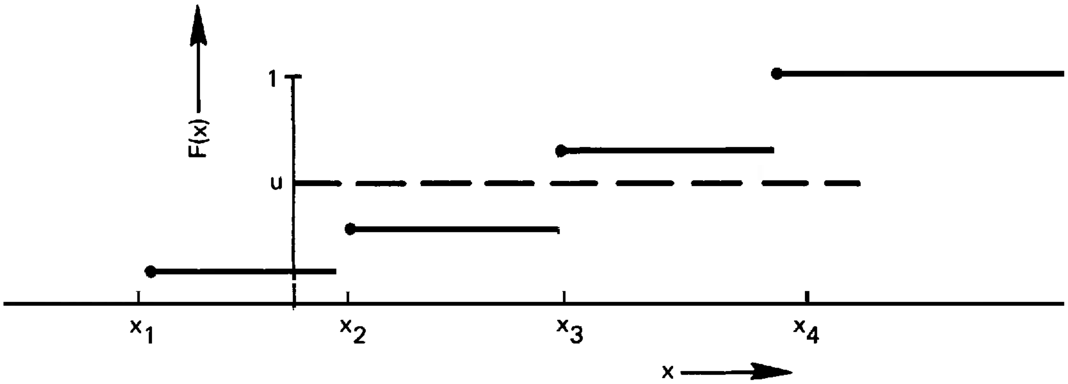


Figure 4-8. Generating a Random Variate from a Discrete Distribution

TABLE 4-2.
MAINTENANCE COST ('EXPRESSEDIN MECHANIC'S TIME)

Operation	Time, hr
Shutdown, prepare for maintenance	1/2
Remove manifold (either intake or exhaust)	2/3
Disassemble one valve	1/3
Overhaul one valve	1-1/4
Assemble one valve	1/3
Replace manifold (either intake or exhaust).	2/3

valves. When a valve fails, it is necessary to shut down the pump and prepare it for maintenance. Each set of valves is covered by a manifold which must be removed after shutdown in order to expose either the three intake valves or the three exhaust valves. There is no downtime cost as the firm has standby pumps to be used during maintenance on the valves.

The company is interested in and wants to evaluate four maintenance procedures which it considers practical:

(I) Repair a valve only when it fails.

(11) Repair all three exhaust valves if one exhaust valve fails, or all three intake valves if one intake valve fails.

(111) Repair all six valves (three exhaust and three intake valves) whenever a pump must be shut down to repair one valve.

(IV) Repair the valve that fails plus all valves which have been in use more than the estimated average service life (560 hr).

The company supplied for the analysis the data shown in Table 4-2.

A cumulative probability distribution for valve service life was constructed from empirical data supplied by the company. Valve life itself, of course, is not a function of the given procedures, assuming the valve remains in use until it fails.

By using the inverse cumulative method, valve life-lengths were generated from this empirical distribution. Simulated experience from a limited set of random draws of valve life-lengths is plotted in Fig. 4-9 for each of the four alternative maintenance policies. It can be seen that because of the different policies, the operations take place at different points in time. A vertical line represents a valve overhaul, or set of 3 in 11, or set of 6 in 111, or a varying number in IV according to the number over the average age of 560 hr.

Policy II will be explained to show the nature of the other charts (and their associated costs). The first valve to fail in the intake manifold was a valve 3, at 440 hr. According to procedure II all three valves, 1, 2, and 3, within the intake manifold are disassembled and repaired. Therefore, all three end their first lives and commence their second lives. It can be seen that valve 3 again fails first at 430 hr, or at about 870 hr (440 + 430) on the clock. Again, all three valves are repaired and started anew. In the third life set, valve 1 fails at 80 hr, or about 950 hr (870 + 80) on the clock, all three valves are repaired, and again started anew. The three valves in the exhaust manifold, 4, 5, and 6, have been operating according to the experience shown.

In policy 111, it can be seen that all six valves are repaired and started anew when any of the six fail. In case IV when a valve fails, any other valve older than its expected life is repaired. For instance, at 740 hr on the clock, valve 1 failed. Valves 2 and 5 were still operating but were over their expected lives of 560 hr and were, therefore, overhauled according to procedure IV. The fact that valve 2 would have lasted until 970 hr would, of course, have been unknown to the repair men.

The analysis for the first 2,300 hr of these four alternatives is shown in Table 4-3. The number of times the different operations were performed was counted from Fig. 4-9 and tabulated in Table 4-3. For instance, up to and including 2,300 hr, policy I had experienced 20 shutdowns and policy 111, 14 shutdowns. However, in policy I a shutdown means one overhaul, or 20 overhauls in total. In policy III a shutdown means six overhauls, one for each valve, or a total of 84 shutdowns (14X 6).

It can be seen from Table 4-3 that alternative I would be the best policy according to this brief simulation. However, alternative IV is within 10 percent of it. These are only averages or, rather, cumulative sums;

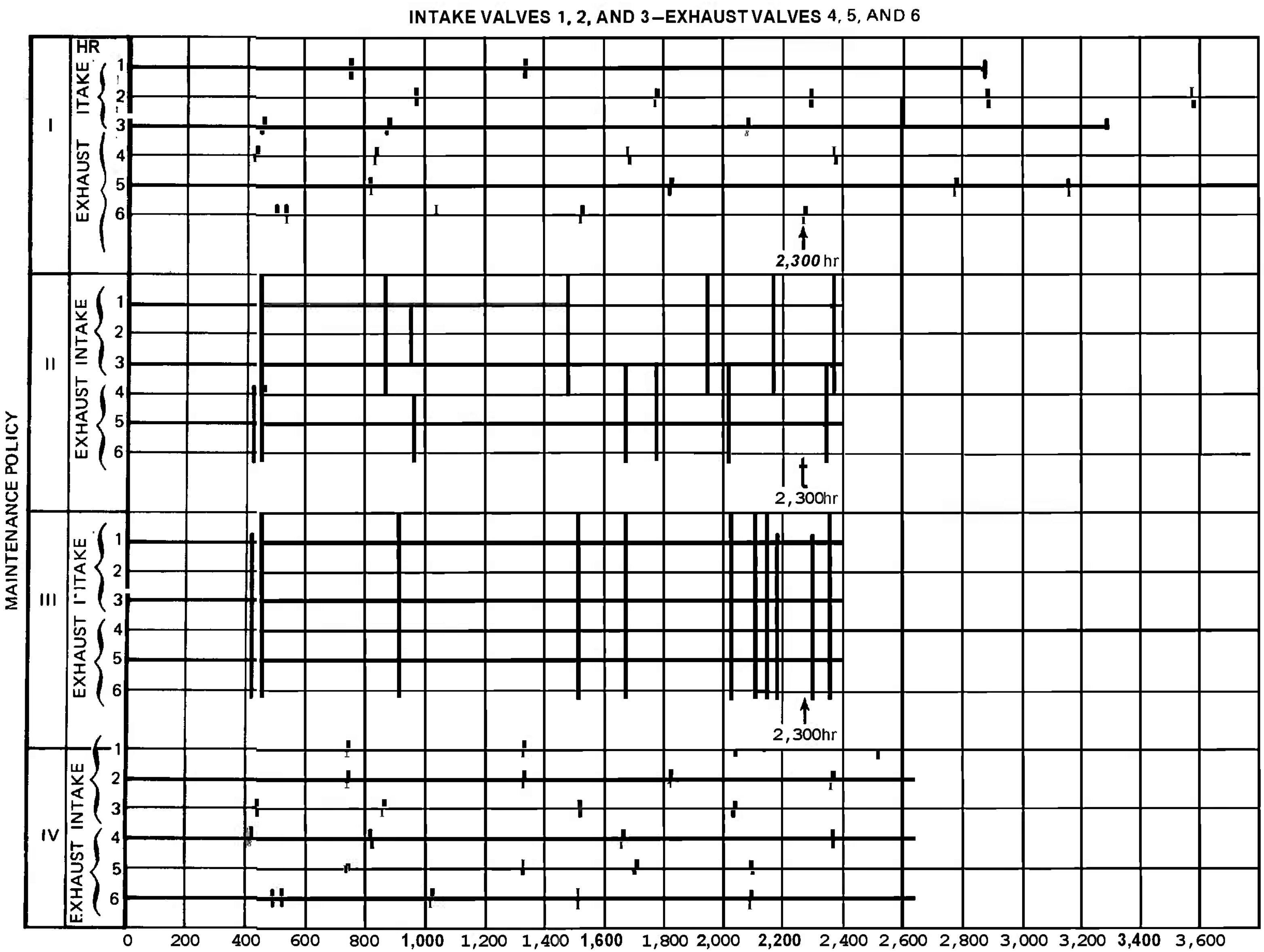

INTAKE VALVES 1, 2, AND 3—EXHAUST VALVES 4, 5, AND 6
MAINTENANCE POLICY
HR
I
II
III
IV
INTAKE
EXHAUST
1
2
3
4
5
6
2,300 hr
2,300hr
2,300hr
0
200
400
600
800
1,000
1,200
1,400
1,600
1,800
2,000
2,200
2,400
2,600
2,800
3,000
3,200
3,400
3,600

TABLE 4-3.
ANALYSIS OF ALTERNATIVE MAINTENANCE POLICIES

Operation	Hr/Op.	I		II		III		IV	
		No.	Times	No.	Times	No.	Times	No.	Times
Shutdown	1/2	20	10	13	6-1/2	14	7	17	8-1/2
Remove intake manifold	2/3	9	6	6	4	14	9-1/3	8	5-1/3
Remove exhaust manifold	2/3	11	7-1/3	7	4-2/3	14	9-1/3	12	8
Disassemble valves	1/3	20	6-2/3	39	13	84	28	24	8
Overhaul valves	5/4	20	25	39	48-3/4	84	105	24	30
Assemble valves	1/3	20	6-2/3	39	13	84	28	24	8
Replace intake manifold	2/3	9	6	6	4	14	9-1/3	8	5-1/3
Replace exhaust manifold	2/3	11	7-1/3	7	4-2/3	14	9-1/3	12	8
Total Time			75		98-7/12		205-1/3		81-1/6

there is no indication of cost variation for each policy. The economics of this problem would justify certainly a longer run than the one used here for demonstration purposes. As the trials are made, it would be advisable to get a cost (in hours of mechanic's time) for each of a set of periods such as every 2,000 hr. From these lists of numbers (costs), the variance inherent in the system could be determined, and some decision reached concerning confidence (statistical or intuitive) in finding the best of the alternatives.

4-5.4 EXPERT JUDGMENT METHOD

The times to perform a maintenance action and the individual maintenance action steps are often based upon statistical facts gathered from historical data banks for existing similar equipments or from statistical sampling of specific observations of representative deployed populations. For new equipment extrapolation should be used as a basis for expert judgments. Generally, in all prediction methods, some portions of the predictions are made using engineering judgment for estimation of the maintenance action task times. Such judgments are most often made when new design complexities appear in the analysis. Usually the judgments are based on:

1. Extrapolation of the observable trends
2. Conjectures about ease of servicing and repair based on inspection of prototypes, mock-ups, and breadboards
3. Analysis of technicians' capabilities and behavioral factors required for the performance of the task steps
4. Design characteristic checklists
5. Time and motion studies of the typical maintenance action steps
6. Field experience
7. Knowledge of exceptional maintainability features of the equipment that might cause a radical change in maintenance requirements
8. Recognition of the maintenance concept limitations at the various levels of maintenance
9. Knowledge of system functional level breakdowns for replacement at various levels of maintenance.

When these basic ingredients are observed in making judgments reliable maintenance time predictors can be obtained.

Of course, much depends upon the specific goal of the expert judgment in terms of the maintainability objectives. One might start with a rough extrapolation from past prime equipment and look for exceptional features that are present in new designs. Expert judgment is an inexpensive and easy process, if based on related experience and if unbiased (Ref. 5, pp. 80-81). The one weakness in expert judgment is the definition of the "average technician capabilities" versus the overall available distribution of technicians' capabilities, when trying to arrive at point estimates of the mean times. The point estimation of the mean time to perform a maintenance function uses the time-frequency summation synthesis method to arrive at, or to verify, the next higher subclass mean time up to the system mean downtime prediction. Since errors in judgment are cumulative they must be controlled by applying the principles cited in par. **4-5.2.**

4-5.5 MATRIX TABULATION METHODS

Complicated relationships among several variables can often be expressed in a matrix format. Maintainability engineers over the past two decades have used various matrix models to describe the need for maintenance tasks and the interrelationship among various elements—such as operational requirements, design characteristics, costs, and integrated logistic support requirements. They usually have been qualitative, but have lacked quantitative mathematical relationships because of the complexity of combining the variables common to the rows and columns of such matrices. This is clearly shown in Ref. **26** in the development of the checklist matrix.

There are several applications of quantitative matrices but these have been limited to small sections of the maintainability prediction problem, such as the interface of the technician and the fault diagnostic action. To be useful, maintainability prediction matrices should display quantitative relationships, such as correlation coefficients, transition probabilities, variances, and quantitative factors of time and frequency.

Two matrix tabulation methods, which might be of significance for maintainability predictions, are the Sympton Matrix (Ref. **33)** developed by **H.** R. Leuba, and the EPRG Symptom-Hypothesis Matrix (Ref. 5, pp. **86-88)** developed by the Electronic Personnel Research Group of the University of Southern California. Both schemes are primarily concerned with the diagnosis of malfunctions and the optimization of diagnostic procedures. Historically, in complex electronics the time required to diagnose what is wrong (fault, verification, isolation, and localization) consumes a major portion of the maintenance downtime (in some cases from ***50*** to **75** percent). Also, the distribution of the diagnostic time has a large variance and usually displays multimodal effects. Thus, an optimization of the diagnostic procedures and step sequences in searching for the "culprit" is certainly in order.

In the paragraphs that follow the ARINC and EPRG matrices are briefly discussed. The discussion is based on the text of Ref. **5,** quoted almost verbatim in places but greatly abbreviated. The interested reader is referred to the original texts (Refs. **5** and **33).**

4-5.5.1 The ARINC Symptom Matrix

This scheme was devised primarily to evaluate check-out and trouble-shooting procedures. However, in addition the scheme yields information which is pertinent to ease of servicing and some quantitative indices of this attribute. The analysis starts with a list of symptoms or gross output states of the prime equipment. Each of the major output states has associated with it one or more subsymptoms. Next, one examines the possible failure modes of each part of the system, and from this output the symptom significance of each kind of failure is determined. Through appropriate summation, one assembles a list of probable causes which are weighted according to their failure rates, and an estimate of the likelihood of each failure mode for each given symptom is tabulated in the form of symptom matrix. This useful tool for diagnostic time prediction analysis starts with the symptom which is where the technician starts. Also, it furnishes a complete rank-ordering of the "potential culprits" in a diagnosis problem. Since the essential output of the symptom matrix is a probability for each possible malfunction, cost factors can be applied in a direct way. If one knows the cost of checking each alternative, and the costs of downtime, then optimal diagnosis sequences are readily calculated.

4-5.5.2 The EPRG Symptom-hypothesis Matrix

The matrices of symptoms produced by defined classes of failure modes of components, stages, or larger functional level items of a system are the basis for developing several measurements and the prediction schemes. The original interest was in scoring the technician's fault localization and isolation proficiency.

R. L. Weis of the University of Southern California in **1963** developed a concept of this kind of problem solving as an iterative process in which, each time, one of a set of alternative tests is selected. The consequence of each selected test is to reduce the initial uncertainty in the situation by some amount. This initial uncertainty was defined as a function of the failure modes of components (or stages, or subsystems, etc.) and the number of outcomes of a test.

The amount of uncertainty reduction for any one test (or move) in a sequence of such responses is then computed, and by applying a decision rule for selecting the "next test" or "next move", the diagnostic test sequence is optimized.

The matrix is composed of a set of conditional probabilities of obtaining a given data output, given one of the hypotheses (malfunctions). The matrix can be computerized to score records of technicians' fault localization and isolation behavior. An extension of the uncertainty reduction concept to predicting diagnostic complexity levels of systems is a very promising development.

It is appealing that source uncertainty reduction measure can be computed from these matrices (ARINC and EPRG) and predictions can be made of how difficult the diagnostic job will be for a given system. Such knowledge can be used for design improvements before hardware is built.

In summary, this chapter discusses the various maintainability factors and characteristics that must be considered in maintainability prediction and allocation. It presents two allocation methods, four prediction methods of MIL-HDBK-472; and, in addition, the prediction methods by extrapolation (smoothing and nonparametric), by simulation, by use of expert judgment, and by matrix tabulation.

REFERENCES

1. MIL-STD-470, ***Maintainability Program Requirements,*** March 1966.
2. R. E. Barlow and F. Proschan, ***Mathematical Theory of Reliability,*** John Wiley & Sons, New York, 1965.
3. MIL-HDBK-472, ***Maintainability Prediction,*** May 1966.
4. AFSC, DH 1-9, ***Maintainability,*** December 1970.
5. J. W. Rigby, et al., Technical Report No. 40, ***Maintainability Prediction: Methods and Results,*** University of Southern California, June 1964.
6. B. S. Blanchard, et al., ***Maintainability Principles and Practice,*** McGraw-Hill Book Co., 1969.
7. DA Pam 705-1, ***Maintainability Engineering.***
8. AFSCM 345-5, ***System Management,*** June 1964.
9. MIL-STD-721B, ***Definitions of Effectiveness Terms,*** 25 Aug. 1966.
10. W. E. Milne, ***Numerical Calculus,*** Princeton Univ. Press, Princeton, N.J., 1949.
11. R. G. Brown and R. F. Meyer, "The Fundamental Theorem of Exponential Smoothing", Operations Research, **9,** No. 5, 673-685 (1961).
12. B. Epstein, "Testing for the Validity of the Assumption that the Underlying Distribution of Life is Exponential", Technometrics, **2,** 83-101, and 167-184 (1960).
13. H. W. Lilliefors, "On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown", Journal of the American Statistical Association, ***64,*** 387-389 (1969).
14. J. Aitchison and J. A. C. Brown, ***The Lognormal Distribution,*** Cambridge University Press, Cambridge, 1969.
15. M. R. Mickey, et al., ***Test Criteria for Pearson Type III Distributions,*** ARL 63-100, Aeronautical Research Laboratories, Wright-Patterson AFB, Ohio, June 1963.
16. A. S. Goldman and T. B. Slattery, ***Maintainability: A Major Element of Systems Effectiveness,*** John Wiley and Sons, Inc., New York, 1967.
17. B. W. Lindgren, ***Statistical Theory,*** 2nd Ed., MacMillan, New York, 1968.
18. R. W. Woodworth, ***Experimental Psychology,*** New York, 1938.
19. ARINC ***Maintainability Prediction, Theoretical Basis and Approaches*** (Rev), 31 Dec. 1963, Publication 267-02-6-420.
20. Federal Electric Corp., A ***Maintainability Prediction Procedure for Designers of Shipboard Electronic Equipment and Systems,*** July 1960.
21. RCA Final Report IV Progress Report, ***Maintainability Prediction Technique,*** 15 March 1962, RADC Report TDR 62-156.
22. Republic Aviation Corp. ***Maintainability and Supportability Evaluation Technique,*** WADD Tech Note 60-82 Part 2, March 1960.
23. P. Sgouros and S. R. Calabro, "Maintainability Prediction in Dept. of Defense Development Progress", IEEE Transactions on Reliability, R-16 No. 1, 25 (1967).
24. NSIA Task Group ***M & R*** Panel, ***Maintainability Predictions Task Group Report, April 1967.***
25. Federal Electric Report ND ***65-36, Corrective Maintenance Burden Prediction Procedure and Procedural Instruction,*** June 1965.
26. American Institute for Research, ***A Development of an Index of Electronic Maintainability—A Research Report,*** Army Signal Corps Agency, 1959.

27. D. R. Bunger & D. R. Dibble, *Technique for Evaluation and Analysis of Maintainability (TEAM),* AIAA Aircraft Design and Technology Meeting, Los Angeles, Calif., Nov. 15-18, 1965, Paper 65-793.
28. F. S. Hillier and G. J. Lieberman, *Introduction to Operations Research,* Holden-Day, Inc., San Francisco, 1967.
29. M. D. MacLaren and G. Marsaglia, "Uniform Random Number Generators", Journal of the Association for Computing Machinery, 12, No. 1, 83-89 (1965).
30. T. H. Naylor, J. L. Balintfy, D. S. Burdick, and K. Chu, *Computer Simulation Techniques,* John Wiley and Sons, Inc., New York, 1966.
31. The RAND Corporation, *A Million Random Digits with 100,000 Normal Deviates,* The Free Press, Glencoe, Ill., 1955.
32. E. H. Bowman and R. B. Fetter, *Analysis for Production and Operations Management,* Richard D. Irwin, Inc., Homewood, Ill., 1967.
33. H. R. Leuba, *The* Symptom *Matrix, An Analytical Tool for Evaluating System-Maintenance Checkout Procedure,* ARINC Research Corporation, Publication 4607-1-294, April 16, 1962.

CHAPTER 5

MAINTAINABILITY DESIGN TECHNIQUES AND INTERFACES

SECTION I

MAINTAINABILITY DESIGN AND MAINTENANCE SUPPORT PLANNING

5-1 GENERAL

In the preceding chapters, the basis for maintainability as a characteristic of system design has been established. Chapter 1 deals with maintainability as a concept. It describes maintainability as one of the important elements of system design (Fig. 1-1). Indeed, the accepted definition of maintainability states that maintainability is a characteristic of design and installation. Chapter 2 shows how maintainability is a significant contributing factor to system effectiveness. In Chapter 3, the various activities with which the maintainability engineering function is concerned are described. Design is one of those activities. The system and equipment features which will promote cost-effective ease of maintenance are of prime concern to the maintainability engineer.

In the final analysis, the design engineer must design maintainability into the equipment. All statements of maintenance concepts and requirements will be of little significance if a conscious design effort is not made to include those specific features of equipment design which will promote ease of maintenance. Maintainability is that portion of the maintenance function over which the designer has control (Refs. 1 and 2).

Although specific attributes of concern for equipment design are listed in many excellent maintainability design guides and handbooks, maintainability has often been relegated to a position of secondary importance by design engineers. It is the function of the maintainability engineer to insure that maintainability considerations do not fall out of the design as afterthoughts during late stages of equipment development so that effective maintenance and maintainability cannot be realized. Rather, the maintainability engineer must see that maintainability considerations are brought to bear as part of total system design from the earliest life cycle phases, as described in Chapter 3.

Of primary significance to design for maintainability is the development, early in the system life cycle, of the system maintenance concept so that all design features will be consistent with operational and logistic support concepts. This is the concern of Integrated Support Planning (Refs. 3 and 4). No longer is it adequate for the design engineer to design to performance and packaging requirements only, leaving maintenance concepts to be developed as a consequence of design instead of as a prerequisite to design. Maintenance Support Planning is the method for determining maintainability design requirements from operational and logistic support concepts. The Maintenance Support Plan—initiated during the latter part of concept formulation—documents the maintenance concept developed from operational and maintenance policies and system requirements, and identifies reliability and maintainability characteristics and requirements for system/equipment design.

Section I of this chapter is concerned with the development of the maintenance concept and maintainability design requirements resulting from the analysis of operational and logistic support requirements and concepts.

Maintainability has been described as dealing with features of equipment design, support, and personnel (Refs. 5 and 6). More specifically, maintainability design is concerned with such system and equipment characteristics as accessibility, controls, displays, test points, test equipment, tools, connectors, mountings and fasteners, labeling and coding, and maintenance information. Of significant importance to maintainability are such additional items as personnel numbers and skill levels, training, human factors, and safety. Reliability engineering is usually concerned only with the physical characteristics of the system or equipment. In contrast, maintainability engineering cannot divorce itself from these personnel-related factors. This is why human factors, safety, and training are engineering disciplines which are so closely allied to maintainability.

Section II of this chapter discusses maintainability design considerations, including specific techniques and interfaces, and the methodology for achieving a cost-effective system/equipment design with respect to maintainability and its allied disciplines. It includes a method, described in Section I, for determining maintainability design requirements resulting from operational and maintenance concepts in terms of the system life cycle, maintenance tasks, system/equipment levels at which maintenance is performed, and specific equipment attributes.

5-2 INTEGRATED LOGISTIC SUPPORT AND MAINTENANCE SUPPORT PLANNING

System engineering (par. 2-2.6) requires that all matters which relate to system design and acquisition be systematically considered. Emphasis is given to integrated logistic support and maintenance support planning as design-influencing considerations which are often as significant to the achievement of a life-cycle, cost-effective system design as the performance characteristics of the prime operating system. Maintenance support planning considerations and their implications for maintainability design are included under the heading of Integrated Logistic Support (ILS) in the total planning process (Refs. 3, 7, and 8).

Integrated Logistic Support is defined as ". . . a composite of all the support considerations necessary to assure the effective and economical support of a system or equipment for its life cycle. It is an integral part of all other aspects of system or equipment acquisition and operation" (Ref. 7). DoD Directive 4100.35 states that "A complete system approach shall be used for planning, analyzing, designing, and managing the incorporation of logistic support into the acquisition of systems." In a joint memorandum to the service secretaries from the Director of Defense Research and Engineering and the Assistant Secretary of Defense (Installation & Logistics) on the subject of revision to the Integrated Logistic Support Planning Guide, 4100.35-G, ILS is identified "not as a separate entity, but as an integral part of the system engineering process" (Ref. 9).

Logistic support planning must begin during the concept development phase even though a formalized program with a detailed ILS Plan may be deferred until the validation phase. In the Army, the ILS Plan includes the Maintenance Support Plan, the Logistic Support Plan, and other related plans (e.g., training plans) as appropriate.

The elements of Integrated Logistic Support are listed in DoD Directive 4100.35 as:

1. The Maintenance Plan
2. Support and Test Equipment
3. Supply support
4. Transportation and Handling
5. Technical Data
6. Facilities
7. Personnel and Training
8. Logistic Support Resource Funds
9. Logistic Support Management Information.

The first of these elements is central to the rest. The plan is a periodically updated document initiated in the latter part of the conceptual phase for an item of military design and during planning for procurement for a commercial item (Ref. 3).

To achieve the required operational capability and availability of Army systems and equipment on a life-cycle, cost-effective basis, logistic support considerations must have a meaningful relationship to design, development, test, evaluation, production, and operations at all stages of the system cycle beginning with early conceptual studies. This requires that the design of operational systems take into account the aspects of logistic support, in view of the available resources and under the conditions and in the environment in which the systems will be used. Thus, trade-offs appropriate to the stage of development must be made to maximize the effectiveness and efficiency of the support system to a degree which is consonant with the overall system operational requirements. It is the purpose of logistic support planning and maintenance support planning to achieve this balance.

The Maintenance Support Plan for a system or equipment is a detailed description of how, when, and where the equipment and each of its end items will be maintained, and what resources will be required to accomplish each maintenance task. It is the major output of maintenance support planning.

Maintenance support planning is illustrated in the flow chart of Fig. 5-1. The input to the process is the set of operational and logistic support policies and constraints stated in the system operational requirement documents. The operational concept is a "Plan For

Use" of the system and is derived from an analysis of system operational requirements developed by the system user or requirements developer, e.g., US Army Training and Doctrine Command (TRADOC). The integrated logistic support concept and its subsidiary maintenance and supply support concepts—prepared by the system (hardware) developer or producer, e.g., US Army Materiel Command (AMC)–constitute a "Plan For Support" of the system. These are derived from the operational concept. It is important to develop, early in the system definition phase, an interface between the Plan For Support and system/equipment design. This interface is a set of design criteria for logistic elements which result from the integrated logistic support, maintenance, and supply support concepts. These criteria should include both qualitative and quantitative statements to provide guideline information to design engineers. One of the important outputs of maintenance support planning is the set of reliability and maintainability design requirements necessary to meet system operation and support requirements.

Among the activities which come under the heading of maintenance support planning and which have impact on maintainability design are:

1. The formulation of maintenance concepts
2. The determination of maintenance tasks and resource requirements
3. The determination of maintainability design requirements from the support viewpoint
4. The performance of analysis and evaluation of development and production hardware configurations to determine the support required.

Maintenance support planning performed during the concept development and validation phases is concerned with applicable maintenance policies and goals, collectively called the system maintenance concept, derived from operational and logistic support concepts. Maintenance policies and goals, from the user's viewpoint, consist of statements—both qualitative and quantitative-concerning system operation, maintenance activities, maintenance resources, and system effectiveness. These, in turn, when logically combined by the system developer, lead to configuration policies and goals and to the resultant implications for system design for maintainability. A method for accomplishing this is given in Ref. **10.** It allows the appraisal of maintenance needs in terms of their effects upon system design and upon life cycle costs, and should result in the establishment of realistic maintenance and maintainability objectives.

Maintenance support planning during validation is concerned with the development of more detailed maintenance concepts, specific maintenance task and resource requirements, specific design features, and the prediction, demonstration, and evaluation of maintenance and maintainability to the lowest required equipment and end-item levels. Maintenance Engineering Analysis Data Sheets (TM **38-703-3**) are used to document maintenance support planning decisions.

5-3 OPERATION CONCEPT

As described in par. **3-2.1.1**, the determination of the operational concept of the system—prepared by TRADOC to specify the required operational capability—is one of the primary activities of the concept development phase. It is the starting point for all system development planning. The operational concept includes:

1. A description of the threat or operational need
2. A description of the anticipated operational environment
3. A description of mission and performance envelopes and system operating modes resulting from threat and mission analyses
4. A determination of mission profiles, operational time factors, and system and equipment utilization rates
5. An elaboration of system effectiveness criteria and requirements in mission-oriented terms to include maintainability
6. A determination of the system life cycle, including system deployment, logistic endurance factors, and out-of-service conditions
7. A description of other system conditions and constraints.

The description of the threat or operational need details why and when the system is needed, the intended purpose for which the system is to be designed, and the resultant effect of not meeting the need. The anticipated operational environment includes geographic, physical, political, legal, and social factors which influence the anticipated operational need. Of specific concern to the maintainability engineer are the geographic and physical environments which dictate the constraints under which the system must be operated and supported. For example, the requirement for operation in extreme cold climates necessitates the layout of controls far enough apart so that the operator or maintenance technician wearing arctic-type gloves can grasp and turn the controls, a human factors consideration. For the maintainability engineer, low temperature extremes require consideration of lubrication, cleaning, and adjustments and what maintenance tasks can be performed under these conditions. Chapter 10, AMCP 706-115 (Ref. 91), describes the environmental conditions with which the designer must be concerned.

Once the operational need and environment have been analyzed and described, it is possible to synthesize and analyze various mission concepts and scenarios and to determine how well these alternative approaches meet the stated needs. In fact, mission analysis is performed together with the needs or threat analysis in a closely linked iterative process.

Among the factors which must be considered are the force structure in terms of types and composition, unit divisions, the number of systems/equipments required per operational installation, mobility, distance from support facilities, the length of the supply pipeline, survivability as a result of enemy action, vulnerability to enemy action and accidental damage, safety of personnel and equipment, and geographical and other physical environment factors.

The result of iterative, threat-and-missionanalyses is a set of mission and performance envelopes within which the system must operate and be supported. Mission profiles may then be constructed from the mission and performance envelope descriptions, indicating the duration and frequency of the various time periods described in par. 2-2.6.1. These include alternate modes as a result of varying mission requirements or as a result of a failure of some part of the system. Goldman and Slattery (Ref. 12) discuss the use of mission profiles and their effect on maintainability.

System performance and effectiveness requirements can now be established along with the operational time factors, and system and equipment utilization rates. This information is essential for both reliability and maintainability design, and for the trade-offs necessary between these disciplines and system performance capability in order to meet the specified system effectiveness. Performance and effectiveness parameters and criteria can be specified and weighted with respect to their contribution to mission success.

The threat and mission analyses, along with the mission profiles and utilization factors, allow the duration of the system life cycle to be defined in detail. This includes policies regarding system deployments, the duration of such deployments, inactive periods, and out-of-service conditions such as overhaul or modification.

Another result from the development of mission profiles, one which is of specific importance to maintainability design, is the determination of a set of logistic endurance factors. The logistic endurance factors indicate when support activities may take place, what may be done at these times, when preventive maintenance actions are allowable and for what duration, the influence of mission criticality on maintenance and logistic support, required turn-around times, and other operational influences on logistic support and support design.

Analysis of these logistic endurance factors, along with stated logistic support doctrine and policies, form the framework for the development of an overall integrated logistic support concept for the system. This approach represents a radical departure from past practice in which an overall logistic support concept rarely existed, often resulting in a system with equipments designed by different contractors under different and loosely defined support ground rules and with conflicting and inconsistent support policies and practices.

The integrated overall/logistic support concept forms the basis for maintenance support planning and supply support planning, as shown in Fig. 5-1. All of the preceding considerations form the operational concept to which the system design must respond.

5-4 MAINTENANCE CONCEPT

The development of the maintenance concept is the central activity of maintenance support planning. The maintenance concept defines criteria for maintenance activities and resources allowable at each of the specified maintenance levels. It is derived from the operational and ILS concepts of the system and from the policy statements which form the constraints and boundaries of the support system as expressed in requirements documents. The maintenance concept serves two purposes:

1. It provides the basis for the establishment of maintainability design requirements.

2. It provides the basis for the establishment of maintenance support requirements in terms of tasks to be performed, frequency of maintenance, preventive and corrective maintenance downtimes, personnel numbers and skill levels, test and support equipment, tools, repair parts, facilities, and information.

For example, if the maintenance policy is that no external test equipment is allowable for organizational level maintenance, one design implication is that built-in test features must be incorporated to allow any necessary checkout and alignment at this level. Or, if the corrective maintenance policy is no repair at the organizational level except to replace failed items, then the

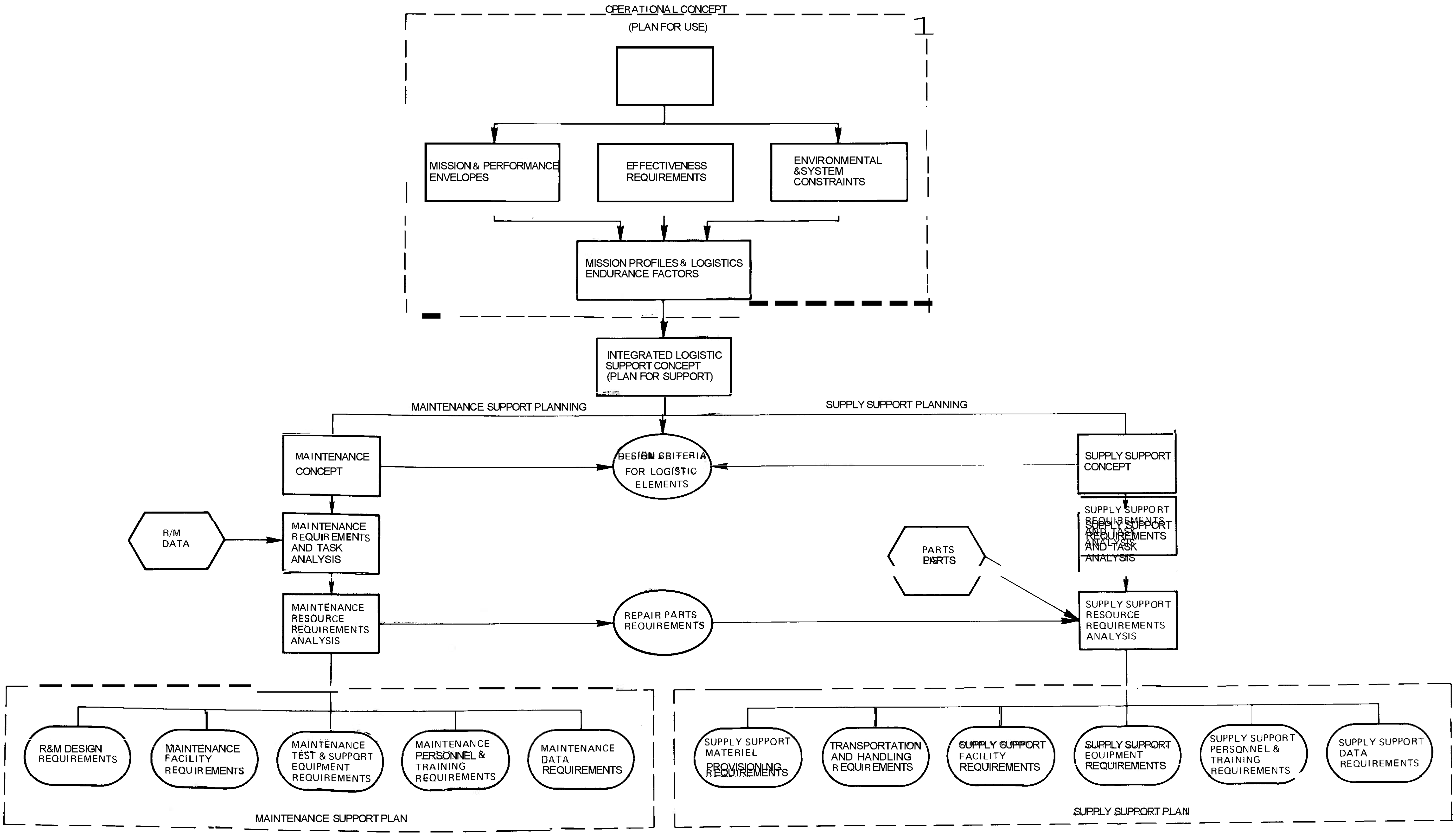


Figure 5-1. Flow Chart of Maintenance/Supply Support Planning

design implication is the use of modular design to the maximum extent feasible. This also means that sufficient spare modules must be provided at the organizational and direct support levels to meet specified effectiveness and readiness requirements. If the operator or repairman at the organizational level must monitor, calibrate, or adjust his equipment, then adequate training and technical manuals must be provided along with necessary calibration and alignment equipment and tools which are not built into the prime equipment.

The maintenance concept must be both realistic and sufficiently definitive to meet the needs of the system/design engineers and the requirements of logistic support planners.

Since the primary purposes for which a system is acquired are intimately related to some set of missions, analysis of the implications of maintenance policies on system design starts logically with mission and operational requirements—the operational concept. The maintenance concept is concerned with policies and goals pertaining to:

1. Operational states of the system
2. Maintenance activities
3. Maintenance and support resources
4. System effectiveness.

These categories may be further subdivided as shown in Table 5-1.

5-4.1 LEVELS OF MAINTENANCE

The Army has four maintenance level categories (Table 5-2):

1. Organizational maintenance
2. Direct support maintenance
3. General support maintenance
4. Depot maintenance

The titles may change but these levels of maintenance are essential to keep equipment in the field in operating condition.

5-4.1.1 Organizational Maintenance

Organizational maintenance is that maintenance normally authorized for, performed by, and the responsibility of, a using organization on equipment in its possession. It includes inspecting, servicing, cleaning, lubricating, adjusting, and the replacement of parts, minor assemblies, and subassemblies. Organizational level personnel are usually fully occupied with the operation and use of the equipment, and have a minimum amount of time for detailed maintenance or diagnostic checkout. Personnel at this level usually do not repair the removed items but forward them to the next higher level if maintenance is to be performed on the item.

Maintenance performed by the equipment operator usually consists only of inspecting, cleaning, servicing, and adjusting the equipment. Maintenance done by the organization repairman consists of making minor repairs and replacements.

Mobility requirements generally limit the amount of tools, test equipment, and supplies available at the organizational level. The design engineer should plan accordingly.

5-4.1.2 Direct Support Maintenance

Direct support maintenance is that maintenance normally authorized and performed by designated maintenance activities in direct support of using organizations. This category of maintenance is limited to the repair of end items or unserviceable assemblies in support of using organizations on a return-to-user basis. Direct support also furnishes supplies and other services directly to the user. Direct support units are designed to provide close support to combat troops and to facilitate tactical operations. This mobility requirement limits the equipment and supplies, and, therefore, the repair jobs that can be undertaken.

At this level, failed components and equipments are repaired by replacement of parts and subassemblies. These units are authorized larger amounts of repair parts and maintenance equipment than the using organization which the unit supports.

5-4.1.3 General Support Maintenance

General support maintenance is that maintenance authorized and performed by designated organizations in support of the Army supply system. Normally, general support maintenance organizations will repair or overhaul material to required maintenance standards' in a ready-to-issue condition based upon applicable supported Army area supply requirements.

This level of maintenance is performed by units organized as semifixed or permanent shops. They also serve lower levels within a given geographical area. These units perform work that overflows from direct support units, but rarely deal directly with the equipment user. The primary function of a general support unit is to repair those items that cannot be repaired by direct support units.

A high degree of specializationcan be expected at the general support level of maintenance. Mobility requirements are also less stringent and permit more complex maintenance operations.

TABLE 5-1.
CLASSIFICATION OF MAINTENANCE POLICIES AND GOALS

A. OPERATIONAL STATES
  1. Inactive Period
  2. Scheduled Downtime Period
  3. Operational Demand Period
    a. Standby
    b. Alert
    c. Reaction
    d. Mission
    e. Deactivation

B. MAINTENANCE ACTIVITIES
  1. Preventive Maintenance
    a. Service
    b. Inspection
  2. Corrective Maintenance
    a. Detection
    b. Diagnosis
    c. Correction
    d. Verification
  3. Maintenance Level
    a. Organizational
    b. Direct Support
    c. General Support
    d. Depot

C. RESOURCES
  1. Personnel
    a. Operators
    b. Maintenance Technicians
  2. Equipment
    a. Prime
    b. Support
  3. Facilities
  4. Repair Parts and Supplies
  5. Information (Publications and Data)

D. EFFECTIVENESS
  1. Downtime
    a. Detection Time
    b. Diagnostic Time
      (1) Localization
      (2) Isolation
    c. Correction Time
      (1) Primary
      (2) Secondary
    d. Verification Time
      (1) Alignment and Calibration
      (2) Checkout
  2. Reliability
  3. Availability or Operational Readiness
  4. Dependability

5-4.1.4 Depot Maintenance

Depot maintenance is that maintenance which, through overhaul of economically repairable materiel, augments the procurement program in satisfying overall Army requirements, and, when required, provides for repair of materiel beyond the capability of general support maintenance organizations.

Depot maintenance level organizations are stable and mobility is not a problem. Depot maintenance may be performed in shops in the continental United States or in shops established by the overseas theater commander for selected items. This level of maintenance provides the major supply bases in an overseas theater with end items and with the parts and supplies required to maintain and repair end items. Facilities are available for completely overhauling and rebuilding equipment. Depot maintenance functions also include repair and reclamation services that are beyond the capabilities of general support maintenance.

5-4.2 MAINTENANCE POLICIES

Maintenance policies concerning the operational states of the system dictate the allowable maintenance tasks and actions which may be performed at the various maintenance levels during the different operational time periods of the system and its equipment, as described in par. 2-2.6.1. Whether the system is performing a mission, is in a standby or alert period, or is in an inactive or scheduled downtime period restricts the maintenance and support activities which can be allowed during each of these operational time periods. It is necessary, therefore, for maintenance policies to be specifically stated in system requirement documents in order to guide the system and equipment designers and logistic support planners.

Maintenance policies concerning maintenance and support activities include policies about preventive and corrective maintenance tasks and the maintenance levels at which these tasks may be performed. In addition to dictating which maintenance actions are allowed to be performed at each maintenance level and during which operational states, these policies give specific guidance to maintainability engineers.

Maintenance policies with regard to resources indicate to the maintainability engineer the skill levels of personnel, both operators and maintenance technicians, which must be available at the various mainte-

TABLE 5-2.
CATEGORIES OF MAINTENANCE IN A THEATER OF OPERATIONS
(Ref. 11)

Category			Direct Support Maintenance	General Support Maintenance	Depot Maintenance
Former Echelon	First	Second	Third	Fourth	Fifth
Done Where	Wherever the Equipment is	In Unit	In Mobile and/or Semi-Fixed Shops		In Base Depot Shop
Done by Whom	Operator	Using Unit	Division/Corps/Army		Theater Commander Zone and/or Z/I
On Whose Equipment	Own Equipment				
Basis	Repair and Keep it		Repair and Return to User or Stock		Repair for Stock
Type of Work Done	Inspection Servicing Adjustment Minor Repairs and Modification		Inspection Complicated Adjustment Major Repairs and Modification Major Replacement Overload from Lower Echelons		Inspection Most Complicated Adjustments Repairs and Replacement Including Complete Overhaul and Rebuild Overload from Lower Echelons

nance levels and for which the system must be designed. Resource policies also indicate maintenance concepts with regard to test and support equipment, such as the extent to which built-in test techniques will be used, the extent to which automatic test and checkout is allowed, policies concerning general purpose vs special test equipment, calibration equipment requirements, and policies relating to tools and other support equipment necessary to the maintenance support of systems and equipments.

Resource policies dictate facility concepts, spares and repair parts policies including inventory control and stocking levels and locations, and repair/discard criteria. They also include policies concerning maintenance information such as technical manuals, provisioning documentation, and field data. These policies assist not only the maintainability engineer but logistic supply support and facility planners as well.

Maintenance policies concerning system effectiveness include many of the quantitative requirements which bear upon system availability and operational readiness. These include such measures as availability, dependability, reliability, and maintainability. They also include allowable preventive and corrective maintenance downtimes and logistic supply delay times. Effectiveness policies are the bridge between operational and support requirements for the reliability and maintainability engineers and system/equipment designers. They have a significant impact on maintainability design.'They provide further guidance for prediction, allocation, demonstration, test, and evaluation of maintainability as the system moves from the requirement phase to actual development, design, test, production, and operational phases.

Table 5-3 lists some representative examples of maintenance policies which might be included in the maintenance concept for a system. All of these policies do not apply to every system. Actual policies to be used for a given system are derived as part of logistic support and maintenance support planning, as described in par. 5-2 and as illustrated in Fig. 5-1. These policies should be combined in logical sets, as described in Ref. 10, in such a manner as to optimize operational readiness (availability) and life cycle cost.

5-5 DEVELOPMENT OF MAINTAINABILITY DESIGN REQUIREMENTS

In Chapter 1, maintainability engineering is described as including those actions taken by a system or equipment designer during engineering development to assure that the system/equipment, when produced, installed, and operated, can be effectively maintained. In order to determine what information is of importance in designing for maintainability, it is necessary to delineate those factors which, in combination, make up maintenance tasks or actions.

Maintenance activities may be partitioned into two major subsets—preventive maintenance and corrective maintenance. Preventive maintenance is that maintenance performed, preferably on a scheduled basis, for the purpose of retaining an item in a satisfactory operating condition. It includes periodic test, monitoring, servicing, and inspection. Corrective maintenance is that maintenance performed to restore an equipment to operating condition after a failure or other malfunction has occurred. Corrective maintenance includes detection, diagnosis, correction, and verification (Refs. 10 and 12). Detection of a fault may also occur during preventive maintenance. The relationship between these primary subsets of maintenance is illustrated in Fig. 5-2. **A** more detailed partitioning of corrective maintenance activities, including the secondary maintenance loop for rear echelon repair, is shown in Fig. 5-3.

Although design for maintainability must include both preventive and corrective maintenance considerations, critical problems often center around corrective maintenance since this involves the restoration of failed items to an operable state-often during a mission and within a relatively short time period. It is evident that time is the critical parameter in corrective maintenance and, therefore, an essential factor in maintainability design.

Since maintainability is associated with the design features of a system/equipment in order to facilitate maintenance, we would like to analyze maintenance tasks as a function of time and provide those maintainability design features which will minimize maintenance task times. Time enters maintainability considerations in two ways:

1. In terms of long-term or life characteristics of the system, e.g., reliability characteristics and time between overhaul
2. In terms of short-term characteristics or the ability to keep an operating system in operation (preventive maintenance), or to restore an inoperable system to operational status (corrective maintenance).

5-5.1 MAINTENANCE TIME PHASES

Corrective maintenance tasks may be separated into the following sequential time phases (Refs. 10 and 12):

TABLE 5-3.
EXAMPLES OF MAINTENANCE POLICIES

1. Preventive maintenance actions will be performed only during scheduled downtime periods, except for mission-ready test and checkout performed prior to the start of a mission.

2. Checkout, alignment, adjustment, and minor corrective maintenance actions which can be performed within the expected standby interval may be performed during the standby period.

3. Only checkout and minor adjustments may be performed during alert periods. Such checkout and minor adjustments will not require opening or disassembly of equipment.

4. No maintenance will be performed during the mission period. Alternate modes or degraded performance will be used, or redundant units which are automatically switched into operation for failed items will be provided in order to meet critical system performance and effectiveness requirements.

5. Only urgent corrective maintenance, limited to replacement of readily accessible plug-in modules, will be permitted during the mission period.

6. Servicing, adjusting, calibration, and other preventive maintenance actions, as well as deferred corrective maintenance and equipment modification, may be performed during inactive time periods.

7. During operational demand periods, maintenance actions will be limited primarily to those which can be performed by the operator or organizational repairman. Requests for directional support assistance during operational demand periods will be minimized.

8. Critical performance functions will be monitored during operational demand periods.

9. Only general purpose or standard test and support equipment will be used to maintain equipment at organizational and direct support levels. Built-in test features will be used at the organizational level to the maximum extent feasible.

10. Because of mission and mobility requirements, those maintenance tests and adjustments which must be performed at the'organizational level will be incorporated into prime equipment. Only simple accessory hand tools will be required.

11. Simple, positive adjustments and indicators will be provided for organizational level test and alignment. Indicators will be based on go/no-go, lo-go-hi, or reference mark indications. Quantitative measurements shall not be required.

12. Precise alignment and calibration of equipment will be performed at the direct support or general support levels. Organizational level repairmen will not be required to perform such calibration nor carry calibration equipment.

13. Repairs at the organizational level will be made by replacement of failed modules or end items only. Failed items will either be discarded or sent to the direct support level for interchange and repair in accordance with repair/discard policies.

TABLE 5-3.
EXAMPLES OF MAINTENANCE POLICIES (Cont.)

14. Redundant standby units will be provided to maintain system operation in the event of a failure of a mission critical item. In the event of a failure in a redundant unit in a system that must operate continuously with no allowable inactive or scheduled maintenance time, restoration of the failed unit will be possible while the system is performing its function. Such restoration activities will not degrade system performance nor be hazardous to maintenance personnel.

15. Maintenance by operators of equipment will be limited to simple visual checks, inspections, servicing, cleaning, lubrication, and adjustment requiring only simple hand tools and built-in test features. No detailed test and repair skills will be required of operators.

16. Maintenance by organizatiohal repairmen will be limited to simple detection and diagnostic routines to replace a failed module with a spare, plus any necessary alignment and checkout. Organizational repairmen also may perform more detailed preventive maintenance tasks and checkouts beyond the capability of the equipment operator. No detailed test and repair skills will be required.

17. Prime equipment will be designed to have ready access for maintenance. Quick-opening fasteners will be used.

18. Insofar as possible, provision will be made to store small handtools and replacement parts – such as fuses, pilot lights, and plug-in items – in the equipment.

19. Replaceable items will be plug-in and require a minimum amount of clamps or fasteners consistent with environmental requirements.

20. Repairs at the organizational and direct support levels will not require special facilities, such as special energy requirements or clean rooms.

(NOTE: In the following policies, the letters X and *Y* should be replaced by actual system/ equipment quantitative requirements. Where two or more items or parameters are given in parentheses, the appropriate one should be selected.)

21. Modules whose replacement cost is less than X dollars will be designed for discard at failure. ***(X to be determined during system definition studies.)***

22. Items whose replacement cost is more than *Y* dollars will be designed for repair. ***(Y to be determined during system definition studies.)***

23. Items and modules whose replacement cost is between X and *Y* dollars will be designed for either discard or repair in accordance with selected repair/discard criteria. Such items will be repaired or discarded at the appropriate maintenance level as determined by repair/discard trade-off studies.

24. The ***MTTR*** will not exceed X minutes.

25. The Maximum Repair Time ***(percentile to be specified)*** will not exceed *Y* minutes.

TABLE 5-3.
EXAMPLES OF MAINTENANCE POLICIES (Concl.)

26. The system/equipment ***(inherent, operational)*** availability will be at least 0.*XX*.

27. System dependability will be at least 0.*YYY* when operating in the specified mode.

28. Mission critical parameters will be ***(periodically monitored, automatically sensed)*** and an alarm given to indicate failure within X minutes after the failure has occurred.

29. A system failure will be localized to the ***(equipment, unit, assembly, module)*** level within X minutes after the failure has been detected.

30. A system failure will be isolated to the replaceable/repairable item or module within X minutes after the failure has been localized.

31. A failed item/module will be replaced/repaired within X minutes after the failure has been isolated.

32. It will be possible to align the repaired item and verify system/equipment effective operation within X minutes after the correction has been completed.

33. The probability of having a repair part or replaceable module when needed at the organizational level will be at least 0.*XX*.

34. The probability of having a repair part or replaceable item when needed at the direct support level will be at least 0.*Y*Y.

35. Repair parts will be carried at the organizational or direct support levels for those repairable items which have an MTBF of less than *Y* hours.

36. Organizational level maintenance will be limited to those tasks which can be performed by an organizational repairman with the following skill levels: ***(specifics to be furnished during system conceptual and definition studies.).***

37. Direct support maintenance will be limited to those tasks which can be performed by a maintenance technician with the following skill levels: ***(specifics to be furnished during system conceptual and definition studies).***

38. Requirements for special training for organizational and direct support level maintenance personnel will be minimized and will be consistent with defined general aptitudes and skill levels of such personnel.

39. Preventive maintenance tasks performed at the organizational and direct support levels will be able to be accomplished within the following defined time periods: ***(to be specified from defined operational demand profiles and logistic endurance factors).***

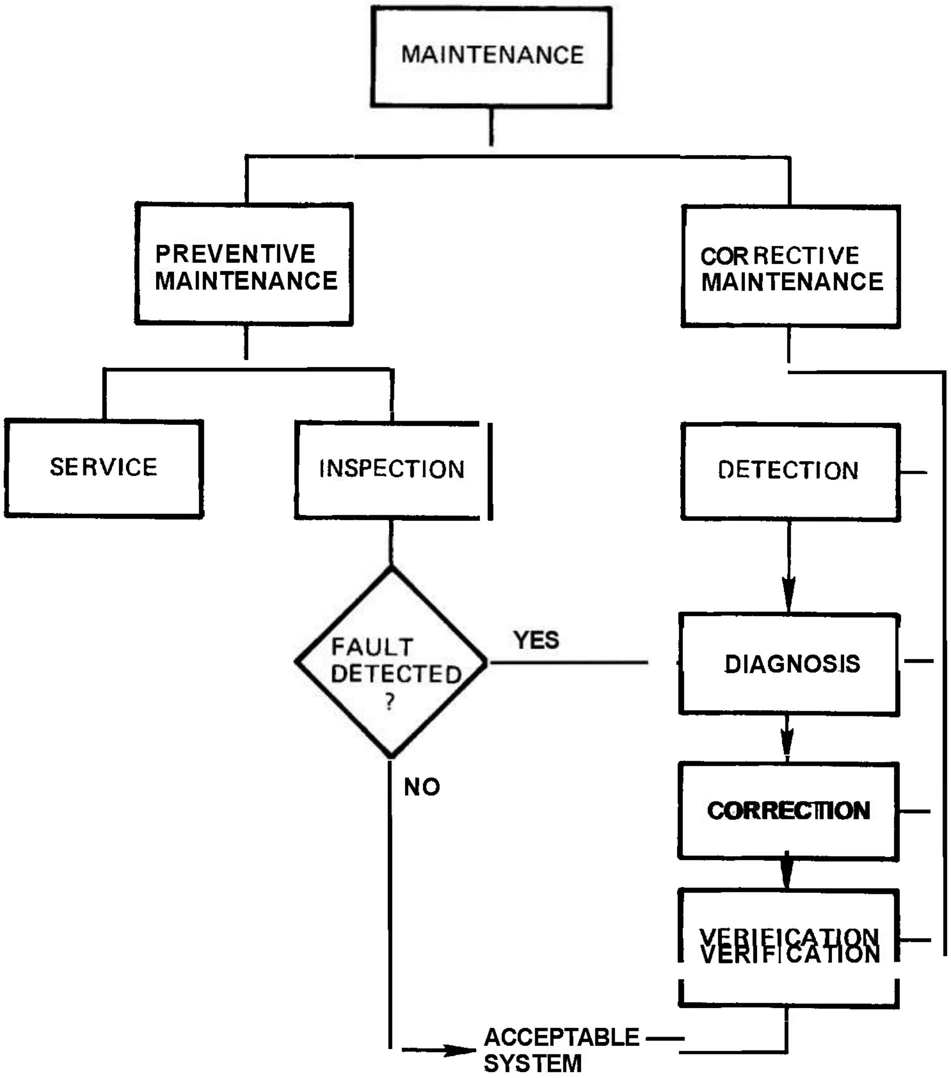


Figure 5-2. Primary Subsets of Maintenance Activities

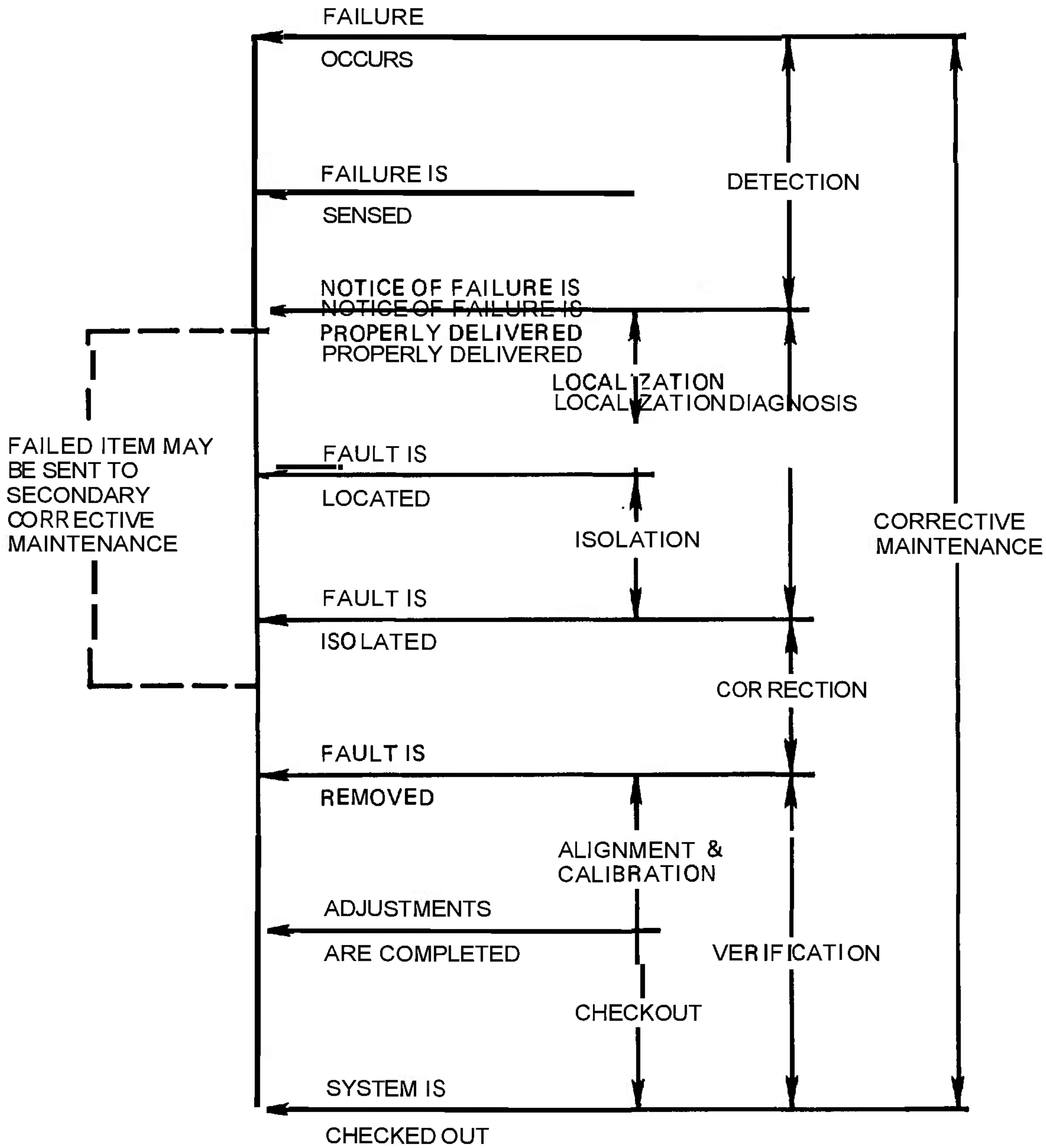


Figure 5-3. Major Events and Activities Comprising Corrective Maintenance

1. Detection time—recognition of a fault
2. Diagnostic time—fault location and isolation
3. Correction time—replace or repair
4. Verification time—test and align.

These four maintenance time phases may be compared to similar phases in the treatment of a patient by a doctor. The *detection* phase corresponds to the symptomatic phase in the human; a fault may have occurred and must be detected by some means of sensing symptoms or symptomatic responses before any corrective action may be prescribed.

A fault may be classified in one of three categories. First, the fault may be one that allows the system to perform but at a reduced capability or efficiency; this is called *degradation.* Initially, the result of the fault is a reduction in performance capability. If allowed to continue, however, the degradation may result in either complete loss of performance or a permanent state of reduced performance capability.

The second type of fault is *critical* failure in which there is either a reduction in performance below acceptable levels or a complete lack of performance. In this case, performance within acceptable limits cannot be restored without taking corrective action.

The third type of failure is due to the Occurrence of a *catastrophic* event, from which there is no recovery; this failure occurs precipitously.

These types of failures may occur in any system. In a military system, degradation is a reduction in system performance capability below a prescribed minimum level of effectiveness. This type of failure can be inhibited or minimized by proper preventive maintenance action. With regard to design for maintainability, this means providing for periodic test, inspection, or servicing, or for monitoring certain critical performance parameters either continuously or periodically.

Critical failures are correctable by taking corrective maintenance actions, which require means for locating and isolating the fault to a replaceable/repairable item, correction of the fault, and verification of system performance. Maintainability design features for this include accessibility, test points, test equipment, displays, connectors, fault indicators, maintenance instructions, and other features which are discussed in Section II of this chapter. In the case of a catastrophic event, the equipment may be nonrepairable and must be replaced.

The *diagnostic* phase, the second of the maintenance time phases, includes localization and isolation of the fault in order that proper corrective measures may be taken.

The third time phase, the *correction* phase, is that time period in which something is actively done to remove the fault and to restore the system to acceptable operating condition. Correction is a matter of gaining access to the faulty item and of removing it, and replacing it with a good item or repairing it in place.

The fourth and final time phase, the *verification* phase. This includes alignment, adjustment, calibration, test, and final verification by checkout.

5-5.2 CORRECTIVE MAINTENANCE DOWNTIME

There are two types of downtime of concern in the corrective maintenance (restoring) operation. One is called *active downtime* or *active repair time,* depending upon whether detection time is included or not, during which repair actions described in par. 5-5.1 are actively taken by the maintenance technician. This is illustrated in Fig. 5-4. The other time category is *waitingor delay time,* during which the maintenance technician is able to do little or nothing towards actively restoring the equipment to operating condition. Delay time is normally defined to include administrative time and logistic supply time.

It is helpful to distinguish between these two types of time for a number of reasons. First, *active repair time* can be controlled by design. *Delay time,* to a large extent, cannot, being primarily a function of operating and environmental conditions of the system and the availability of personnel and system resources. Second, active repair times usually may be described by a statistical distribution such as the lognormal, while delay times may be generally characterized by a distribution such as the exponential distribution.

As defined in Chapter 2, *operational availability* includes both types of time while *intrinsic availability* includes only elements of active repair time. Because of the difficulty of measuring delay times and demonstrating operational availability during system test and evaluation, most specifications use active repair time measures, such as *MTTR* and M_{MAX} (see par. 1-6), and intrinsic availability as requirements for maintainability and system effectiveness. In recent years, attempts have been made to use operational availability and mean active downtime or mean downtime (see par. 2-2.4.2) to more nearly simulate actual operating conditions.

In general, maintenance actions taken during the four corrective maintenance downtime periods and the maintainability design characteristics required to minimize these downtime periods are different. These are discussed in par. 5-8.

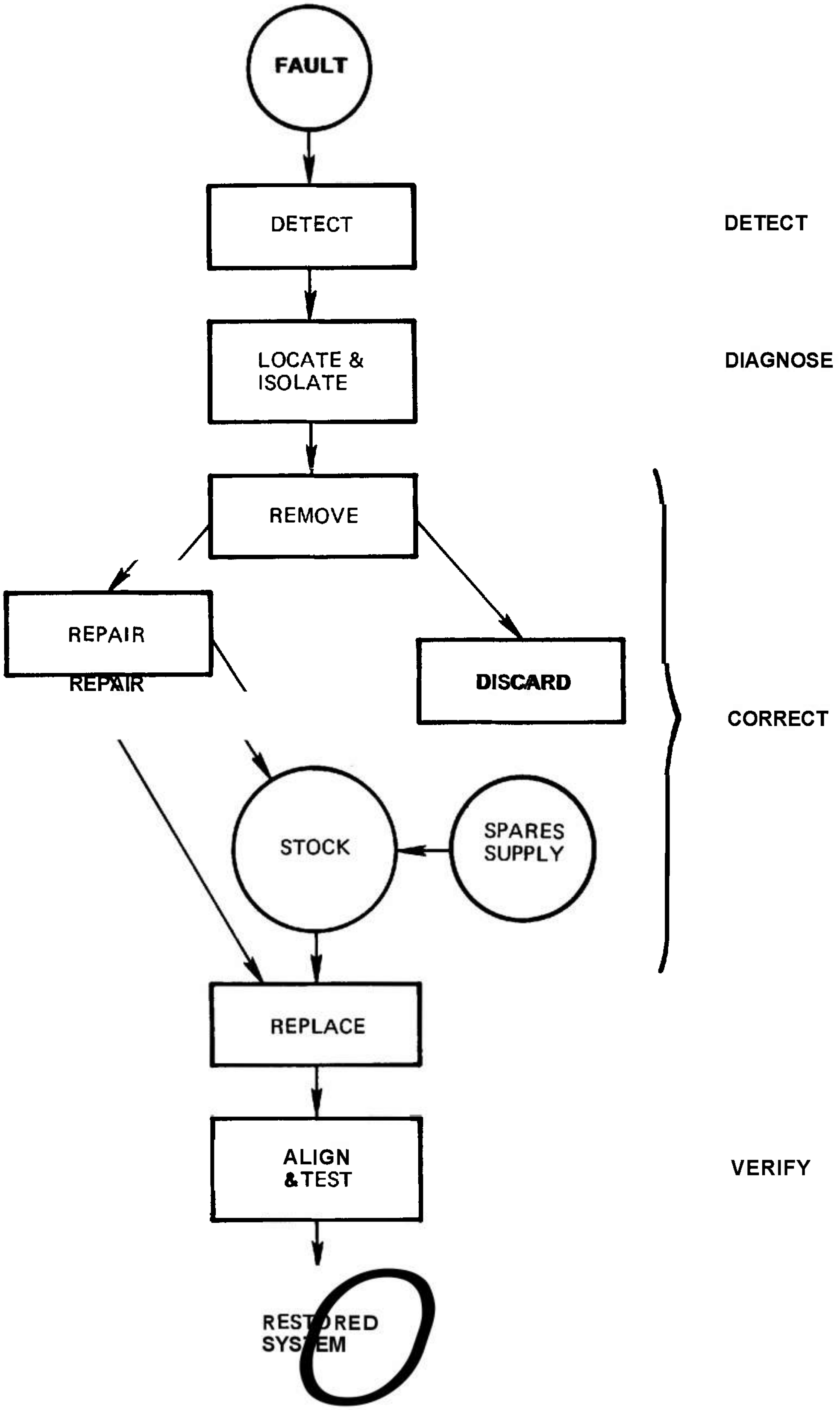


Figure 5-4. Steps in Active Downtime

SECTION II

MAINTAINABILITY DESIGN CONSIDERATIONS

5-6 MAINTAINABILITY AND SYSTEM DESIGN

In Chapter 3, it is pointed out that maintainability considerations must be included in all phases of the system life cycle if a cost-effective, supportable system design is to be achieved. The means for carrying out such a life cycle system approach is the ***system design process*** (Fig. 5-5). In order to examine the interactions of maintainability with system design (Fig. 1-1), one must first understand the system design process.

At the beginning of their book ***System Engineering,*** Goode and Machol indicate that "for more than a decade, engineers and administrators have witnessed the emergence of a broadening approach to the problem of designing equipment. This phenomenon has been poorly understood and loosely described. It has been called ***system design, system analysis,*** and often ***the systems approach"*** (Ref. 13). In the ensuing years since Goode and Machol's book was published, there have been numerous efforts to describe system engineering and system design (Refs. 10 and 14-17). As a result, a clear pattern of the system design process is now evident.

An accepted definition of engineering design is: "Design is defined as an iterative, decision-making process for developing engineering systems or devices whereby resources are optimally converted into desired ends" (Refs. 14 and 17). Fig. 5-5 is a model of this definition of the system design process. It represents a feedback control system for transforming a set of inputs into outputs in an optimal (economic) manner, within allowable constraints, in order to meet stated needs in accordance with a defined measure of system worth or effectiveness. It is applicable throughout all the life cycle phases. A more complete description of the system life cycle and the system design process is contained in Refs. 10 and 17.

The input to the design process is information—information about the need for the system, the system operational environment, constraints on the system, its design, use, and support, and any other pertinent information.

The step labeled "formulation of a value model" is often called "defining the problem". It involves gathering and organizing the pertinent information about the system objectives and constraints. But more than defining the problem, it also involves the essential task of formulating the criteria of system worth or effectiveness (the value model) by which the system alternatives will be evaluated. Without such evaluation criteria, system optimization is not possible.

Once the need (problem) has been defined and the system effectiveness criteria established, alternative means for satisfying the system requirements may be synthesized. These alternatives are then analyzed or tested and the results evaluated against the established effectiveness criteria. A decision may then be made as to whether the design is optimal or whether iteration is needed. It is a rare occurrence for an optimal design to be achieved the first time through the design process. Rather, a number of iterations are usually required in which design parameters are varied in order to meet stated performance requirements or to reduce the uncertainty of unacceptable or marginal performance. This iterative process is called optimization and is the feedback loop shown in Fig. 5-5. Occasionally, the result of evaluation and iteration will require that the value model (system effectiveness criteria) be modified. This is shown as the dotted feedback path in Fig. 5-5. Risk analysis techniques should be employed in the decision process. Ref. 92 provides basic information on risk analysis.

When the design has been frozen, it is then communicated to others for implementation. Such communication is in the form of drawings, specifications, reports, test and acceptance procedures, manufacturing instructions, installation and support information, operating and maintenance instructions, personnel and training requirements, etc.

The system design process may be applied to maintainability design. Maintainability engineering, then, is concerned with the logical processing of those system design factors about which maintainability is concerned, and the techniques for so doing in accordance with the system design process.

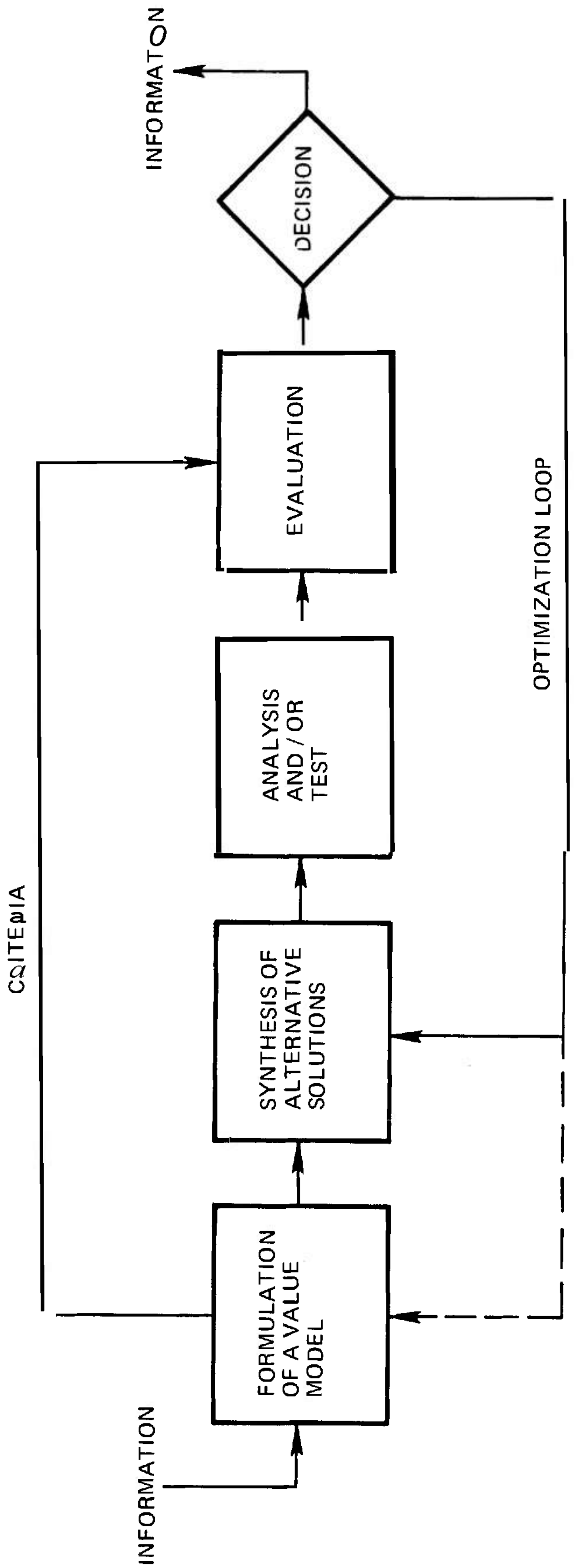


Figure 5-5. The System Design Process

5-6.1 INPUT INFORMATION—THE BACKGROUND FOR MAINTAINABILITY DESIGN

Input requirements with respect to maintainability are often incomplete and in primitive form. In order to utilize this information during system design, the maintainability engineer must find answers to the following statements:

1. Why the system is being designed—the operational need.

2. What the environmental and policy considerations are—operational and resource constraints, maintenance policies, integrated logistic support concept, applicable maintainability standards, design handbooks, and guidelines.

3. What the maintenance objectives are—maintenance concepts, system effectiveness requirements, cost, and other support criteria.

4. When the system can be maintained—mission profiles, logistic endurance factors, preventive vs corrective maintenance.

5. Where it is to be supported–organizational, direct support, general support, or depot levels.

6. How it can be supported—repair/replace/discard policies, depth of maintenance, use of standby redundancy, periodic test and checkout, overhaul.

7. Who is to support it-operators, organization repairmen, rear level maintenance' technicians.

The maintainability engineer should obtain field feedback information on similar systems and environments to assist in the formulation of answers. The Army and other service data banks should be carefully searched for pertinent data; The Army Maintenance Management System (TAMMS) is such a bank. In the past, maintenance feedback information from the field has generally been limited in scope and difficult to use for design analysis. A number of programs have been initiated by the military services and industry to improve the quality and reliability of such information.

5-6.2 FORMULATION OF THE VALUE MODEL (EFFECTIVENESS CRITERIA)

The input information is used to formulate both qualitative and quantitative maintainability objectives during system synthesis, evaluation, and optimization in order to achieve the best possible maintainability design within the established constraints. Quantitative maintainability criteria such as allowable downtimes, turn-around times, time between overhauls, maintenance and support costs, and repair/discard criteria contribute to the establishment of measures of effectiveness to be used in evaluating the system design for maintainability during each of the life cycle design stages. Quantitative maintainability criteria are discussed in Chapters 1, 2, and **4.**

Of particular importance during the formulation of the value model is the analysis of maintenance policies and goals stated in system requirement documents, along with such quantitative system effectiveness requirements as availability, dependability, and mission reliability, and the defined operational capability of the system. It is the maintainability engineer's role to perform these analyses and to interpret the design requirements to the equipment designers.

5-6.3 SYNTHESIS OF MAINTAINABILITY DESIGN MODELS

The synthesis of models useful for maintainability design is complicated by the fact that not all of the physical variables with which maintainability is concerned can be quantified. In addition, maintainability is concerned with such man-machine interfaces as human engineering factors and safety, which, when measurable, often require the use of subjective and stochastic measures or simulation techniques.

No simple or general model of maintainability design is available. However, it is possible to construct a conceptual model combining those items of concem to maintainability engineering. For example, maintainability engineering is certainly concerned with minimizing system (end item) maintenance downtimes. It is also concerned with the system level at which maintenance actions will be performed and with the specific maintainability attributes of the system. Finally, it is concerned with the maintainability design activities which are performed at each of the design stages of the system life cycle. Such a model is discussed in par. **5-7.**

5-6.4 MAINTAINABILITY ANALYSIS

Maintainability analysis is principally concerned with the prediction and demonstration of the resultant efforts of maintainability design considerations, usually measured by calculating, estimating, or measuring downtime under simulated operating conditions. It is also concerned with the determination of the effects of maintainability design characteristics of the equipment on required maintenance resources.

Typical analytic techniques include simulation, maintainability prediction, allocation, and demonstration tests depending upon the design stage of the system life cycle. Statistical techniques are required. Chapters

1, **4,** and 6 discuss measures of maintainability and existing prediction and demonstration techniques. In addition to the current techniques described in Military Standards (Refs. 18 and 19), the prediction and evaluation techniques need to be developed which can be more usefully applied during the early life cycle planning and design phases. Also needed are quantitative data representative of current design packaging techniques such as the use of solid-state devices, microelectronics, and mechanical items.

Maintenance engineering analysis (Refs. 3 and 20) is also an important analytical tool for maintainability analysis during the design stages for determining both maintenance task times and actions as well as for determining maintenance resources.

5-6.5 MAINTAINABILITY EVALUATION

Maintainability evaluation compares the results of maintainability analyses against system effectiveness criteria in order to obtain a decision as to whether the design is acceptable or further iteration' is desired. Maintainability must be evaluated not only against its own stated requirements and reliability, but also against higher level system effectiveness requirements such as availability, dependability, and reliability.

The earliest life cycle evaluations are concerned with maintainability predictions and allocations from the system level to equipment levels. These occur during system definition and preliminary design stages. Lower level allocations and predictions follow during detail design stages, and finally, the results of maintainability tests and demonstrations are evaluated during the test and evaluation stage. Additional maintainability evaluations may occur during operational system testing by the user.

5-6.6 DECISION

Decisions must be made during each of the design stages of the system life cycle. These decisions will normally be either to leave one design stage and enter the succeeding one or to iterate the existing stage because some criterion has not been optimally met or because some constraint boundary has been exceeded. Within each design stage many subdecisions will be made. Occasionally, the decision will be to return to a preceding stage.

Both interior and exterior decisions must be made with regard to maintainability design. Interior decisions are those which are made to reduce downtime or to modify some design attribute to effect a change or trade-off in one or more of the maintenance task times, independent of other system parameters. Exterior decisions are those which affect other system parameters such as reliability, safety, or supply support. Risk analysis techniques should be employed in reaching decisions.

As each design stage milestone is reached, certain design data and other information should be evolved and presented in proper form to serve as a basis for design review (Ref. 21). These serve to facilitate program management decisions.

5-6.7 OPTIMIZATION

As mentioned earlier in this section, optimization is the iterative feedback loop which is used to modify the system model, analyze the resulting change, evaluate, and decide. The process is repeated until the marginal cost of additional iteration is no longer commensurate with the expected increase in benefits (effectiveness).

It is possible to optimize maintainability requirements independently of other system parameters based on specified maintainability criteria only. This is maintainability suboptimization. If availability or dependability are specified as the system effectiveness requirements, then it is possible for the system engineer to allocate and trade-off reliability and maintainability requirements, a higher level optimization than that of maintainability alone.

A trade-off may or may not be an optimization. The difference is whether the value of the higher level criterion function which relates the items being traded off is changed or not. For example, if a trade-off is made between reliability and maintainability for a constant availability (along an availability isocline), then no optimization has been effected with respect to system effectiveness; we have just swapped or "traded off" one resource (reliability) for another (maintainability) with no increase in benefit. If, on the other hand, a trade-off is made which increases the value of availability, then optimization is being performed. Such trade-offs are described in Refs. 12 and 22. Additional discussion of trade-offs and trade-off techniques is given in par. 5-8.7.

5-6.8 OUTPUT INFORMATION

The final output of the system design process is information. This information includes the design characteristics appearing in the detailed drawings reflecting the finished design. These design characteristics include maintainability features. The output information also includes the data resulting from maintainability design analyses, predictions, demonstration tests, evaluations, design reviews, maintenance engineering analyses, and other pertinent information. The total maintainability engineering effort should indicate, with a high degree of confidence, that systems/equipments produced and installed in accordance with the design data package will operate with the required effectiveness if maintained and supported as specified.

At each stage of the design process, certain information must be available to allow design review and evaluation. It should be presented in such a format as to facilitate the evaluation and decision-making process. Examples of the kinds of information required for design reviews are given in Refs. **20, 21**, and **23.**

Information for reliability and maintenapce plans is part of this output.

5-7 A MODEL OF MAINTAINABILITY DESIGN

Maintainability design is concerned with providing those system/equipment features which will facilitate preventive and corrective maintenance. Figs. **5-2** and **5-3** indicate that both preventive and corrective maintenance are concerned with similar activities. The emphasis in preventive maintenance is (1) to service those system elements known to have short wearout lives and which can be expected to fail or degrade in a succeeding time period unless serviced, and (2) to inspect those system elements or their performance whose failure or performance degradation is critical to system operations and mission success. The emphasis in corrective maintenance is to promptly detect, diagnose, and correct a system failure, and to verify that the system has been restored to proper operation.

Maintainability engineering is concerned with providing those features in the design of a system which will facilitate maintenance activities at the system maintenance level that will be effective. It is also concerned with how these system attributes should be considered at each stage of design.

5-7.1 MAINTAINABILITY DESIGN REQUIREMENTS

It is readily apparent that the design features to minimize each of the active downtime segments discussed in par. 5-5 will differ to a considerable extent. For example, to minimize detection time, one would want to consider such equipment features as performance monitoring and failure alarms. To minimize diagnostic time, one should provide adequate test points, test equipment, and logical troubleshooting procedures. To minimize fault removal, one should provide for rapid access, ease of replacement, plug-in spares. To minimize verification time, simple and unambiguous mechanical alignment procedures, adequate controls and indicators and rapid, logical checkout features should be provided.

It is possible to consider separately the equipment characteristics for each of these major downtime areas and to delineate the design requirements for them at each stage of the system life cycle. This should make evident both independent and interdependent optimization (trade-off) possibilities. (Use the DOD Work Breakdown Structure (WBS).)

Maintainability design may be considered to be a function of four primary parameters—maintainability design factors f, maintenance task times t, system level at which maintenance is performed ℓ, and life cycle design stage s (Ref. 10). Thus:

$$M = f(f, t, \ell, s) \tag{5-1}$$

As a first step in formulating a model of maintainability design, the relationships among these primary vectors must be delineated. These may be shown in a four-dimensional dependency matrix, as in Fig. 5-6. Each of these primary vectors is, in turn, multidimensional.

The intersections depicted by the X's on each of the plane surfaces indicate that the two factors are related. A projection of the X's into the third or fourth dimensional volumes further indicates that there is a third or fourth parameter relationship existing. For example, tools are of concern at the assembly level, detailed design stage, and corrections task time, but not at the system definition stage or during detection time.

Since the use of a four-dimensional conceptual model, as indicated in Fig. 5-6, offers practical difficulties, it is helpful to unfold the six interrelated planes into flat two-dimensional representations. Matrices of maintenance task times, maintainability design factors, system levels, and life cycle design stages are shown in the unfolded views of Figs. **5-7** and **5-8.** (NOTE–These figures are presented as a conceptual illustration and not as a guide.) We will delay the discussion of Figs. 5-7 and 5-8 until Tables **5-4** and 5-5 have been introduced.

A search of a number of existing maintainability design guides and related documents indicates those primary equipment characteristics which should be emphasized during design for maintainability (Ref. 10). These are listed in Table **5-4** and discussed in detail in Ref. 11 and also in Refs. **24-38.** Those most frequently

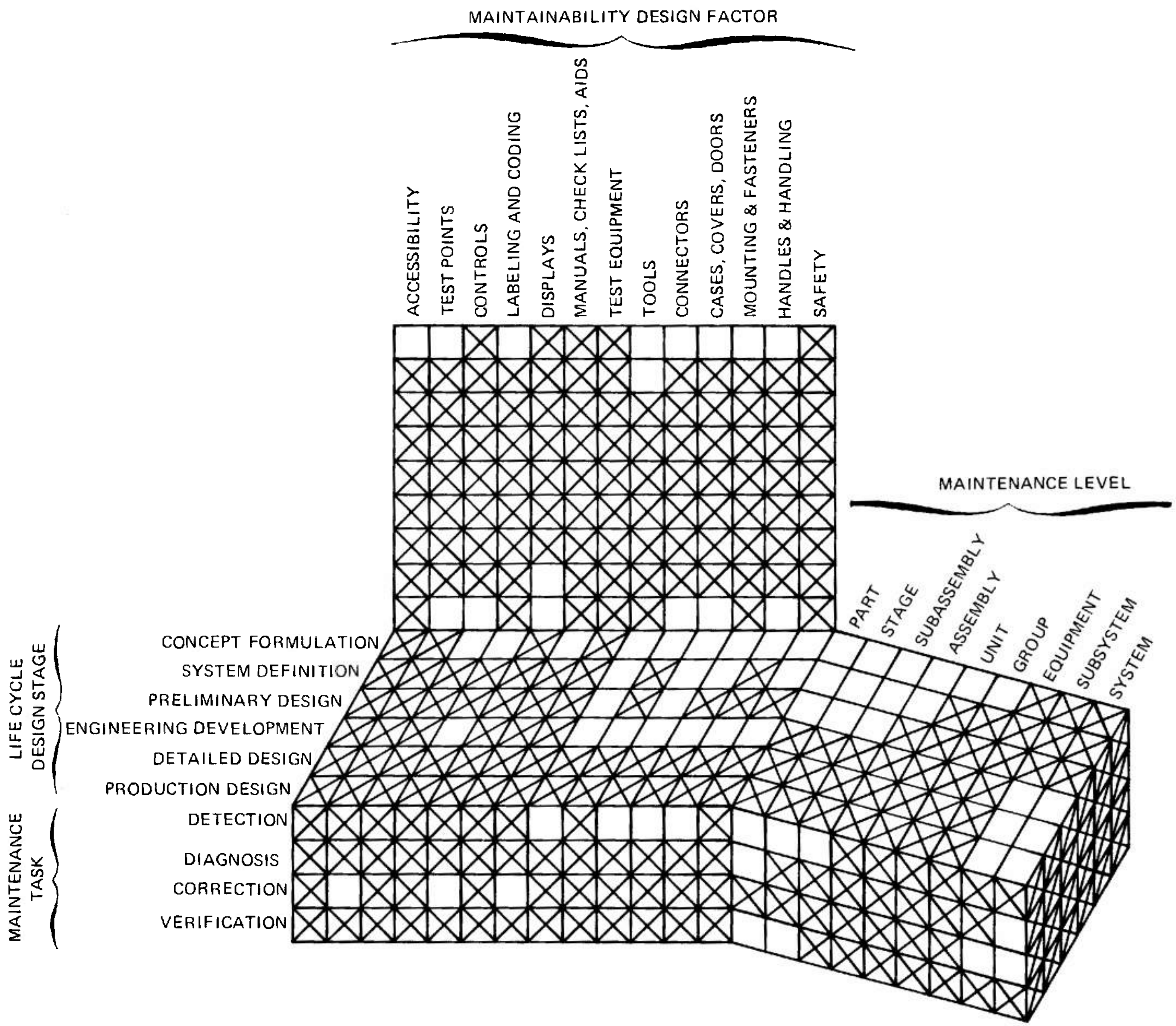


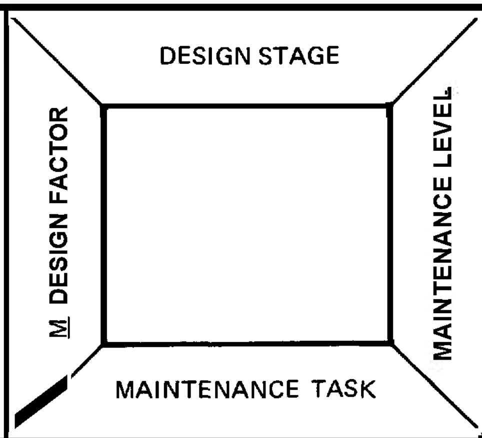
Figure 5-6. Maintainability Design Relationships

MATRIX 1

	ACCESSIBILITY	TEST POINTS	CONTROLS	LABELING & CODING	DISPLAYS	MANUALS, CHECKLISTS, AIDS	TEST EQUIPMENT	TOOLS	CONNECTORS	CASES, COVERS, DOORS	MOUNTING & FASTENERS	HANDLES & HANDLING	SAFETY
CONCEPT FORMULATION	1	1			3		3						
SYSTEM DEFINITION	4	3	1	1	4	1	4		1			1	
PRELIMINARY DESIGN	4	4	3	1	4	3	4		3		1	1	1
ENGINEERING DEVELOPMENT	1	4	4		4	3	3						
DETAILED DESIGN	4	4	4	4	4	4	4	4	4	4	4	4	4
PRODUCTION DESIGN	4	4	4	4	4	4	4	4	4	4	4	4	4

MATRIX 3

	SYSTEM	SUBSYSTEM	EQUIPMENT	GROUP	UNIT	ASSEMBLY	SUBASSEMBLY	STAGE	PART
CONCEPT FORMULATION	4	4	3						
SYSTEM DEFINITION	4	4	4	3	1				
PRELIMINARY DESIGN	1	4	4	3	3	1			
ENGINEERING DEVELOPMENT			3	3	3	3	4	4	1
DETAILED DESIGN			4	4	4	4	4	3	3
PRODUCTION DESIGN			3	3	3	4	4	4	4



MATRIX 2

	ACCESSIBILITY	TEST POINTS	CONTROLS	LABELING & CODING	DISPLAYS	MANUALS, CHECKLISTS, AIDS	TEST EQUIPMENT	TOOLS	CONNECTORS	CASES, COVERS, DOORS	MOUNTING & FASTENERS	HANDLES & HANDLING	SAFETY
DETECTION	1	3	1	3	4	4	4		1				1
DIAGNOSIS	3	4	3	4	4	4	4	1	3	3	1	1	4
CORRECTION	4		1	4		4		4	4	4	4	4	4
VERIFICATION	3	4	3	4	4	4	4	1	3	3	1	1	3

MATRIX 4

	SYSTEM	SUBSYSTEM	EQUIPMENT	GROUP	UNIT	ASSEMBLY	SUBASSEMBLY	STAGE	PART
DETECTION	4	4	4	3	3	1			
DIAGNOSIS		4	4	4	4	4	3	1	
CORRECTION			4	4	4	4	4	3	1
VERIFICATION	3	4	4	4	3	3	1		

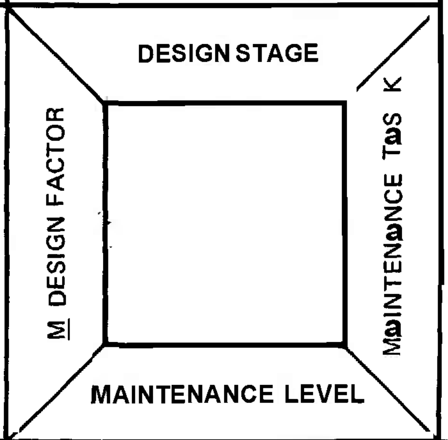
Figure 5-7. Maintainability Design Relationships

MATRIX 1

Design Stage	ACCESSIBILITY	TEST POINTS	CONTROLS	LABELING & CODING	DISPLAYS	MANUALS, CHECKLISTS, AIDS	TEST EQUIPMENT	TOOLS	CONNECTORS	CASES, COVERS, DOORS	MOUNTING & FASTENERS	HANDLES & HANDLING	SAFETY
CONCEPT FORMULATION	1	1			3		3						
SYSTEM DEFINITION	4	3	1	1	4	1	4		1			1	
PRELIMINARY DESIGN	4	4	3	1	4	3	4		3		1	1	1
ENGINEERING DEVELOPMENT	1	4	4		4	3	3						
DETAILED DESIGN	4	4	4	4	4	4	4	4	4	4	4	4	4
PRODUCTION DESIGN	4	4	4	4	4	4	4	4	4	4	4	4	4

MATRIX 6

Design Stage	DETECTION	DIAGNOSIS	CORRECTION	VERIFICATION
CONCEPT FORMULATION	3	3	1	3
SYSTEM DEFINITION	4	4	4	4
PRELIMINARY DESIGN	4	4	4	4
ENGINEERING DEVELOPMENT	4	4	4	4
DETAILED DESIGN	4	4	4	4
PRODUCTION DESIGN	1	3	3	3



MATRIX 5

Maintenance Level	ACCESSIBILITY	TEST POINTS	CONTROLS	LABELING & CODING	DISPLAYS	MANUALS, CHECKLISTS, AIDS	TEST EQUIPMENT	TOOLS	CONNECTORS	CASES, COVERS, DOORS	MOUNTING & FASTENERS	HANDLES & HANDLING	SAFETY
SYSTEM			1		4	3	3						1
SUBSYSTEM	1	1	3	3	4	3	4		4	3	4	1	4
EQUIPMENT	3	3	4	4	4	4	4	1	4	4	4	3	4
GROUP	3	3	4	4	4	4	4	1	4	4	4	3	4
UNIT	4	4	4	4	3	4	4	3	4	4	4	4	4
ASSEMBLY	4	4	4	4	3	4	4	4	3	4	4	4	4
SUBASSEMBLY	4	4	3	4	1	4	3	4	3	3	3	4	4
STAGE	3	3	1	4		3	3	4	1	1	3	3	4
PART	1			3		1	1	3			1		3

MATRIX 4

Maintenance Level	DETECTION	DIAGNOSIS	CORRECTION	VERIFICATION
SYSTEM	4			3
SUBSYSTEM	4	4		4
EQUIPMENT	4	4	4	4
GROUP	3	4	4	4
UNIT	3	4	4	3
ASSEMBLY	1	4	4	3
SUBASSEMBLY		3	4	1
STAGE		1	3	
PART			1	

Figure 5-8. Maintainability Design Relationships

mentioned characteristics are ordered in Table 5-5. In addition, relative importance of each of these characteristics for the various active downtime phases is shown on a cardinal ranking number as defined in Table 5-5. There is no ranking factor of 2. This is intended to indicate that there is a greater difference in weight between items ranked 3 and items ranked 1 than between those ranked 4 and 3 and items ranked 1 and 0. Although the 13 design factors chosen as most often mentioned obviously represent the collective judgment and efforts of many investigators, this does not imply that only these characteristics are important for any given system, nor necessarily in the order given. For any system, these will depend upon the defined maintenance and system operational policies and goals.

To translate the maintainability design relationship matrices into usable design tools, the ordinal rankings must be converted into a scale of cardinal values and then modeled into a set of analytic or empirical functions, where possible. These expressions may then be manipulated with the aid of analytic and computational techniques to arrive at the exact design details to be used. The importance of the ranking on an *ordinal* scale is to allow the maintainability engineer to focus his attention, particularly during early system design stages, on the more important design factors. In a sense, the *ordinal* rankings give a sensitivity dimension to the problem.

A study of these matrices—refer to Figs. 5-7 and 5-8—reveals the following maintainability design considerations:

1. ***Matrix I.*** A look at the horizontal rows shows that all maintainability design factors are important in the detailed design and production design stages. A look at the vertical columns indicates that displays and test equipment are important considerations in all design stages. Therefore, during concept formulation, at-

TABLE 5-4.
MAINTAINABILITY DESIGN CHARACTERISTICS

Manuals, Checklists, Charts, Aids	Equipment Units
Labeling and Coding	Interconnecting Wires & Cables
Test Equipment	No Maintenance Induced Faults
Tools	Sensitivity — Stability — Criticality
Test Points	Components
Functional Packaging	Interchangeability
Controls	Servicing Equipment
Adjustments and Calibrations	Size and Shape
Displays	Modular Design
Test Hookups	Cabling & Wiring
Test Adapters	Ease of Removal (and Replacement)
Marginal Checking	Operability
Weight	Personnel Numbers
Handles and Handling	Personnel Skills
Cases, Covers, Doors	Safety
Openings	Work Environment
Accessibility	Illumination
Mounting and Fasteners	Training Requirements
Connectors	Failure Indication (location)
Installation	
Standardization	
Lubrication	
Fuses and Circuit Breakers	

TABLE 5-5.
MOST OFTEN MENTIONED MAINTAINABILITY DESIGN FACTORS
AND THEIR CARDINAL WEIGHTING

Times Mentioned	Maintainability Design Factor	Detect	Diagnose	Correct	Verify
16	Accessibility	1	3	4	3
14	Test Points	3	4	0	4
14	Controls	1	3	1	3
13	Labeling and Coding	3	4	4	4
12	Displays	4	4	0	4
12	Manuals, Checklists Charts, Aids	4	4	4	4
12	Test Equipment	4	4	0	4
12	Tools	0	1	4	1
12	Connectors	1	3	4	3
11	Cases, Covers, Doors	0	3	4	3
10	Mounting and Fasteners	0	1	4	1
10	Handles and Handling	0	1	4	1
10	Safety	1	4	4	3

WEIGHTING CODE

0 — not a factor
1 — not ordinarily important
3 — might be important
4 — necessity

tention should be focused on displays and test philosophy insofar as maintainability design factors are concerned. During system definition, this is expanded to include accessibility and test points, and so on, through all design stages.

2. *Matrix 2.* An examination of maintainability design factors vs maintenance tasks reveals, that the same factors are important in the diagnostic and verification periods. (This is essentially because they both are related to test and checkout.) In addition, the informational items (test points, labeling and coding, displays, manuals and checklists, and test equipment) are important in the detection period, while items related to access and mechanical actions are important in the correction period.

3. *Matrix 3.* A look at the relationships existing between design stages and maintenance levels shows clearly the shift in emphasis from system to lower levels as the design cycle moves from concept to detailed design. System and subsystem definition, and maintainability and allocation of tasks to lower levels should be well established by the end of the Preliminary Design Stage. This does not imply that one can forget about these levels from then on, since interface and total system integration considerations need to be constantly reviewed.

4. *Matrix 4.* A similar shift in emphasis occurs between maintenance level and maintenance tasks. Whereas detection is of extreme importance at the system level, diagnosis is very important at subsystem, equipment, and lower levels. Similarly, correction becomes important at equipment and lower levels. Verification shifts the emphasis up again to the higher levels to assure that the system has been successfully restored.

5. *Matrix 5.* This matrix shows that nearly all of the maintainability design factor considerations are of major significance in the intermediate levels from equipment through subassembly and many of them at the subsystem and system levels.

6. *Matrix 6.* The final matrix shows that all maintenance tasks are of fundamental importance at all life cycle design stages. This is because downtime is the fundamental measure of maintainability and maintainability is, in turn, a fundamental parameter of system design and system effectiveness.

In summary, as a first step in design, for example, Figs. 5-7 and 5-8 indicate that fault detection is concerned primarily with informational items such as test equipment, manuals and checklists, displays, test points, and labeling; this concern exists from system to unit levels.

The next step is to develop functional relationships among these parameters. This is a difficult task requiring empirical approaches. Very little has been done in this area, and it offers much room for continued research. Par. 5-7.2 illustrates how such relationships would be approached for fault detection as a function of maintainability design factors.

5-7.2 AN ILLUSTRATION OF MAINTAINABILITY DESIGN-FAULT DETECTION

To illustrate how the preceding matrices might be used in maintainability design, detection time will be examined in more detail. As indicated in par. 5-5.2, most maintainability specifications and reports exclude detection time from their definitions of Active Repair Time. However, fault detection can be a significant part of downtime, and since it is, to a large extent, design-controllable, attention should be given to maintainability design to minimize fault detection time.

A fault is assumed to exist upon the Occurrence of one of the following events, as defined in par. 5-5.1:

1. Degradation of performance but above minimum acceptable level
2. Critical failure
3. Catastrophic failure.

Now let us examine the elements that constitute the fault detection process:

1. Sensing the parameters which are subject to change and whose sensitivity to change is such that there is a reasonable probability that performance will be degraded below acceptable levels
2. Comparing the performance (or change in performance) of these parameters against established standards
3. Indicating the actual change in performance of substandard parameters or above/below acceptable level (GO/NO-GO) status
4. Causing a failure indication and/or alarm to be .registered.

This process is shown in Fig. 5-9 and is further illustrated in the Fault Detection Flow Chart, Fig. 5-10.

There are a number of ways in which fault detection may be implemented. The choice depends upon:

1. Criticality of the parameter, unit, or system involved with regard to mission success (availability or dependability requirement)

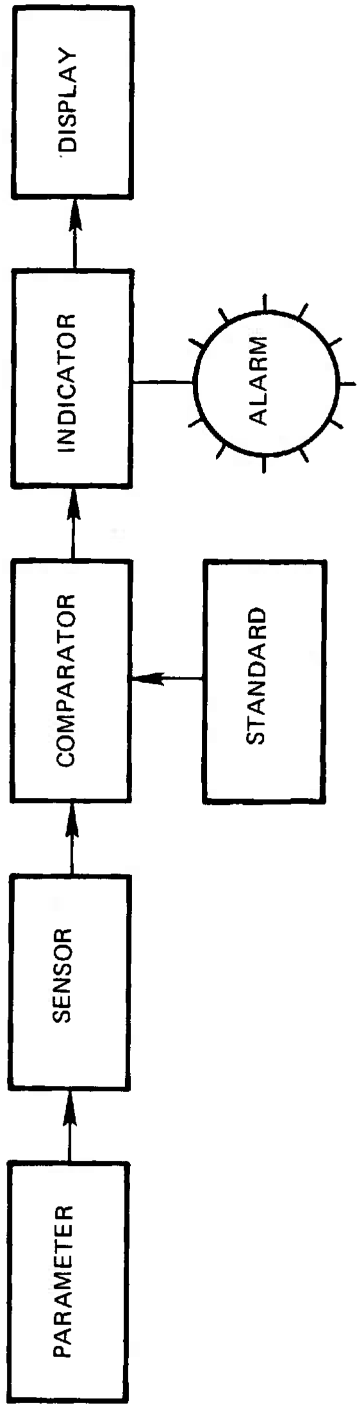


Figure 5-9. The Fault Detection Process

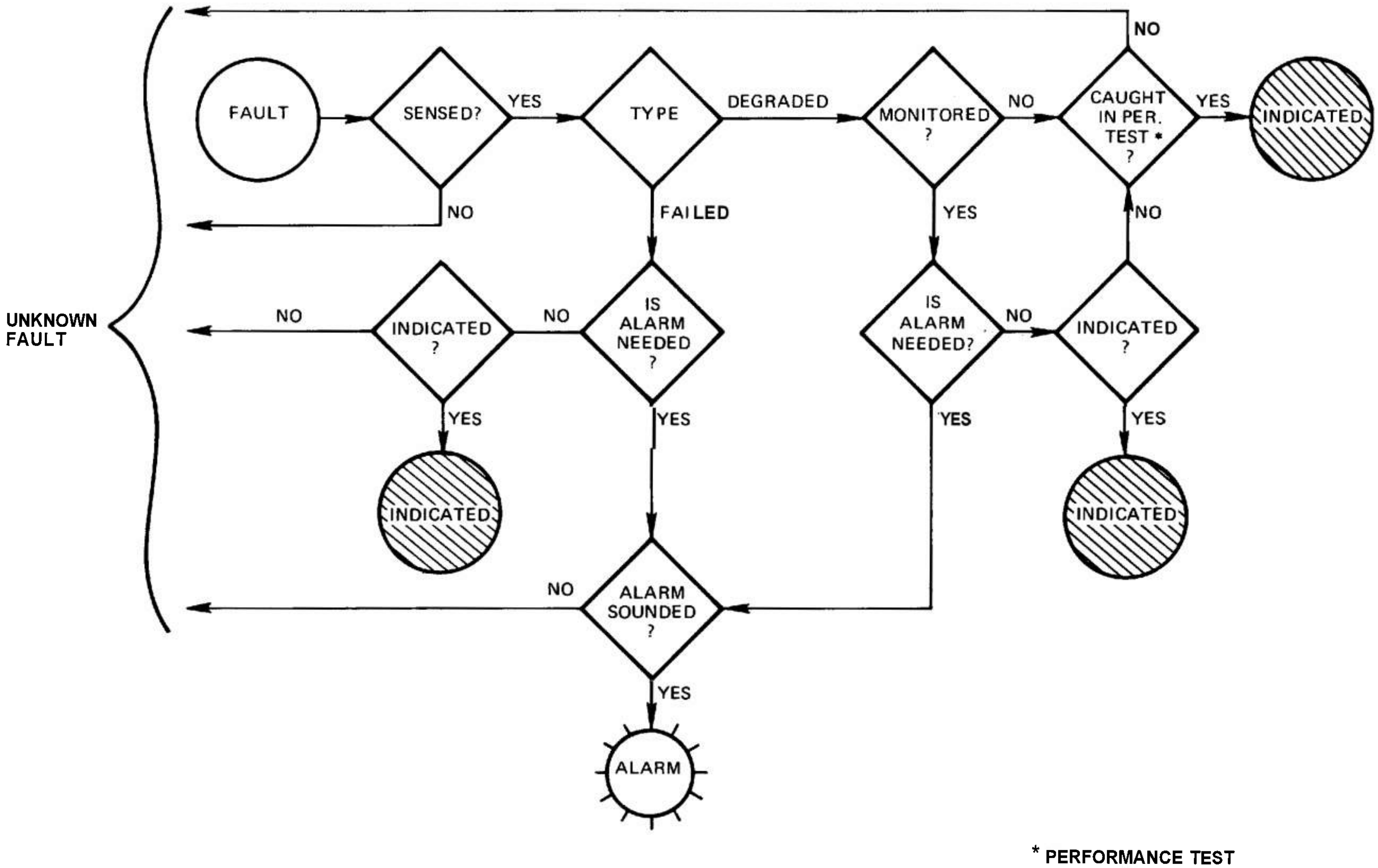


Figure 5-10. Flow Chart—Fault Detection Process

2. Failure rate of the parameter, unit, or system

3. Overall operational maintenance and test policy adopted

4. Cost-effectiveness of the proposed scheme.

Among the possible fault detection strategies are:

1. Automatic monitoring with alarm indication only

2. Automatic fault indication with alarm for critical parameters and with degradation level/failure indicators or recorders

3. Periodic test and inspection by maintenance personnel

4. Periodic self-test and calibration by the operator

5. Abnormal operations noted by the operator, based on his experience and knowledge of the equipment operational characteristics.

In Table 5-5, detection time is seen to be a function of a number of maintainability design factors. Expressing detection time t_{det} as a function of the more important factors (rating of 3 or 4), we obtain

$$t_{det} = F(TP, LC, D, MC, TE)$$

where

TP = Test Points
LC = Labeling and Coding
D = Displays
MC = Manuals, Checklists, Aids
TE = Test Equipment

The test equipment factor is one of the more significant of these. Depending upon the operational requirements, the nature of the equipment under consideration, its importance to mission effectiveness, the overall maintenance policy, and cost considerations, a decision may be made from one extreme of no specific provision for fault detection to the other extreme of fully automatic monitoring of all parameters of importance. The effect of these on detection time and their costs can be expected to vary somewhat, as shown in Fig. 5-11.

It is readily seen how the other design factors are affected by the test equipment factor. For example, Automatic Monitoring and Automatic Test Equipment imply a large number of built-in sensors and references. Built-in Test Equipment (BITE), General Purpose Test Equipment (GPTE), and Special Purpose Test Equipment (SPTE) imply fewer sensors but more external test points (for detection). Operator failure-sensing implies few or no test points. Similarly, automatic monitoring or testing imply simple GO/NO-GO display or alarm provisions, whereas general purpose checks and operator sensing and indication imply more interpretive types of display. Automatic features imply relatively little or no manual or checklist information requirements, and simple labeling and coding. The manual features require significant amounts of these items.

Automatic detection features require less accessibility. Periodic test, on the other hand, may require test point, control, and display accessibility.

5-7.3 MONITORING AS A FAULT DETECTION TECHNIQUE

There are several ways in which faults may be detected. One is by continuous monitoring of the sensitive system parameters. Another is by intermittent monitoring or sampling. A third is by periodic test. The method chosen should be determined by the system effectiveness requirements of each system, subsystem, or parameter; its sensitivity to change or degradation; operational demand requirements; and cost. *Monitoring* is defined here as the process of determining change in the parameter, function, or item under examination, based on its own state of operation in the system, and without the injection of external stimuli. *Test* is defined as the interruption of the normal operation and the injection of a standard test signal. This signal may be either a self-test signal injected automatically at periodic or random intervals or upon call, or an externally introduced stimulus.

Monitoring can be a valuable and effective means of fault detection and location and, to some degree, fault isolation. There has been a tendency in the past to monitor too many items—for example, such details as individual resistance values or current and voltage in electronic equipment, even where changes in these items have little effect on system performance.

The monitoring or detection of variation in nearly every component and parameter only encumbers the maintenance function, and causes a greater potential reduction in system maintainability and reliability than monitoring to a defined subsystem or functional level. This is largely due to the added complexity and to false alarms which put the system down. Only those system performance parameters that materially affect system performance should be monitored for fault indications. Sensitivity analysis can be used to determine which parameters these should be. A properly designed monitoring system based on a logical analysis of system requirements and maintenance policy can be an invaluable aid to minimizing system downtime.

To provide adequate system effectiveness, it is necessary to have a rapid and timely indication of system failure or degradation. There are several ways in which this may be accomplished. Gross system failure will

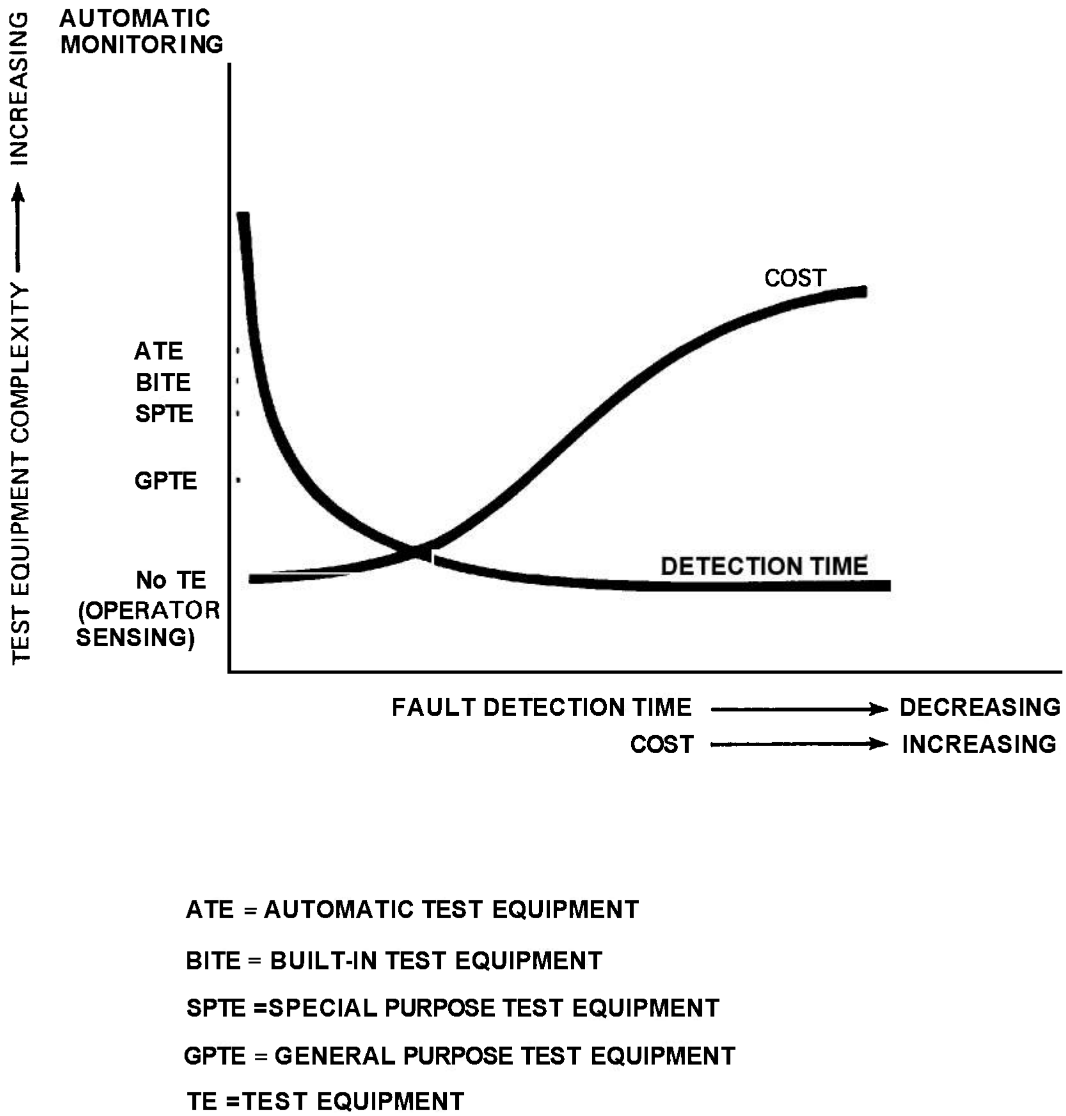


Figure 5-11. Test Equipment Complexity vs Detection Time and Cost

often be recognized by the operator if the system is in use. Degradation may sometimes be recognized by the operator even before a built-in monitor does so. It is necessary, however, to take other measures to minimize uncertainties or the risk of failure to detect an inoperable condition which is critical to mission success.

Monitoring or testing may be on a GO/NO-GO basis, by LO-GO-HI sectors, or by means of actual value indication. All of these may be quantitative in nature. The first of these, however, merely indicates that the parameter or function being monitored is in one of two binary states, i.e., above or below an acceptable level. The second method gives an indication of shift in parameter value or marginal warning while the third gives an actual numeric reading, as an absolute value or percentage. These latter two types of indication are of value if the trend is indicative of the degradation in performance so that preventive maintenance actions can be taken at appropriate times. The information available may also be used for reliability and maintainability analysis, and thus for improvement of system design.

Monitoring implies an "on-line" operation. What should be monitored depends upon the system effectiveness numeric; the *MTBF* of the various subsystems, equipments, and functional elements; and the sensitivity of the system to variations in these. Whether continuous or intermittent monitoring should be used is determined by the following factors:

1. Total number of monitoring points
2. Whether system availability and operational demand allow for test and preventive maintenance time periods
3. Whether monitor indications at the equipment or module location are satisfactory or whether a central monitoring (and perhaps recording) facility is feasible
4. Whether multiplexing is feasible.

Of the various equipments that make up a system, some will be performing vital functions at all times, some part of the time, while others will be performing largely auxiliary functions. The extent to which monitoring should be carried out is a function of such considerations.

5-8 MAINTAINABILITY DESIGN CHARACTERISTICS

The maintainability design characteristics of a system or equipment include those equipment features and design factors which will promote a decrease in downtime and an increase in availability. Factors to consider are:

1. Ease of maintenance
2. Minimization of preventive and corrective maintenance tasks to be performed
3. Minimization of the logistical burden by a reduced need for maintenance and support resources, such as personnel numbers and skill levels, support equipment, repair parts, and special maintenance facilities
4. Reduction in support costs.

Many of the principal factors which affect design for maintainability are listed in Table 5-4 and are related schematically in Fig. 5-12. The primary maintainability design characteristics have been ranked in Table 5-5.

Specific features and their effect on maintainability design are discussed under par. 5-8. These include the use of checklists, packaging, standardization, test and checkout, human factors considerations, safety, tradeoffs, and cost considerations. More detailed design guideline information is contained in Refs. 11 and 24-38. In particular, Refs. 11 and 31 are especially recommended.

5-8.1 CHECKLISTS

Checklists are an important and useful aid for system and equipment designers to insure that all essential design factors which influence the maintainability characteristics of the system/equipment have been given adequate consideration, in much the same way that Table 5-1 serves as a checklist to insure that maintenance policies and goals have been properly considered in developing the maintenance concept and maintenance support plan.

Maintainability checklists are counterparts to and should be included in maintainability design guides. Checklists may be used in three ways:

1. By maintainability and design engineers for considering the influence of specific maintainability design features (Ref. 39, Appendix A; Ref. 40)
2. By system engineers and project managers for design reviews during the various system life cycle phases, particularly in the design stages (Fig. 3-9) (Refs. 11, 21, 41).
3. By maintainability engineers for maintainability prediction (Refs. 19 and 39, Appendix B).

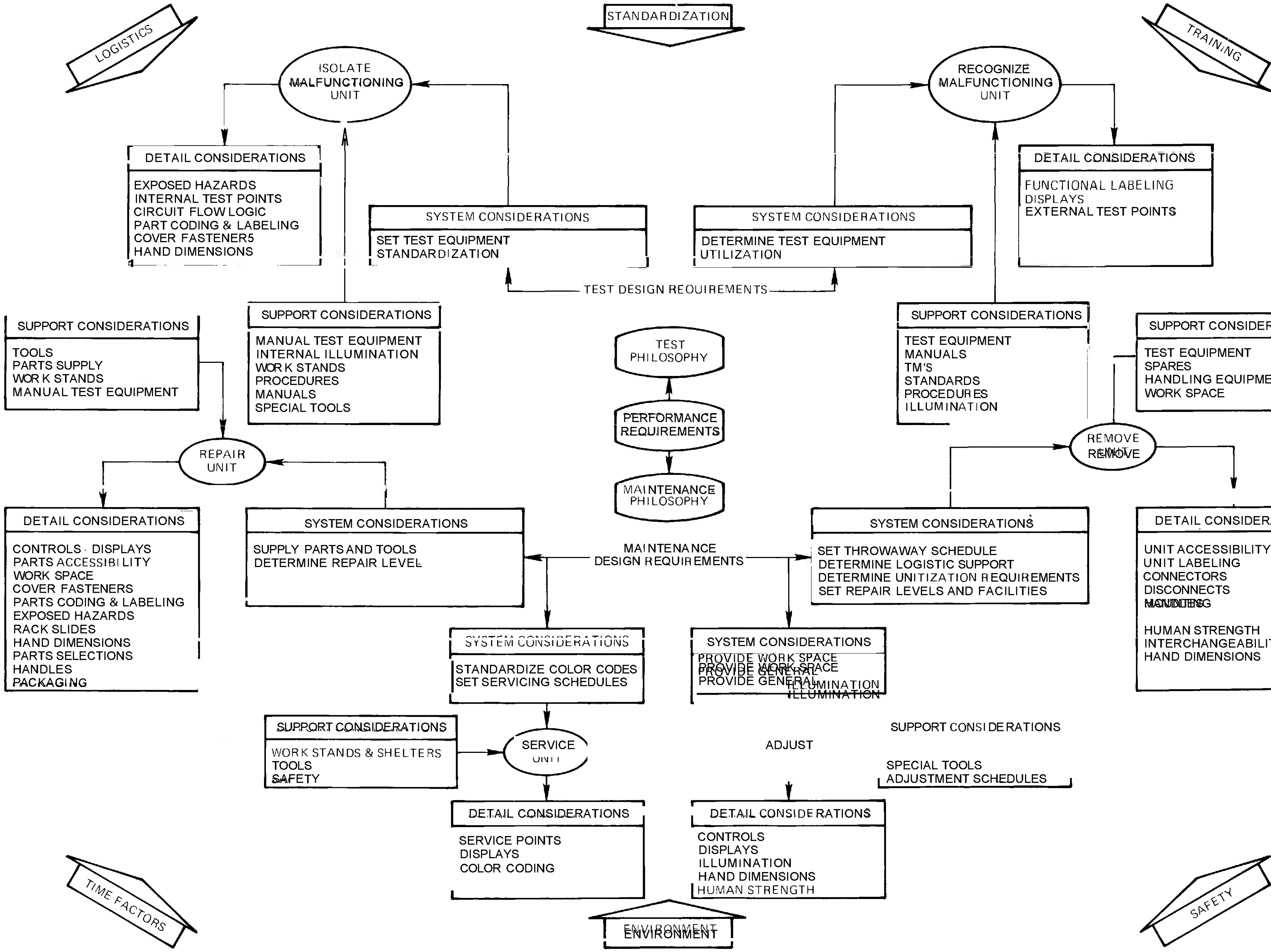


Figure 5-12. Maintainability Design Considerations

5-8.1.1 Checklist Information Pertaining to Maintenance Downtimes

Checklist information with respect to maintenance downtimes includes those items which will help reduce preventive maintenance times (servicing and inspection) and the corrective maintenance downtime components (detection, diagnosis, correction, and verification). One way to reduce downtime is to improve the reliability of the system and its components, with regard to both failure characteristics and wearout life. Reliability and maintainability features must be carefully examined for cost-effectiveness trade-off. This should be of concern to design engineers as well as to reliability and maintainability engineers.

In addition to examining the trade-offs between reliability and maintainability, maintainability and design engineers must seek means for minimizing preventive and corrective maintenance tasks. These include:

1. Reducing the frequency of scheduled maintenance actions
2. Simplifying or eliminating the preventive maintenance tasks
3. Increasing the time between scheduled overhauls
4. Reducing each of the corrective maintenance downtime elements
5. Reducing the requirements for highly trained specialists
6. Maximizing the amount of standardization and interchangeability of components and modules.

Table 5-6 is a checklist for use in minimizing maintenance downtimes.

5-8.1.2 Checklists for Consideration of Maintainability Design Factors

Engineers concerned with the maintainability characteristics of systems and equipment generally agree on the importance of certain design factors (Table 5-5). Many of these are interdependent. For example, accessibility includes considerations of safety, cases, covers, doors, handles, mounting, fasteners, connectors, locations of test points, and human factors. In order to assure that proper consideration is given to these factors, it is essential that conscious attention be given to them during all system planning and design phases. Design reviews (par. 10-6) are one means of assuring such consideration. Checklists help in both design and review situations.

Checklists and design guidelines applicable to the factors listed in Table 5-5 are given in AMCP 706-134 (Ref. 11).

5-8.2 PACKAGING

The manner in which the equipment is packaged* is a dominant factor in its maintainability. The layout of parts, components, and assemblies, their mounting, access and ease of removal or repair all contribute significantly to ease of maintenance and maintenance downtime, and should be designed to facilitate the required or expected maintenance operations. The majority of these items can be located and packaged in a variety of ways and places. Among the factors which should be considered are:

1. Accessibility preferences and requirements
2. Modularity or unitization requirements
3. Standardization requirements
4. Reliability factors
5. Operating stress, vibration, temperature, and other environmental considerations
6. Producibility and other manufacturing requirements
7. Requirements for built-in test and test points
8. Characteristics peculiar to each item, such as
   a. Size, weight, and clearance
   b. Fragility and sensitivity with resultant protection needs
   c. Servicing, adjusting, or repair needs and procedures
   d. Clearance requirements for removing and replacing each item
   e. Tool and test equipment access
   f. Specific phenomena such as critical lead length, weight balance, and heat dissipation.
9. Safety.

5-8.2.1 Accessibility

Accessibility is one of the prime considerations with regard to both maintainability and equipment packaging. In 1966 it was stated that "gaining access to equip ment is probably second only to fault isolation as a time-consuming maintenance activity, and when automatic fault-isolation equipment becomes available, it unquestionably will be first" (Ref. 42). It can probably be stated today, with the advent of automatic fault-finding a reality in many systems, that accessibility is indeed the number one problem (Table 5-5).

Accessibility can be defined as the relative ease with which an assembly or component can be approached for inspection, repair, replacement, or servicing. Ineffective maintenance is often the result of inaccessibility.

*Not to be confused with packaging for shipment.

TABLE 5-6.
MAINTENANCE DOWNTIME CHECKLIST

1. Have servicing and inspection intervals been maximized or otherwise chosen to assure maximum material readiness consistent with mission and operational requirements?

2. Have adequate provisions been made for inspection – such as appropriate access doors, inspection windows or ports, test points and displays, inspection instructions located next to inspection points, proper working height for the operator or technician, adequate lighting, and safety?

3. Have provisions been made to facilitate rapid fault detection?

4. Are mission critical performance parameters automatically monitored?

5. Do performance monitoring features include means for degradation measurement and failure prediction?

6. Where automatic monitoring is not feasible, are provisions for periodic checkout provided?

7. Can significant failures be readily localized to the affected equipment, assembly, or unit?

8. Are provisions for automatic fault location feasible?

9. Have features been provided to isolate the failed item to the replaceable or repairable module(s) or part?

10. Are indicators or alarms provided and located in such places and manner as to assist the maintenance technician to locate and isolate the failed item promptly?

11. Have provisions been made for rapid and ready access to failed items?

12. Are the failed items readily replaceable or repairable?

13. Has a replace/repair/discard policy been established?

14. Have adequate spares and repair parts been provided and located so as to facilitate interchange time?

15. Have adequate controls, displays, adjustments, test points, and checkout procedures been provided to facilitate alignment, calibration, and checkout of the unit after repair has been made?

16. Have built-in test features been provided to facilitate verification of corrective maintenance actions?

The technician will tend to delay or omit maintenance actions, make mistakes, and accidentally damage equipment if he cannot adequately see, reach, and manipulate the items on which he must work. Poor accessibility to routine service and inspection points and parts of equipment reduces the efficiency and increases the time of the maintenance operation. If it is necessary to dismantle a given component completely or partially to reach a given part, the availability of the equipment decreases and maintenance costs increase. Controls, check points, inspection windows, lubrication, and pneumatic and hydraulic service points are built into the equipment so that it can be kept operating at peak performance. If these service points are inaccessible, routine maintenance becomes difficult.

Accessibility, however, when considered separately, does not constitute maintainability. The mere fact that a technician can "get at something" does not mean that he can maintain it. Accessibility requirements are determined by the necessary maintenance action, which may be visual, physical, or both, depending on whether the task is inspection, servicing, adjusting, repairing, or replacing. Generally, they represent two needs: access to an item for inspection and testing, and space in which to adjust, repair, or replace it.

Well designed equipment accesses are essential for ease of maintenance, and should be provided whenever a maintenance procedure would otherwise require removing a case or covering, opening a fitting, or dismantling a unit. Before designing equipment accesses, the engineer should list the parts of the equipment that have to be reached, their failure rate, and the operations that are likely to have to be performed on each part. The access then should be designed to make those operations as convenient as possible. Table 5-7 gives recommended equipment accesses.

Factors affecting accessibility include:

1. Operational location, setting, and environment of the unit

2. Frequency with which the access must be entered

3. Maintenance functions to be performed through the access

4. Time requirements for the performance of these functions

5. Types of tools and accessories required by these functions

6. Work clearances required for performance of these functions

7. Type of clothing likely to be worn by the technician

8. Distance to which the technician must reach within the access

9. Visual requirements of the technician in performing the task

10. Packaging of items and elements behind the access

11. Mounting of items, units, and elements, behind the access

12. Hazards involved in or related to use of the access

13. Size, shape, weight, and clearance requirements of logical combinations of human appendages, tools, units, etc., that must enter the access (Ref. 11).

Once access has been gained to an area in which an assembly or part is to be repaired or replaced, access to that particular item must be provided. Guidelines for the designer in planning for ease of maintenance include:

1. Locate each unit in the equipment in such a way that no other unit or equipment has to be removed to get to the unit.

2. Locate assemblies and parts so that structural items and other parts do not block access to them.

3. If it is necessary to put one unit behind another, place the unit requiring less frequent attention in back of the one requiring more frequent attention.

4. Do not locate a unit in a recess, or behind or under structural members, floor boards, operator's seats, hoses, pipes, or other parts of the equipment that are difficult to remove unless this serves some purpose, such as protecting the unit.

5. Removing any line replaceable unit (LRU) should require the technician to open only one access.

6. Units generally should be designed for removal through the front rather than through the side or back of the equipment.

7. Units should be removable from the installation along a straight or moderately curved line; they should not have to be juggled around comers.

8. Place assemblies and parts so that sufficient room is available for the use of test probes and other tools needed.

9. Place throw-away items so that they can be removed without the necessity of removing other items.

10. Design each assembly so that it need not be removed in order to troubleshoot any of its components.

11. Use plug-in modules wherever economically feasible.

Many more specific guidelines have been published for particular types of equipment, as listed in the references to this chapter. In particular, AMCP 706-134, ***Maintainability Guidefor Design,*** (Ref. 11), contains design guidelines and checklists applicable to the design of Army equipment.

5-8.2.2 Logical Flow Packaging (Functional Modularization)

Logical flow packaging or functional modularization is a packaging method in which a conscious effort is made by design engineers to locate and package components and subassemblies in self-contained functional units in order to facilitate both the operation and maintenance of a system in accordance with some functional relationship. Although broad in its applications, functional modularization is specific in its use by maintainability engineers as a design factor for complex systems. The paragraphs that follow are taken from Ref. 11.

Modularization refers to the separation of equipment into physically and functionally distinct units to facilitate removal and replacement. It denotes any effort to design, package, and manufacture a group of parts and elements in an aggregate which can be considered as an undivided whole. Modularization enables systems, assemblies, and subassemblies to be designed as removable entities.

The modular concept covers the range of complete black-box equipment built on a single structure to the smallest removable subassembly. The significance of modular construction lies in its degree of use. For example, a module may consist of nothing more than a single operating circuit in a system, i.e., the system reduced to the smallest operating function possible, or it may consist of modules built on modules to form the overall equipment function. The degree to which the concept is applied depends on the particular application of the equipment and its practicality and cost.

Modular construction should be incorporated or designed into the product whenever practical, logistically feasible and combat suitable, or where elimination or reduction of personnel training and other similar advantages will result.

The concept of modularization creates divisible configuration, which is more easily maintained. Troubleshooting and repair of unitized assemblies therefore can be performed more rapidly. Utilization of these techniques to the fullest extent improves accessibility, makes possible a high degree of standardization, provides a workable base for simplification, and provides the best approach to maintainability at all maintenance levels.

**TABLE 5-7.
RECOMMENDED EQUIPMENT ACCESSES**

Desirability	For Physical Access	For Visual Inspection Only	For Test and Service Equipment
Most desirable	Pullout shelves or drawers.	Opening with no cover.	Opening with no cover.
Desirable	Hinged door (if dirt, moisture or other foreign materials must be kept out).	Plastic window (if dirt, moisture or other foreign materials must be kept out).	Spring-loaded sliding cap (if dirt, moisture or other foreign materials must be kept out).
Less desirable	Removable panel with captive, quick-opening fasteners (if there is not enough room for hinged door).	Break-resistant glass (if plastic will not stand up under physical wear or contact with solvents).	
Least desirable	Removal panel with smallest number of largest screws that will meet requirements (if needed for stress, pressure, or safety reasons).	Cover plate with smallest number of largest screws that will meet requirements (if needed for stress, pressure or safety reasons).	Cover plate with smallest number of largest screws that will meet requirements (if needed for stress, pressure, or safety reasons).

Another important advantage of unitized or modular construction from a maintenance viewpoint is the division of maintenance responsibility. Modular replacement can be accomplished in the field with relatively low skill levels and few tools. This accomplishes a prime objective of maintainability—reduce downtime to a minimum. Defective modules can be discarded (if nonmaintainable), salvaged, or sent to a higher maintenance level for repair.

Modular design cannot be applied to all types of equipment with equal advantage. Its greatest application has been in electronic equipment. It has application in complex equipment of other types, but becomes increasingly difficult to exploit in simpler devices. Following are a few additional advantages of modular construction:

1. New equipment design can be simplified and design time shortened by use of previously developed standard "building blocks".

2. Current equipment can be modified with newer and better functional units that replace older assemblies of component parts.

3. The standard "building blocks" can be manufactured by fully automated methods.

4. Maintenance responsibilities can be divided among the various maintenance levels best equipped to fulfill them.

5. The recognition, isolation, and replacement of faulty units is facilitated, permitting rapid maintenance at the user level, with consequent reduction of downtime.

6. Training of user maintenance personnel will take less time and cost less (Ref. 11).

7. The use of automatic and semi-automatic diagnostic techniques is facilitated by functional packaging, inasmuch as modularization allows for the ready prediction of such faults as occur in a system. This is made possible by the extent to which packaging permits the employment of programmable test sequences with a highly developed capability for isolating faults. Once a fault has been quickly located in a small unit, standard fault-isolation procedures developed for that unit can be used for repairing the item, or more advantageously, modular replacement can be made, thereby reducing system downtime (Refs. 11 and 42).

Logical flow packaging is based on the following:

1. Circuits, parts, and components are packaged in an arrangement parallel to their functional relationships as established by logic diagrams.

2. Methods and subassemblies are selected so that, insofar as possible, only single input and output checks are necessary to isolate a fault within an item.

3. Clear indication is given of the unidirectional signal flow within a given piece of equipment.

In order to make use of logical flow packaging, one must construct functional block diagrams which relate the logical flow of information, signals, or energy in the equipment or assembly, or use timing logic, test logic, or maintenance logic diagrams. The half-split technique (Ref. 43) or Design Disclosure Format (Refs. 23 and 24) are two of the many techniques useful for logical flow packaging, both of which are maintenance related.

The advent in recent years of solid state devices requiring less power and thus reducing heat dissipation needs, along with microelectronics and cordwood packaging techniques, has permitted packaging densities several orders of magnitude higher than heretofore possible. These advances now allow multifunction module packaging and, together with advances in test techniques, permit testing of a replaceable module to be accomplished automatically.

As contrasted with logical flow packaging, standard packaging methods have no clear-cut procedures. Rather, the final product is packaged by balancing a number of factors such as heat loss, component size, unit size and weight, and design and manufacturing convenience rather than ease of maintenance. The logical flow method is superior in minimizing downtimes, reducing the requirements for high skill levels and in optimizing the amount of information gained per unit of test time.

Because of the rapid advances in packaging methods just described, methods of mounting parts, components, and subassemblies have changed markedly in the past ten years. In addition to higher density packaging, advances have been made in materials, cooling methods, fastening methods, and in construction techniques. It would be impractical to discuss these in detail. Parts Four and Five of AMCP 706-134, along with other publications, discuss such methods (Ref. 11).

5-8.3 STANDARDIZATION AND INTERCHANGEABILITY

Standardization is a design feature for restricting to a minimum the variety of parts and components that will meet the majority of the requirements of a system. It denotes any effort to select, design, or manufacture parts, components, assemblies, and equipment, or associated tools, service materials, or procedures, so they are either identical to or physically or functionally interchangeable with other parts—the so-called "form, fit, and function" criterion.

Standardization should be a primary goal whenever the design configuration of a system or equipment is considered. Standardization significantly reduces both the acquisition and the support costs of a system over its life cycle, as well as resulting in increased maintainability and reliability.

Standardization may occur at many levels. In addition to part, component, and assembly standardization, it should be applied to types and models within the Army and should also be applied insofar as feasible across product lines and AMC Commodity Commands (Intra-Service) and across the services DoD-wide (Inter-Service). There exist both Defense Standardization and Federal Standardization Programs which should be closely adhered to (Ref. 11).

The scope of the ***Defense Standardization Program*** includes the standardization of materials, components, equipment, fasteners, and processes as well as the standardization of engineering practices and procedures essential to the design, procurement, production, inspection, application, preservation, and preparation for delivery of items of military supply.

The congressional mandate to standardize the Federal Supply System applies to all areas where specific benefits can be anticipated. A vigorous standardization program is of mutual concern to both industry and the Government. Eliminating and/or preventing excessive item variations results in economies in tooling, engineering, manpower, and in the size of both Governmental and industrial inventories.

The primary goals of standardization are to:

1. Reduce the number of different models and makes of equipment in use.

2. Maximize the use of common parts in different equipment.

3. Minimize the number of different types of parts, assemblies, etc.

4. Use only a few basic types and varieties of parts, etc. to ensure that those parts are readily distinguishable, compatible with existing practices, and used consistently for given applications.

5. Control, simplify, and reduce part coding, numbering practices, and storage problems.

6. Maximize the use of standard off-the-shelf items and components.

7. Maximize the use of interchangeable parts.

Standardization, however, is not intended to inhibit design improvement effects. Before improvement efforts are undertaken, it should be established that the value of design improvement outweighs the advantages of standardization. Rather than being a matter of initiative or freedom, the lack of standardization seems largely attributable to poor communication among designers, contractors, users, buying agencies, subcontractors, and their divisions and agencies. It is suggested that the maintainability effort concern itself with this lack of communication and assume responsibility for ensuring and coordinating compatibility and uniformity in design (Ref. 31).

Standardization must be applied at all stages of design, as well as to items already in the supply system. Wherever practical, it is required that standard parts, components, and subassemblies be used, Standardization decreases the number of unique component items and design prerogatives in system development and production.

While standardization is highly desirable for maintainability, it must be realized that standardization cannot be permitted to interfere with technical advances. Consequently, standardization is a continuous process rather than a static condition.

A key factor in reducing the overall and long range costs of logistical support is to design so as to standardize for both physical and functional ***interchangeability.*** Due consideration to standardization during the development of a new system will provide for rapid and easy interchange and replacement of parts and subsystems under all conditions. This is the ultimate result of effective standardization. Both Government and industry should see that their efforts are coordinated toward this achievement (Ref. 11).

When standardization is carried to the practical maximum in system design, certain major advantages are gained by the support activities required for the completed system as follows:

1. Both the types and the quantities of spares normally are reduced because of the increased system reliability obtained by design. This, obviously, reduces overall support costs.

2. Training requirements for support personnel are reduced, principally by the simplification of circuits and functions resulting from the application of standardization design principles; moreover, the number and types of support personnel required are also reduced.

3. In the same way, requirements for technical publications are greatly reduced in quantity, as well as in the amount of detail to be covered.

4. The varieties and quantities of test equipment required to support a system are reduced.

5. In general, standardization design reduces the need for support facilities of all kinds (Ref. 42).

Interchangeability, as a maintainability design factor, is closely related to standardization, in that it is through standardization that interchangeability is realized. As defined by maintainability engineers, interchangeability is a design policy whereby any given part or unit, so specified, can be substituted in an assembly or system for any like part or unit in accordance with the principles of standardization. Functional interchangeability is attained when a part or unit, regardless of its physical specifications, can perform the specific functions of another part or unit. Physical interchangeability exists when any two or more parts or units made to the same specifications can be mounted, connected, and used effectively in the same position in an assembly or system.

In order to attain maximum interchangeability of parts and units in a given system, design engineers must insure:

1. That functional interchangeability exists wherever physical interchangeability is a design characteristic.

2. That physical interchangeability does not exist wherever functional interchangeability is not intended.

3. That wherever complete (functional and physical) interchangeability is impracticable, the parts and units are designed for functional interchangeability, and adapters are provided to make possible physical interchangeability wherever practicable.

4. That sufficient information is provided in job instructions and on identification plates to enable a user to decide definitely whether or not two similar parts or units are actually interchangeable.

5. That differences are avoided in the size, shape, and mounting, and in other physical characteristics.

6. That' modifications of parts and units do not change the ways of mounting, connecting, and otherwise incorporating them in an assembly or system.

7. That complete interchangeability is provided for all parts and units that are intended to be identical, are identified as being interchangeable, have the same manufacturer's number or other identification, and have the same function in different applications. This is especially important for parts and units whose failure rates are high (Ref. 42).

Interchangeability requirements should be determined from consideration of field conditions as well as from that of economy of manufacture and inspection. Liberal tolerances are essential for interchangeability. Tight tolerances do not themselves increase quality or reliability; on the contrary, unnecessarily close requirements may increase manufacturing costs without tangible gains in accuracy. Specifying tolerances closer than required is uneconomical in cost and time. Tolerances should be assigned to component features for position, concentricity, symmetry, alignment, squareness, and parallelism when the control of these factors is important for correct functioning or correct assembly. Tolerances assigned to components should be reviewed carefully, however, to prevent unnecessary difficulties in production or inspection from being imposed without real functional or assembly necessity.

Insofar as is possible and practical—and where interchangeability design considerations do not degrade equipment performance, increase cost, or reduce inherent maintainability or reliability-equipment should be designed with the minimum number of sizes, types, assemblies, subassemblies, and parts possibly requiring replacement. Like assemblies, subassemblies and replaceable parts should be according to MIL, AN or MS standards where possible and should be electrically, mechanically, hydraulically or otherwise interchangeable, both physically and functionally, regardless of manufacturer or supplier (Ref. 11).

The advantages gained from effective interchangeability are essentially the same as those gained by standardization. In addition, the provision of interchangeability is essential to effective standardization.

5-8.4 HUMAN FACTORS CONSIDERATIONS

One of the most important aspects to be considered in equipment design—regardless of the configuration, size, operation, or application of the item—is that it must be capable of being operated and maintained by man, the variable factor upon which human factors engineering is based. The system engineering concept applies not only to equipment but also to the human beings who operate and maintain the equipment (see Fig. 1-1). People are used or involved in every equipment system, because equipment systems are always built for some human purpose; they exist to serve some human need. Men decide when and how to use machines; men feed inputs to and base their actions on outputs from machines. Machines work well only if the men operating and maintaining them can and do perform their jobs satisfactorily. The system engineering concept, therefore, must be of a man-machine system.

A man-machine system is any system in which men and machine interact in performing a function. The system might be a large aggregate, such as an Army mobile force composed of men and combat vehicles, or it might consist of a single man and a single machine, such as a radio operator and a radio. It follows, therefore, that human factors engineering may be defined as the application of data and principles about human performance to the planning, design, and development of components, equipments, and systems.

The basic objectives are to improve and maximize the field performance and reliability of man-machine systems, particularly with respect to human factors. These include problems involving speed and accuracy of operation, operational reliability and maintainability, minimization of operator training and skill requirements, safety, and operation under stress (Ref. **45**).

The Army vehicle is a good example of a man-machine system in which the operator plays a commanding role or actively intervenes in the system from time to time. The man reacts to inputs from the speedometer and other displays, inputs from the road and outside environment, noise from the engine, feedback to his muscles from the steering wheel, and other stimuli. From these inputs he makes decisions to perform certain control movements. These movements affect the machine, which in turn furnishes new and different inputs to the driver.

We consider such a man-machine interaction as a closed-loop system because it calls for continuous interaction between the man and machine (Fig. **5-13**). An open-loop system is one in which the interaction between man and machine is intermittent rather than continuous. For example, a communication system in which the talker gets no feedback as to whether the message has been received would be considered an open-loop system (Ref. **43**).

Systems are designed and built by people. There are no self-maintained systems. Systems do not replace their own burned-out vacuum tubes, transistors, light bulbs, or failed modules, or solder their own connections. People do all these things. For these reasons, one could argue that all equipment systems are man-machine systems. Nonetheless, systems vary enormously in the degree to which they. involve human operators in any active sense. The system of traffic lights that regulates the traffic flow of any large city operates independently of human operators. Once the lights and regulating mechanism are installed, the lights go on and off automatically. In systems of this type, the role of the human being is largely that of designer, builder, and maintenance man (Ref. **46**).

Human factors are primary considerations in stating requirements, in developing hardware to meet these requirements, and in testing the acceptability and suitability of the item to meet operational and environmental conditions. In the preparation of a statement of requirements or military characteristics, an overriding consideration is that the equipment must be controlled by, operated by, and maintained by men. In the development and test of the equipment, the objective is to determine the functional suitability of the man-machine combination (Ref. **47**).

Designing and developing the various machine components of a system require the abilities of many different types of engineers. Just as one engineer is made solely responsible for power requirements, another for aerodynamic properties, etc., it is mandatory that an engineer be made responsible for human factors. Problems generated in the absence of a human-factors representative on the design team are usually discovered very late in the system development and are disproportionately expensive to solve by hardware redesign. If not solved in terms of hardware redesign, they might have to be solved even more expensively by a selection and training program devised to identify men who might be capable of fitting into the system after prolonged training. Thus, if a man-machine system is to perform at its best—for no better reason than that of economy-design must start with, and revolve around, the human components and their capabilities (Ref. **43**).

The designer must have the user (operator) in mind when he designs an equipment. He should be able to describe exactly what the operator has to do in operating (or maintaining) the equipment. Too often this task of writing down the job or task has been left to a human factors specialist. The designer should learn to do this himself, if for no other reason than that it forces him to anticipate difficulties he may have been creating for the user. Both designers and human factors specialists are also concerned with engineering questions, such as, "Should a function be performed manually by an operator, or should it be made automatic?" This question cannot be answered with a simple statement of "yes" or "no". There are, however, certain factors which may be considered in arriving at a fairly sound decision (Ref. **48**).

Human engineering is concerned with the following:

1. Man and his characteristics and capabilities
2. Man and his environment
3. Man as a system component
4. The man-machine interface.

The first category includes anthropometric (body measurement) data, man's sensory capabilities, his psychological makeup, his information processing capability, and his adaptability by means of learning. The second category includes the impact of the environment on man's capability to think and act and to perform certain tasks, including the effects of the physical, physiological, and psychological environment.

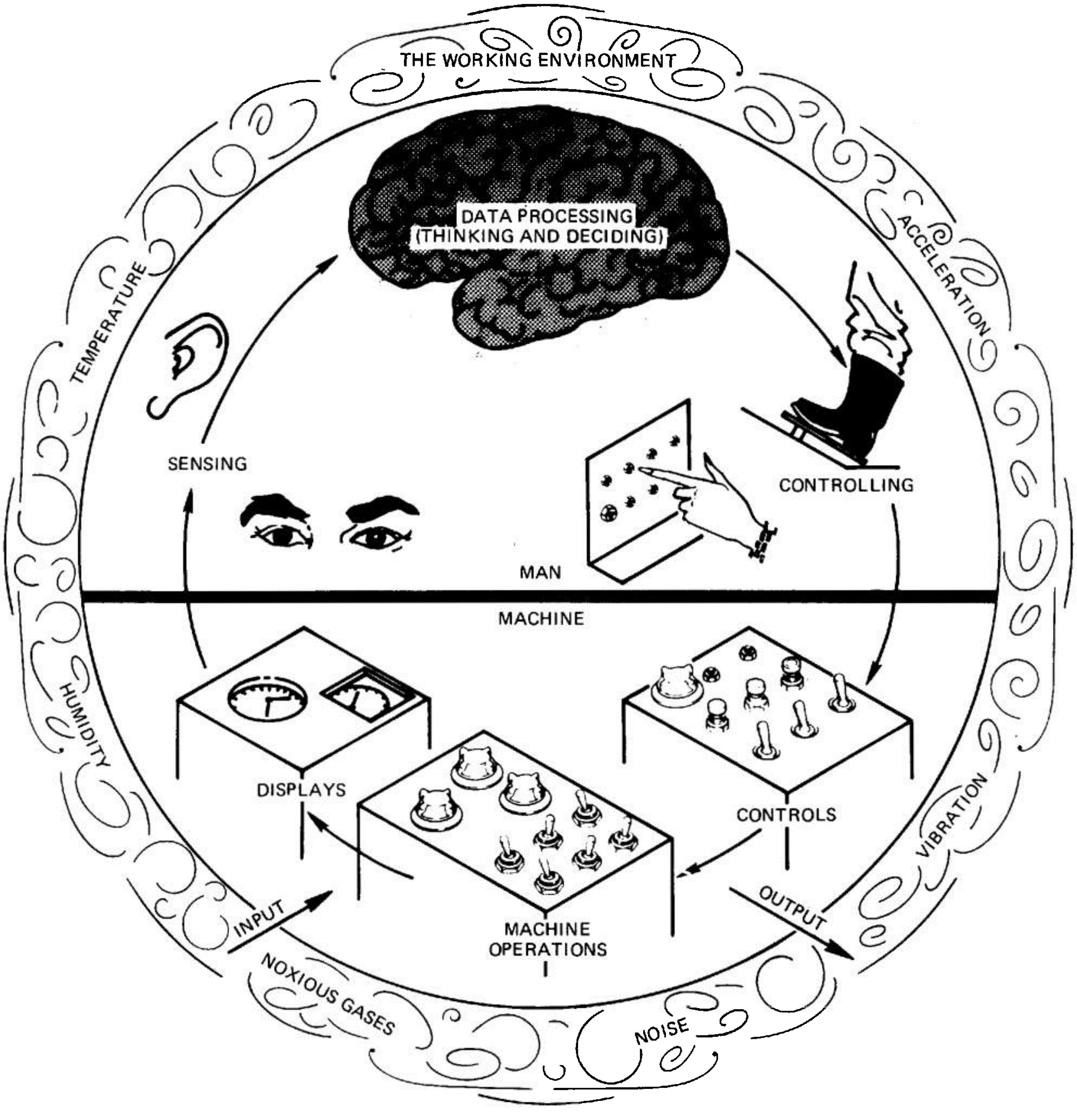


Figure 5-13. The Man-Machine System
Adapted from *Man-Machine Engineering,* by A. Chaparis. Copyright 1965 by Wadsworth Publishing Company, Inc. Reprinted by permission of the publisher, Brooks/Cole Publishing Company, Monterey, California.

The third category treats man as one component of the man-machine system. In effect, man is the sensor, data processor, and controller in the man-machine system (Fig. **5-13**). Finally, in the fourth category, the designer must consider those elements and characteristics which must be designed into the hardware/software portions of the system (e.g., displays, controls, sensors, test points, operating and maintenance information) and which will optimize the man-machine combination — the man-machine interfaces.

When the man and machine are considered in this fashion, it immediately becomes obvious that, to design the machine component properly, the capabilities and limitations of the man and his role in the system must be fully taken into account. Such consideration of man's abilities is the only way of achieving insight into the best ways in which he can be used as a component (Ref. **43**).

It is the purpose of this portion of the handbook to discuss these vital human factors considerations and their impact on design for maintainability. Detailed human engineering considerations are given in the many excellent textbooks and design guides referenced and will not be repeated here. In particular, Refs. **43, 46, 48,** and **49** are recommended as basic treatises on human engineering. Refs. 11, **24, 26, 30, 31, 32, 38,** 43, **48, 49,** and **50** contain valuable detailed design guideline information.

5-8.4.1 Man and His Characteristics and Capabilities

5-8.4.1.1 Human Body Measurement (Anthropometry)

The human body, its structure and mechanical function, occupies a central place in man-machine design. Failure to provide a few inches, which might be critical for the operator, can jeopardize the performance and safety of both man and machine. With proper forethought, these critical inches usually can be provided without compromising the design (Ref. **43**).

One important consideration, therefore, in designing for maintainability is information on body measurements. This information is required in the earliest design stages to ensure that equipment will accommodate operators and maintenance men of various sizes and shapes. Anthropometry is concerned with human-body measurement. Such measurement normally includes body dimensions, range of motion of body members, and muscle strength. While most of us are familiar with body dimensions that the tailor takes in altering or tailoring suits, the anthropometrists usually measure other body dimensions, as well as ranges of motion and strength. Certain of these are particularly 'pertinent for the design of seating arrangements, workspaces, controls and displays, sizes of access openings, and sizes and weights of units which can be lifted or carried by one man, etc.

Anthropometric data are usually presented in upper and lower percentiles, ranges, and medians (or means). With information of this type, the designer, who usually will not be able to accommodate all possible sizes, can decide where to make the cutoff. He must, of course, design equipment so that all members of the population for which it is designed can operate and maintain it; but at the same time, he might have to inflict less efficient or less comfortable circumstances on a small percentage of the population, i.e., those individuals having extreme measurements.

MIL-STD-**1472,** Human ***Engineering Design Criteria*** (Ref. **50**), states "Design shall insure operability and maintainability by at least 90 percent of the user population. The design range shall include at least the 5th and 95th percentiles for design-critical body dimensions." It further states that the use of anthropometric data shall take the following into consideration:

1. Nature, frequency, and difficulty of the related tasks
2. Position of the body during performance of these tasks
3. Mobility or flexibility requirements imposed by the tasks
4. Increments in the design-critical dimensions imposed by the need to compensate for obstacles, projections, etc.
5. Increments in the design-critical dimensions imposed by protective garments, packages, lines, padding, etc.

Tables of 5th and 90th percentile body measurements including static, dynamic, range of motion, and weight and strength data are included in Refs. 11, 43, **48,** and **50.**

5-8.4.1.2 Man's Sensory Capability and Psychological Makeup

Man, as part of the man-machine system, contains many useful sensors. In addition to the five major senses of sight, hearing, taste, smell, and touch, man can also sense temperature, position, rotation and linear motion, pressure, vibration, and acceleration (shock). Because man contains an information processing system and control system which is particularly sensitive to small changes in these sensations over a wide range, he can automatically recognize and react to such changes. Thus, man is often the best detector of changes in performance or other system conditions.

Men and machines have different capabilities and limitations. Although they sometimes can do the same thing equally well, more often one is better than the other. Men can do some things better than machines, and machines can do some things men cannot. Such differences in capability must be considered in detail when designing systems—they are important in deciding which jobs to assign to men and which to assign to machines. Differences in capability also often determine how a machine should be designed to be used most effectively by a human operator (Ref. 43). Refs. 11, 43, and 48 contain details of man's sensory capabilities.

5-8.4.1.3 Man as an Information Processor

Man has certain advantages and disadvantages as an information processor as compared with machines. The rapid advances in machine information processing capabilities (computers) in recent years have narrowed the differences to some extent. For example, the use of adaptive control and pattern recognition techniques have allowed some computers to have an acquired learning capability.

Man has good long-term memory for generalized experience, but rather poor immediate memory for most sensory functions. This is especially so in audition. His access time is slow, compared with that of a computer, but he is able to recall generalized patterns of previous experience to solve immediate problems. As yet, no computer can do this. Man learns to do numerical computations, but in the main, his time constants are such that he is a relatively poor numerical computer when under stress. No computer can match him, however, for the more qualitative, nonnumerical computations (Ref. 48). Ref. 48 contains data concerning man's information processing capability.

5-8.4.1.4 Man's Adaptability

Man is adaptable. He is able to make use of learning and experience to alter his reactions and behavioral patterns. His is truly an adaptive control system. Man reacts to psychological as well as physical needs. Among these psychological needs are comfort, security, safety, anxiety, fatigue, boredom, reward, punishment, and motivation. His effectiveness and efficiency are a function of these psychological factors. They must be taken into account by the designer.

Man is very flexible and can perform well in many different jobs if his limitations are not ignored. As the requirements placed on him become more complex, however, this same flexibility may result in a decrement to system performance. Use the machine to relieve the man of as many routine jobs as possible, but use the man to supply the judgments and flexibility of which machines are incapable (Ref. 48).

Several human factors experts have prepared lists of statements which compare man to machine. Ref. 48 contains a composite of several such lists.

5-8.4.2 Environmental Considerations

Conditions under which equipment—especially military equipment—must be supported vary widely, and in all too many instances are extremely adverse. This is true whether we speak of conditions imposed by the physical environment or the physical and psychological conditions resulting from strain, fatigue, or prolonged worry of the operator.

The machine components of man-machine systems are normally designed to give maximum performance within specified environmental limits; when these limits are exceeded, both performance and reliability suffer. Some support is required under all conditions.

Machines often have failed to fulfill their missions, not because they were poorly designed or badly constructed, but because they demanded more of the operator than was humanly possible considering the environment.

In contrast to equipment, the design of which can be changed, the human being has inherent and relatively inflexible "design" characteristics. The only alternative available is, wherever possible, the exercise of control over environmental conditions to provide reasonably acceptable working conditions (Refs. 42 and 43).

System and equipment designers must be aware of the effects of the environment and take these into account in their designs. There are several types of environments of concern. These include:

1. The geographic or physical environment
2. The operational or working environment
3. The human environment.

The physical environment includes such factors as temperature, humidity, noise, vibration, shock (acceleration), radiation, wind, pressure, salty atmosphere, toxic fumes, sand and dust, insects, fungi, ice, and rainfall. The ***working environment*** includes the arrangement of operating and maintenance work spaces, operational or mission conditions, illumination, acoustics, ventilation, time of day, duration of work, and numbers and skill levels of personnel involved. The ***human environment*** pertains to physical, physiological, and psychological capabilities and limitations of the human being. All of these environments are of specific concern to human factors engineering, and thus to design for maintainability.

Just as the reliability of equipment will be enhanced if the designer assumes that the equipment will be used at the extremes of the various environmental conditions and provides, in his design, features which will allow the equipment to work at these stress levels, the maintainability of equipment will also be enhanced if the designer analyzes the maintenance tasks which must be performed under these environmental conditions and takes these into account in his design.

Geographical-physical environmental conditions are described in detail in Ref. 11. In addition, environmental conditions in all three of the previously listed categories are described in the referenced human engineering guides (Refs. 43 and 48-50).

Successful maintainability design must incorporate consideration of the effects of the working environment on human performance. The environment (both natural and induced) in which maintenance is to be accomplished can have a profound effect on the efficiency with which a technician can carry out his assigned duties. While the design of the physical environment *per se* may not be a responsibility of the design engineer, environmental factors must be considered in equipment design for rapid, accurate, and safe maintenance. For example, design of prime and support equipment for a nuclear-powered system without regard for the maintenance environment would be obviously inappropriate since radiation hazards may require remote handling which, in turn, may require the design of special features into the equipment (Ref. 30).

Among the environmental items which have significant impact on maintainability are temperature, humidity, air circulation, lighting, noise level, vibration, and work space arrangement. Some of these are interrelated, for example, temperature, humidity, and air circulation.

5-8.4.2.1 Temperature, Humidity, and Air Circulation

Temperature, relative humidity, air circulation, and the purity of air all affect human performance. For practical purposes, temperature, relative humidity, and air movement are often combined, and as such, are referred to as "Effective Temperature" or ET. "This is an empirical index that expresses the combined effects of these three characteristics in terms of the subjective feeling of warmth. When the ambient air is completely saturated (100 percent relative humidity) and air velocity is zero, the value of ET is that of the air temperature. Any combinations of temperature, humidity, and air movement that produce the same subjective feeling of warmth are given the same ET value" (Ref. 43). Fig. 5-14 depicts a family of ET curves.

The optimum temperature range for personnel at work obviously varies with the type of work being done and the conditions under which it is being done. For most general purposes, the range of 65° to 70°F is recommended, even if the relative humidity is rather high. In order for maximum efficiency to be obtained, air conditioning should be provided if the temperature exceeds 90°F. The recommended range of 65° to 70°F may be moved upward if the work to be done is light, or downward, if heavy. (See also Fig. 10-5 in Ref. 11).

Prolonged exposure to temperatures below optimum (65°F) may adversely affect work performance. Temperatures below 50°F frequently produce a stiffening of the fingers and a consequent loss of full manual dexterity. When a man has to work in heavy clothing and wear lined gloves or mittens, his efficiency is reduced.

Relative humidity (RH) affects human performance adversely if combined with temperatures that are below or above the optimum. The RH range from 30 to 70 percent is generally acceptable if optimum temperatures for comfort are maintained. At temperatures above optimum, comparatively small rises in RH usually have significantly adverse effects on both comfort and performance.

Proper ventilation is essential to efficient performance over a period of time in an enclosed work area. An adult at work requires 1,000 ft^3 of fresh air per hr. From 20 to 40 ft^3 per min is the recommended rate of air circulation in enclosed work spaces in cold weather; it should be increased slightly in hot weather.

Whenever toxic materials in the air constitute a menace, adequate measures to protect personnel must be taken. Ventilation alone is insufficient. Either the source of contamination must be closed off, or the personnel must be issued protective devices (Ref. 42).

Additional information with regard to the effects of temperature, humidity, and air flow is given in Refs. 11 and 48.

5-8.4.2.2 Illumination

A technician needs appropriate illumination if he is to properly perform the tasks assigned to him; accuracy, speed, and safety suffer when he cannot see clearly what he is to do. On the other hand, it must be realized that adequate illumination will not always be available. Accordingly, the designers of equipment should, as far as is possible, develop their designs to permit maintenance work to be performed effectively under the poorest lighting conditions that are anticipated; to this end, they should acquaint themselves with all of the circumstances that may reduce available

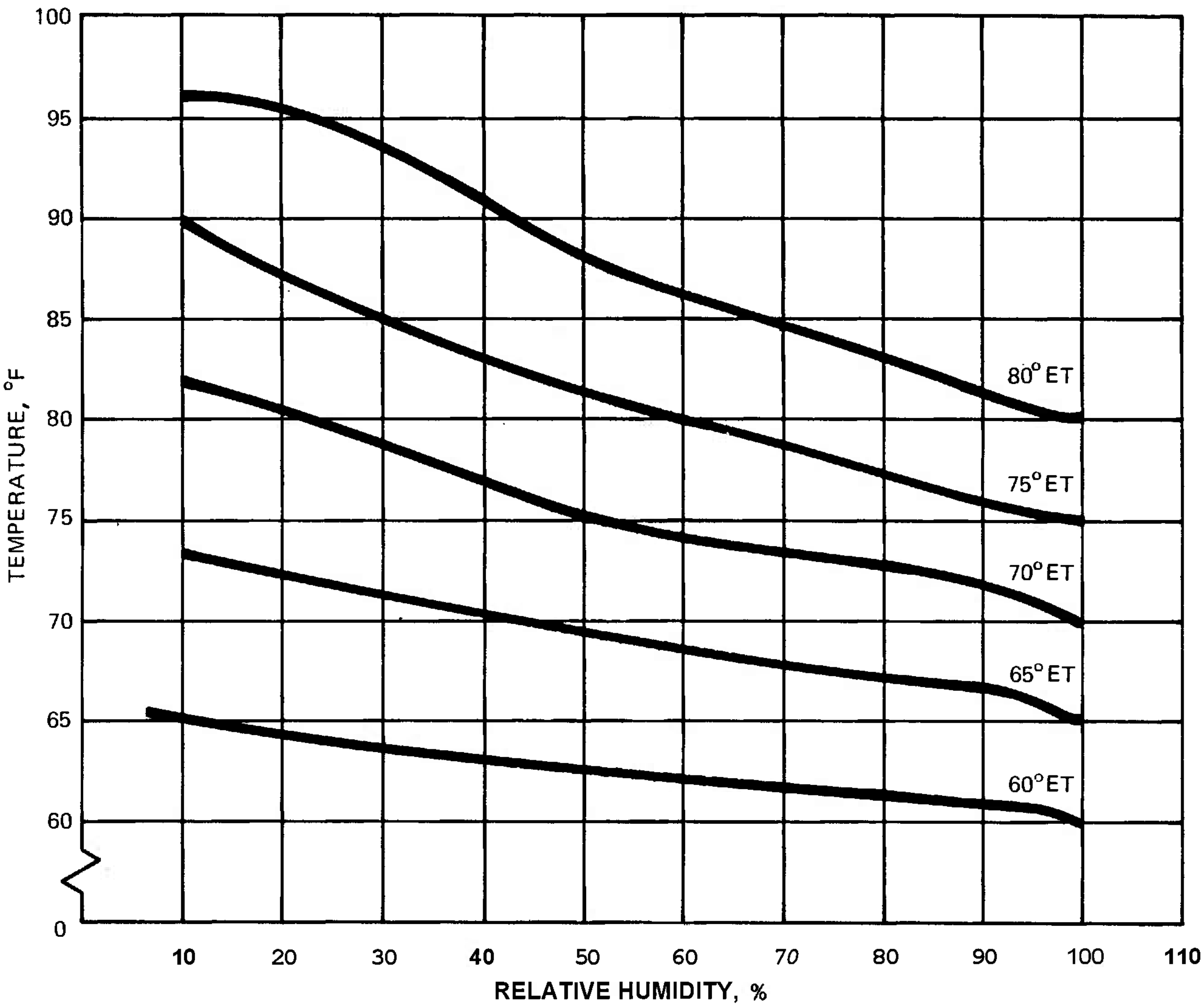


Figure 5-14. Effective Temperature (ET) Curves

illumination. If only a flashlight is expected to be available, the equipment to be developed should be designed so that maintenance work on it can be done with illumination from a flashlight. The employment of plug-in modules and readily accessible and easily replaceable units of light weight will be helpful in this respect (Ref. 42).

There are several important factors that should be considered in the design of any lighting system:

1. Suitable brightness for the task at hand
2. Uniform lighting on the task at hand
3. Suitable brightness contrast between task and background
4. Lack of glare from either the light source or the work surface
5. Suitable quality and color of illuminants and surfaces.

It is difficult to specify exact levels and limitations for all the problems that may arise in designing an efficient lighting system, but analysis recommendations given in Table 5-8 will undoubtedly serve as a safe guide to better seeing for most applications. Design and placement of all lighting elements. should facilitate maintenance and cleaning in order to retain optimum illumination characteristics (Ref. 48). Specific design recommendations with respect to illumination are given in Refs. 30, 43, and 48.

5-8.4.2.3 Noise

Noise is defined as any undesirable sound, even though it might be a meaningful one. The criterion of undesirability is based on the capacity of sound to disrupt communications. Excessive noise in a work area usually reduces the efficiency of the workers, and thus, indirectly, may reduce overall system readiness if the work performed is maintenance; exposure for long periods may result in loss of hearing.

Noise is most clearly described in terms of its two major physical characteristics, frequency and intensity. The frequency of sound is usually measured in hertz (Hz)—the prime factor of pitch. The human ear can detect sounds of frequencies from 20 to 20,000 Hz. Marked individual differences exist, of course, and changes come about with age. Human engineering is concerned primarily with the frequencies to which the ear is most sensitive, namely, those between 600 and 900 Hz.

Personnel exposed for long periods to noise in the range of 4,000 to 6,000 Hz usually suffer major loss of hearing.

The intensity of sound is usually measured in decibels (dB); it is the prime factor in the sensation of loudness. Table 5-9 lists the intensity levels of some common sounds.

Exposure to noise of more than 80 dB may result in temporary or permanent loss of hearing, the extent of damage being determined by the length of exposure.

Excessive noise also affects personnel psychologically. On exposure to it, fatigue occurs more rapidly, ability to concentrate decreases, and annoyance increases. As a result, efficiency declines. Noise conditions in maintenance work areas should be studied and, when necessary, reduced. If reduction is not feasible, the workers should be issued protective devices (Ref. 42). Refs. 11, 30, 42, 48, and 49 contain specific design guidelines with respect to noise.

5-8.4.2.4 Vibration

Vibration is concerned with the effects on human performance of periodic mechanical forces impinging on body tissues. Of interest are vibratory forces the effects of which displace or damage bodily organs or tissue other than those involved in ordinary hearing and/or those that produce perceptible feelings of pain, annoyance, or fatigue. In general, these are high-amplitude, low-frequency vibrations generated by machines of some sort.

The effect of vibrations on the body depends on the physical parameters of the impinging energy; its direction of application with respect to the longitudinal axis of the body; and the mechanical impedance and absorption coefficient of body tissue, organs, and the body as a whole. In addition, because the matching of applied frequencies to the natural frequency of the body and/or its parts will produce resonances, resonant frequencies of the body and its parts assume special importance (Ref. 43).

The human body reacts to vibration and resonating stimuli much the same as does a mechanical system of masses and springs. When the resonant stimulus approximates the natural human-body resonance of about 5 Hz, the person concerned finds this quite disagreeable (Ref. 48).

The parameters of vibration are frequency, amplitude (displacement), velocity, acceleration, and jolt. For a fixed frequency, the last three terms are successive derivatives of amplitude with respect to time.

A detailed discussion of the effects of vibration on the body is given in Ref. 43.

The effect of vibration on the human body depends upon the direction in which these vibrational forces are applied. MIL-STD-1472 states that each direction is to

TABLE 5-8.
RECOMMENDED ILLUMINATION LEVELS (Ref. 42)

Task	Illumination levels, foot-candles		Lighting source
	Minimum	Optimum	
Perception of small detail under low contrast conditions for prolonged periods of time, or where speed and accuracy are essential (Examples: small component repair; inspection of dark materials)	100	125	Special fixture
Perception of small detail under conditions of fair contrast where speed and/or accuracy are not so essential (Examples: drafting; electronic assembly)	50	100	Special fixture
Prolonged reading, desk or bench work, general office, and laboratory work (Examples: assembly work; record filing)	25	50	Local
Occasional reading, recreation, and sign reading where visual tasks are not prolonged (Example: bulletin board reading)	10	20	General
Perception of large objects, with good contrast (Example: locating objects in bulk supply warehouse)	5	10	General
Passing through walkways and handling large objects (Example: loading from a platform)	2	5	General

be evaluated independently in accordance with given limits (Ref. 50).

Equipment that is vibrating when being worked on by a maintenance man creates many small and large problems for him, ranging from the manipulation of controls to the reading of indicators and labels; in any event, his efficiency is reduced. Designer engineers should make every reasonable attempt to eliminate the possibility of vibration from the equipment they are designing. Among the principal means by which this is accomplished, apart from major design features, are vibration insulation, rubber shock mounts, and the cushioning of work platforms and seats (Ref. 42).

5-8.4.2.5 Work Space Arrangement

Work space means the area and volume of space required by personnel in operating and maintaining equipment. Involved in the concept of work space are:

1. The general dimensions and layout of the work area in which the operation or maintenance will be performed.

2. The accessibility of controls, displays, assemblies, and internal maintenance points. Maintenance access to internally located components necessitates insertion of parts of the body, either with or without tools or accessories (Ref. 30).

TABLE 5-9.
SOUND INTENSITY LEVELS (Ref. 42)

Effect on personnel	Intensity level, dB	Remarks
Levels unacceptable as dangerous to personnel	150	Maximum permissible (regardless of the amount of reduction in the ear canal)
	130	Approximate threshold of pain
	120	Loud thunder
	110	Punch press
	100	
Reduction to efficiency may occur above this point	90	City bus
Acceptable noise levels	80	
	70	Heavy traffic
	60	Normal conversation
	50	
	40	Quiet residential area
	30	
	20	Voice whisper
	10	Motion picture sound studio
	0	Approximate threshold of hearing

Many details affecting men and equipment must be considered in the layout of work spaces, but it is seldom possible to provide optimum conditions throughout the design.

The designer should obtain information about requirements before beginning to design the work space. Unless the following information is available, the designer might include undesirable characteristics that are not detectable until late in the development stage:

1. Purpose or mission of the system
2. Mission profile or detailed steps in conducting typical and atypical missions
3. Tolerances allowable in the performance of the system — accuracy, speed, etc.
4. Effects on system performance when various tolerances are not met
5. Specific tasks that the operator must perform — sequences to be followed, relative importance of each task, relative frequency and time duration for each task
6. Inputs to the operator — information that he needs to accomplish his specific tasks
7. Outputs of the operator—data provided by the operator to influence the system
8. Anticipated environmental conditions — temperature, humidity, noise, illumination, vibration, ventilation, radiation, altitude, body position, accelerative forces, etc.
9. Specific pieces of equipment already committed to the design
10. Maintenance access and clearance requirements (Ref. 48).

The space provided for maintenance is not primarily a convenience; it is a requirement to insure an acceptable level of operating efficiency. In laying out a work space, consideration must be given to the methods by which the equipment in it will be maintained, especially if more than a normal amount of space will be required to perform maintenance. Maintenance access requirements are of primary importance in the location of equipment.

Specific details regarding accessibility are given in par. 5-8.2.1. Specific details and guidelines regarding the layout of work spaces are given in Refs. 43 and 48.

5-8.4.3 Human Factors Elements in Designing for Maintainability

The reliability of a system or equipment is concerned primarily with its inherent failure and life characteristics. Maintainability, on the other hand, is concemed with the servicing, inspection, diagnostic, and repairability characteristics of the system/equipment. Maintainability is dependent upon both the operator and maintainer, and this involves the man-machine interface. An outstanding difference between reliability and maintainability, therefore, is the degree of dependence on human factors. Failure to consider human factors in the design will result in increased maintenance problems as well as reduced effectiveness and readiness.

It should come as no surprise, therefore, that all of the maintainability design factors listed in Table 5-5 concern human factors. They epitomize the man-machine system. Therefore, it is no coincidence that most of the maintainability design guides cited in Refs. 11 and 24-38 were put together by human factors engineers. They were the first to recognize the importance of the man-machine interfaces to maintainability and to give specific attention to them (par. 1-1.1). Accessibility, as has already been discussed, is the primary man-machine interface for maintainability.

Test points, controls, displays, labeling and coding, manuals, checklists, and aids are all man-machine interfaces concerned with man's sensing, data processing, and control characteristics as part of the man-machine system (Fig. 5-13). Test equipment, tools, connectors, cases, covers, doors, mounting, fasteners, handles, and handling are also part of the human factors design considerations to facilitate man's role in performing maintenance tasks. They serve no prime purpose in the equipment other than assisting man.

It is essential that design engineers ascertain the conditions under which the equipment they are designing will be used and maintained. If these conditions will be extreme, the design must be altered as much as is practicable in order to protect the equipment. Regarding support activities, such alterations would be made to reduce to a minimum the number of tasks to be performed, and to provide that the tasks which could not be eliminated can now be performed with ease and rapidity under the conditions expected. In conjunction with such efforts, the system engineers should become familiar with the necessary support equipment, such as mobile maintenance facilities and cold weather clothing available (Ref. 42), and the constraints they impose.

Regardless of thoroughness of training and level of skills attained, a technician will and does make mistakes, and such errors frequently cause equipment malfunction, with varying consequences. A driver fails to fill the radiator of his truck, with the result that the engine overheats and the truck stops on the road—inconvenient but not serious. A technician fails to put a cotter pin in a castellate nut in the flight-control linkage of an aircraft, with the result that control of the plane is lost in flight, the plane crashes, and all aboard are killed—very serious. Maintenance requirements are so demanding that all too often they leave no room for human error, yet, man being what he is, personnel failure cannot be completely eliminated. For example, a report by one of the military services revealed that in a period of 15 months, errors made in the maintenance of aircraft contributed to 475 accidents and incidents in flight and ground operations, with 96 aircraft being seriously damaged or destroyed and 14 lives lost. A study of these accidents revealed that many of the failures that produced them occurred shortly after completion of periodic inspections. It also showed that many of the mistakes were repetitious. The conclusions arrived at with regards to the basic causes of these human failures were:

1. Inadequate basic training in the relevant maintenance practices, policies, and procedures
2. Lack of training in the maintenance of the types and modules of the equipment being maintained
3. Inadequate or improper supervision
4. Inadequate inspection.

It follows from this that both operators and technicians need all the assistance the designers of equipment can give them for the effective support of equipment. The principal goals toward which the designers work for this purpose are:

1. Reducing to a minimum the number of support tasks to be performed for each system
2. Designing equipment so that the support tasks required can be performed easily and simply by personnel of specified skills working in specified environments
3. Designing equipment with features that make it difficult or impossible for a task to be performed improperly or incompletely (Ref. 42).

Finally, safety as a design consideration is of importance to both man and machine with regard to protection of life and injury to the man and damage or destruction of the equipment. Safety is discussed in par. 5-8.5.

The maintainability engineer and system/equipment designer should become thoroughly familiar with the handbooks and design guides referenced in this chapter, if a successful total man-machine system is to be realized.

5-8.5 SAFETY

Safety is a condition created by either the nonexistence of hazards or by the utilization of devices provided to give protection against hazards. As such, it is an important objective of man-machine designers. Absolute safety is not attainable, first, because not all hazards can be designed out of machines and, second, because operators and technicians cannot be relied upon to observe safety procedures at all times; as "Murphy's Law" states, "If there is a wrong way to do something, sooner or later someone will do it that way" (Ref. 42).

Safety is one of the important parameters of system design, along with performance, packaging, reliability, maintainability, and human factors (Fig. 1-1). Although safety is often thought of in terms of the prevention of injury or death to personnel, it must also be considered with respect to damage to or loss of equipment, and the resultant effects upon operational readiness and system effectiveness. Indeed, safety is defined as "freedom from those conditions that can cause injury or death to personnel, damage to or loss of equipment or property" (Ref. 51).

System safety engineering is an element of system engineering involving the application of scientific and engineering principles for the timely identification of hazards and initiation of those actions necessary to prevent or control hazards within the system. It draws upon professional knowledge and specialized skills i the mathematical, physical, and related scientific disciplines, together with the principles and methods of engineering design and analysis to specify, predict, and evaluate the safety of the system (Ref. 51).

Costs in time and dollars and the failure of designers to give special attention to this aspect of their work are reasons for the existence of hazards in equipment which could have been eliminated, had they been dealt with otherwise when the equipment was in the design stage. Nevertheless, the majority of accidents that occur are caused by the human component of the man-machine combination. The person at fault may be the equipment operator or the technician charged with its maintenance. A designer who works successfully to minimize hazards in the equipment he is designing can do much to reduce the number of accidents resulting from its operation, but even by employing the best design principles and test procedures, he cannot reduce them to zero.

Inasmuch as hazards cannot be completely designed out of systems, it is imperative that those that remain be clearly recognized and that measures be provided to protect against them. Guards are needed to protect operators and technicians from moving parts, electrical charges, sharp edges and points, high temperature chemical contamination, etc.; in addition, warning signs should be conspicuously placed near dangerous items, and audible warning devices should be added to indicate very dangerous conditions (Ref. 42).

System safety requirements, though normally considered as being essentially in the same general category and scope as reliability and maintainability, may in certain cases be the antithesis of not only the reliability but also of performance requirements. An often obvious, but sometimes subtle, aspect of most hardware systems is that for almost every energy-related functional requirement (propulsion, explosive bolts, separation, radar transmission, lifting, etc.) there is a corresponding requirement to control the actual or potential energy so that it is not inadvertently expended in a manner which results in an undesired, destructive, or injurious incident. This same control requirement concept exists for toxic and corrosive chemicals and materials as well (Ref. 52).

This innate safety requirement is also traditionally recognized on such potentially hazardous materiel as electroexplosive devices (EED's), bombs, rocket motors, propellants, radiation sources, high voltage or high pressure subsystems, and material handling equipment. Numerous safety regulations, specifications, contract exhibits, and technical studies have been documented and published on methods to control the inherent hazards of these items (Ref. 52).

The hazards associated with maintenance and other human tasks performed on equipment are not as well documented. They generally must be considered for each individual system and equipment by means of the performance of safety analyses.

System safety management, as an element of system program management, is intended to insure that:

1. Safety consistent with mission requirements is designed into the system throughout all system planning and acquisition phases.

2. Hazards associated with each system, subsystem, and equipment are identified and evaluated, and eliminated or controlled to an acceptable level.

3. Control over hazards that cannot be eliminated is established to protect personnel, equipment, and property.

4. Minimum risk is involved in the acceptance and use of new materials and new production and testing techniques.

5. Retrofit actions required to improve safety are minimized through the timely inclusion of safety factors during the acquisition of a system.

6. Historical safety data generated by similar system programs are considered and used where appropriate (Ref. 51).

5-8.5.1 System Safety Analysis

There are a number of analytic techniques which are used for system safety analysis. These include hazard analysis, failure modes, effects, and criticality analysis (FMECA), and fault tree analysis. FMECA is also a technique used for reliability analyses, and thus provides a close alliance between reliability and safety. The reliability function is primarily concerned with the assurance that the hardware will accomplish its assigned functions. The safety function is primarily concerned with the assurance that all safety-critical activities have been identified and are controlled—thus minimizing the likelihood of catastrophic events (such as explosion or loss of life) (Ref. 53). Fault tree analysis is also a tool of maintenance diagnostic analysis, and thus provides a close alliance with maintainability. Hazard analysis is similarly closely associated with human factors.

5-8.5.1.1 Hazard Analysis

A hazard is any real or potential condition that can cause injury or death to personnel, or can result in damage to or loss of equipment or property. Hazard analysis is performed in terms of hazard levels. The following hazard levels are defined in MIL-STD-882:

1. Category I—Negligible. Conditions such that personnel error, environment, design characteristics, procedural deficiencies, or subsystem or component failure or malfunction will not result in personnel injury or system damage.

2. Category II—Marginal. Conditions such that personnel error, environment, design characteristics, procedural deficiencies, or subsystem or component failure or malfunction can be counteracted or controlled without injury to personnel or major system damage.

3. Category III—Critical. Conditions such that personnel error, environment, design characteristics, procedural deficiencies, or subsystem or component failure or malfunction will cause personnel injury or major system damage, or will require immediate corrective action for personnel or system survival.

4. Category IV—Catastrophic. Conditions such that personnel error, environment, design characteristics, procedural deficiencies, or subsystem or component failure or malfunction will cause death or severe injury to personnel, or system loss.

Hazard analysis is concerned with identifying potential hazards, classifying them by level, and highlighting those areas which require special design attention to eliminate or minimize the identified potential hazards, particularly in Categories III and IV. Areas to be considered in hazard analyses include:

1. Isolation of energy sources
2. Fuels and propellants: their characteristics, hazard levels and quantity-distance constraints, handling, storage, transportation safety features, and compatibility factors
3. System environmental constraints
4. Use of explosive devices and their hazard constraints
5. Compatibility of materials
6. Effect of transient current, electrostatic discharges, electromagnetic radiation, and ionizing radiation to or by the system. Design of critical controls to prevent inadvertent activation and employment of electrical interlocks
7. Use of pressure vessels and associated plumbing, fittings, mountings, and hold-down devices
8. Crash safety
9. Training and certification pertaining to safe operation and maintenance of the system
10. Egress, rescue, survival, and salvage
11. Life support requirements and their safety implications in manned systems
12. Fire ignition and propagation sources and protection
13. Resistance to shock damage
14. Environmental factors such as equipment layout and lighting requirements and their safety implications in manual systems
15. Fail-safe design considerations
16. Safety from a vulnerability and survivability standpoint; e.g., application of various types of personnel armor (metals, ceramics, and glass), fire suppression systems, subsystem protection, and system redundancy
17. Protective clothing, equipment, or devices
18. Lightning and electrostatic protection
19. Toxic fumes
20. Implosion
21. Nuclear radiation and effects
22. Human error analysis of operator functions, tasks, and requirements (Ref. 51).

Hazard analyses are used to determine safety requirements for personnel, procedures, and equipment used in the installation, operation, test, maintenance, logistic support, transportation, storage, handling, and training, and to evaluate the compliance of system and equipment design with safety requirements and criteria.

A discussion of hazards, their effects, and safe limits, as well as design guidelines for safety is given in various safety manuals and guideline handbooks (see, for example, Ref. 11).

5-8.5.1.2 Failure Modes, Effects, and Criticality Analysis (FMECA)

FMECA is a systematic procedure for determining the basic causes of failure and defining actions to minimize their effects. It may be applied to any level of assembly. In each case, the *failure mode* is described as the particular way in which the item fails to perform its function, independent of the reason for failure—the how, not why. The *failure effect* describes the result of the failure for each possible failure mode—what happens. The *criticality* establishes the category of hazard.

Failure modes include such items as loss of function, loss of output, reduced output, short or open circuit, and rupture. *Failure effects* include mission abort, injury or damage to personnel or equipment, loss of target track, loss of communication, and reduced control. In addition, loss of function or loss/reduction of output may be a failure effect as well as a failure mode. *Failure cause* may be voltage surge, vibration, contamination, overpressure, overheating, wear, or chemical reaction.

As part of FMECA, the reliability engineer tries to determine the causes of failure and the physical mechanisms which cause the failure. He uses stress-strain analysis and the physics of failure to improve item reliability and to inhibit failure. The maintainability engineer uses FMECA to guide him in determining preventive and corrective maintenance tasks to be performed and their frequency. The safety engineer uses FMECA to determine hazards to personnel and equipment.

Thus FMECA is a qualitative means of evaluating reliability, maintainability, and safety of a design by considering potential failures, the resulting effects on a system, and criticality of these effects. Basically, the analysis involves the identification and tabulation of the ways (modes) in which a part, component, assembly, equipment, subsystem, or system can fail. For example, a ball bearing may fail from normal wearout, abnormal wearout, or brinelling. The effect of each mode is identified and the criticality to system and mission operation determined. For example, abnormal wearout will cause increased noise and vibration, with rapid wearing of bearing parts and surfaces, and eventual destruction of bearing and seizing of the item.

In using the analysis, the identified effect may be different depending on the purpose for which the analy-

,is is to be used. In reliability analysis, the effect considered is the effect on the performance of the system or equipment function. In maintainability analysis, the effects include the symptoms by which failure can be identified (such as temperature of the bearing or increased noise or vibration), and the additional parts needing replacement due to damage because of the failure. In safety analysis, the additional effects considered would be damage to adjacent items and equipment, and possible danger to personnel.

The importance of FMECA to safety analysis is the identification of potential hazards and their consequences. A well prepared FMECA for reliability analysis often suffices for safety as well.

5-8.5.1.3 Fault Tree Analysis

Fault tree analysis is a technique to measure system safety by determining the probability that an undesirable event, or fault, will occur. A typical fault tree is shown in Fig. 5-15, where the undesired event, "unsafe failure of protective system", occurs if the system fails in a mode such that *S* would remain energized despite occurrence of the abnormal condition. The fault tree method may be summarized as follows (Ref. 54):*

1. The undesirable event, or fault, whose possibility and probability are to be investigated is selected. This may be inadvertent or unauthorized launch, ignition of an ordnance device, failure of equipment to perform an operation, injury to personnel, or any similar mishap.

2. Functional flow diagrams of the proposed system design are analyzed to determine those combinations of events and failures which could contribute or would be necessary to an occurrence of the fault.

3. Contributory events and failures are diagrammed systematically to show their relation to each other and to the undesirable event being investigated. The process begins with the events which could directly cause the undesirable event (first level), and working back step-by-step through the system to determine combinations of events and failures which could bring about the end result. The diagrams so prepared are called "fault trees". When more than one event is involved, the chart indicates whether they must all act in combination (AND relationship), or whether they may act singly (OR relationship).

4. Suitable mathematical expressions to represent the fault trees are developed using Boolean algebra. These expressions will be the mathematical statement of the AND/OR relationship and can be simplified.

5. The circumstances under which each of the events in the fault tree could occur are determined. This consists of examining each component of the subsystem capable of producing an event in the fault tree, and determining how its failure would contribute to a mishap.

6. An estimate is made of the probability of occurrence of each event or the failure rate of each component or subsystem. This may be from known failure rates obtained by past experience, vendors' test data, comparison with similar equipment, or experimental data obtained specifically for this equipment. These probabilities or failure rates are entered into the simplified Boolean expressions.

7. The probability of occurrence (adversity) of the undesirable event being investigated is determined from the probabilities of occurrence of the contributory events. This procedure will also identify the most influential factors, and any sensitive elements whose improvement would reduce the probability of a mishap.

Certain assumptions are made concerning the characteristics of the components and their operations. These are:

1. Components and subsystems either operate satisfactorily or fail; there is no operation with partial success.

2. Basic failures are independent of each other.

3. Each basic item has a constant failure rate which conforms to an exponential distribution. This assumption can be modified to accommodate other distributions.

Fault tree analysis has two major disadvantages: it is fault oriented and is linear. It is concerned only with those factors which may contribute to occurrence of that fault. Also, like failure effects analysis, it is inadequate in its consideration of when a fault may occur. It is an excellent procedure within these limitations (Ref. 54).

Generally, fault trees serve three purposes:

1. As a tool for accident analysis, a fault tree aids in determining the possible causes of an accident. When properly used, the fault tree often leads to discovery of failure combinations which otherwise would not have been recognized as causes of the event being analyzed.

2. The fault tree serves as a display of results. If the design is not adequate, the fault tree shows what the weak points are and how they lead to the undesired

*Adapted with permission from W. Hammer, "Numerical valuation of Accident Potentials", *Annals of Reliability and Maintainability,* Vol. 5, 1966, American Institute of Aeronautics and Astronautics, New York, pp. 494-500.

event. If the design is adequate, the fault tree shows that all conceivable causes have been considered.

3. For reliability analysis the fault tree provides a convenient and efficient format for the problem description (Ref. 55).

A fault tree is an event-logic diagram relating component failures to a particular system failure. An "event-logic diagram" is a logical representation of the interrelationship of various events occurring within a complex system, such as a missile or a nuclear reactor. The logic diagram is constructed using events interconnected by logic "gates". Each logic gate indicates the relation between a set of "input events" and an "output event". The input events are considered to be causes of the "output events". Output events from most gates serve as input events to other gates. An input event which is not the output of any gate becomes a basic input event. Only a few types of logic gates are used and the logic of each is simple and completely defined (Ref. 55).

Construction of a fault tree begins with definition of the "top" undesired event (the system failure interest). The causes are then indicated and connected to the top event by a logic gate. The procedure is then repeated for each of the causes and the causes of causes, etc., until all events have been fully developed. The events are considered fully developed when the causes have been shown for all events except independent component failures. The latter are considered basic input events. Occasionally, subsystem or equipment failures are used as basic input events if they are independent of all other basic input events (Ref. 55).

The subject of basic input event independence warrants emphasis. All basic input events for a fault tree must be independent; unless the failure can be caused by other events, the causes must be explicitly shown.

Some of the relative advantages of fault tree and FMECA analysis follow (Ref. 53):

1. The fault tree is the optimum technique for multiple failures, whereas FMECA analysis is the optimum technique for single failures (Table 5-10).

2. The fault tree does not require analysis of failures which have no effect on operation under consideration, whereas FMECA analysis provides documentation to insure that every potential single failure has been examined.

3. The fault tree is event oriented. It easily identifies higher level events or events subsequent to failure. The FMECA analysis is hardware oriented. It easily identifies results of failure of any component, subsystem, or system.

4. The fault tree identifies all external influences which contribute to loss—such as human errors, environment, and test procedures—whereas FMECA analysis does not require investigation of as many external influences, and the associated data are not required.

5. The fault tree has a restricted scope with analysis in depth while FMECA analysis has a broader scope with restricted depth of analysis (Table 5-10).

5-8.5.2 Safety and Maintainability Design

Safety and maintainability are closely related. Safety is listed as one of the most important maintainability design factors (Table 5-5). Because technicians must perform maintenance tasks on equipment during both equipment on and off periods, they are exposed to hazards and are subject to accidents. Many of these hazards are created by careless design or insufficient attention to human factors and safety features during design. Some are created by environmental conditions. In addition, the technician may create hazards to himself and other personnel or to equipment if he is careless while performing maintenance tasks.

In addition to the safety analyses described in the preceding paragraphs, the process of maintenance engineering analysis (see Section I of this chapter) determines maintenance tasks and requirements which affect both design for maintainability and design for safety.

Some general guidelines for the design of equipment to provide for the safety of technicians and operators are:

1. Items and subassemblies that will need maintenance should be located and mounted so that access to them may be gained without danger to technicians from electrical charge, heat, moving parts, toxic chemicals, and other hazards.

2. Access openings should be fitted with fillets and rounded edges, and large enough to permit easy entrance.

3. Fail-safe devices should be provided so that a malfunction in one unit or subassembly cannot cause malfunctions in other units or subassemblies, with resultant serious damage to the system and possible injury to personnel.

4. Potential sources of injury by electrical shock should be carefully studied. The effect of electrical shock depends on the resistance of the body, the current path through the body, the duration of the shock, the frequency of the current, the physical condition of the individual, the amount of current and voltage, and the size of the contact area. A designer has some control over these last two factors and should exercise it in

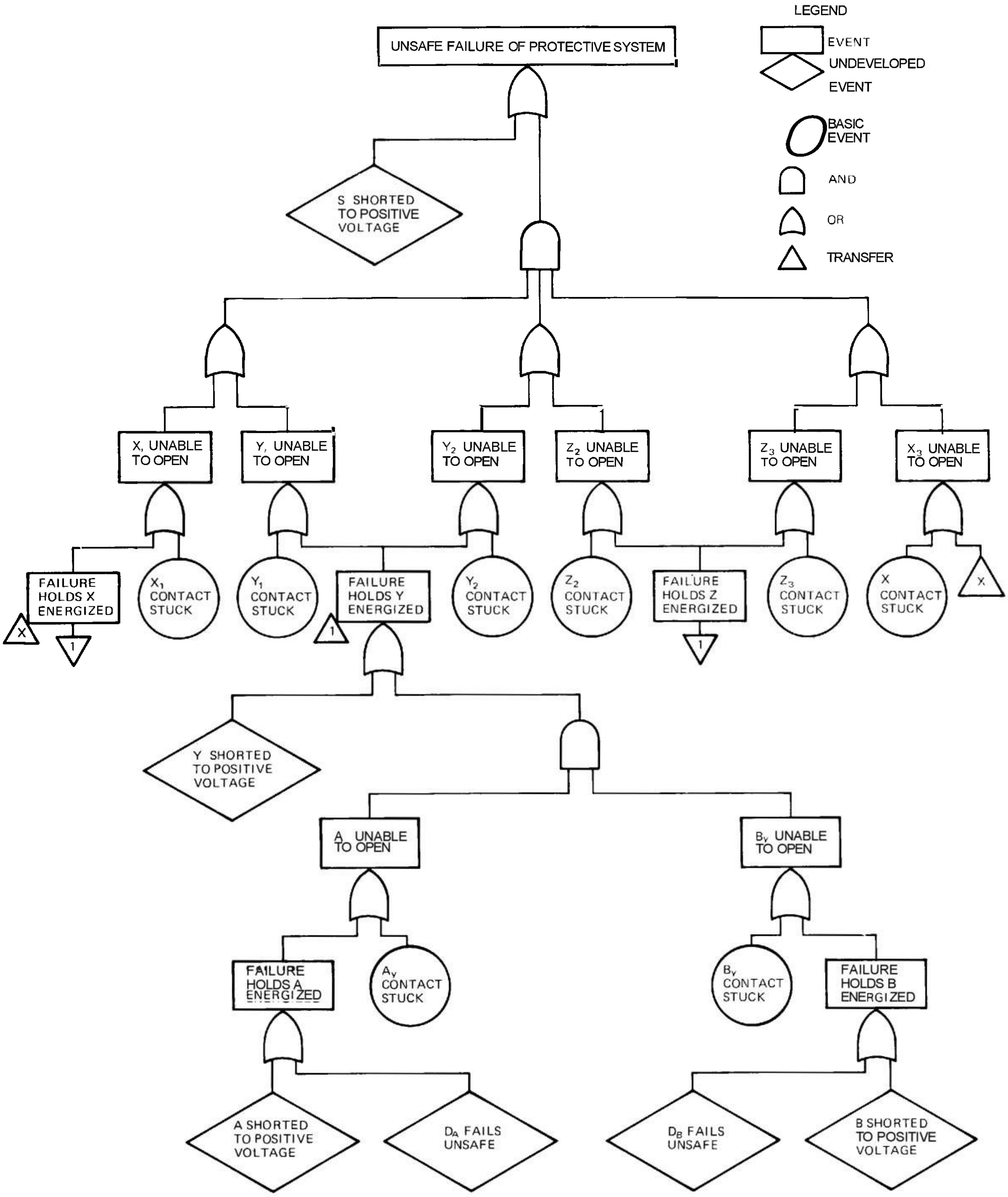


Figure 5-15. Fault Tree Example

Reprinted with permission from R. J. Schroder, "Fault Trees for Reliability Analysis", *Proceedings 1970 Annual Symposium on Reliability,* Institute of Electrical and Electronics Engineers, New York, pp. 198-205.

the interest of personnel safety (Refs. 11 and 42 give tables of electric voltage and current and their lethality).

For specific information on safety guidelines see Ref. 11. In addition to designing equipment so as to make it as safe as possible for technicians and operators, designers should also give attention to the design of work areas for the same purpose. Some design guidelines to this end are:

1. Adequate fire-extinguishing and other fire-fighting equipment-of the proper type—should be made available in areas where fire hazards exist.

2. Emergency doors and other emergency exits should be placed so as to provide maximum accessibility.

3. Eye baths, showers, and other special first-aid equipment should be provided if toxic materials are to be handled.

4. The weight-lifting or -holding capacity of each stand, lift, hoist, jack, and other such equipment should be clearly indicated on the item itself.

5. Guides, tracks, and stops should be provided to facilitate the handling of equipment (Ref. 42).

In general, the attention given to safety features in design will be repaid many times in the conduct of support operations, even though such features alone cannot eliminate accidents. Measures recommended for reducing the number of accidents caused by human error include:

1. Make certain that every man is properly trained to perform his assigned duties.

2. Prepare support procedures which will minimize the chances of human error.

3. See to it that supervisors constantly check that the support procedures established are being properly followed.

4. Alert every man to the hazards involved in the

TABLE 5-10.
SCOPE OF ANALYSES

		Theoretical Fault Tree Mission Loss	
		Practical Fault Tree Catastrophic Loss	**Noncatastrophic loss**
		Human Errors Causing Catastrophic Loss	**Human Errors Causing Mission Loss Only**
Theoretical FMECA Analysis		**Multiple Hardware Failures Causing Catastrophic Loss**	**Multiple Hardware Failures Causing Mission Loss Only**
	Practical FMECA Analysis	**Single Hardware Failures Causing Catastrophic Loss**	**Single Hardware Failures Causing Mission Loss Only**

Reprinted with permission from K. H. Eagle, "Fault Tree and Reliability Analysis Comparison", ***Proceedings 1969 Annual Symposium on Reliability,*** **Institute of Electrical and Electronics Engineers, New York, pp. 12-17.**

work he is doing, and to the possible consequences of his failure to perform his duties correctly.

5. Design mating parts so that they can only be mated and assembled in the correct configuration.

6. Have all work properly inspected.

7. Provide proper tools, and adjust safety equipment.

8. In general, make all workers safety conscious (Ref. 42).

Wherever safety is involved, no one may assume—every one must make certain. Ref. 11 contains a safety checklist for maintainability design.

5-8.6 TEST AND CHECKOUT

The test and checkout features of a system or equipment are an essential part of design for maintainability. Three of the four corrective maintenance downtime categories are concerned with test and checkout. Ten of the thirteen maintainability design factors given in Table 5-5 are influenced to some degree by the system test. and checkout philosophy adopted.

Test features in a system provide the means for performance monitoring, fault detection, fault location and isolation, calibration and alignment, and system checkout. The test philosophy might be, at one extreme, to have the equipment operator be the only test means, or, at the other extreme, to have a complex, completely automatic test and checkout system. In between lie such concepts as manual testing aided by the use of general-purpose or special-purpose test equipment, built-in test, semi-automatic test and checkout, and fully automatic test and checkout. Which of these concepts should be used for any particular system depends upon trade-off among many items. These include:

1. Technician skill levels available
2. System effectiveness requirements
3. Mission criticality of performance features
4. Complexity of the item to be tested
5. Accuracy and precision requirements
6. Number of tests which must be made
7. Frequency with which tests must be made
8. Maintenance level at which tests are made
9. Availability of general-purpose or standardized test equipment
10. Nature of the system to be tested-electronic, mechanical, hydraulic, pneumatic, optical, or combinations of these
11. cost.

Of specific importance to the test and checkout problem is the state-of-the-art of both the prime equipment design and test discipline. It is customary for prime system designers and users to want the test equipment to be an order of magnitude more accurate than the item to be tested. This creates significant problems when the prime system and equipment themselves are pushing the state of the art, often resulting in uncertainties as to which equipment is testing which. These order-of-magnitude and state-of-the-art conditions, coupled with the tendency on the part of equipment designers to concentrate on the prime equipment design to the exclusion of test considerations until late stages of design, make the task of the test equipment designer extremely difficult in the shortened time frame which remains.

Test equipment is required for inspecting systems and components, monitoring system performance, controlling quality of production, and facilitating maintenance. Test equipment is applied to operational systems at all levels of assembly under both actual and simulated operational conditions. For each new weapon system, someone must determine the type and quantity of test equipment required and the level of assembly and maintenance at which it is to be applied.

Test equipment program management must be aimed at achieving cost-effective utilization of existing and future test equipment, reducing current and future proliferation of test equipment, and eliminating undesirable duplication in development of new test equipment. Total test equipment program management cannot take place in a vacuum—it must take place in conjunction with, and as an integral part of, weapon system design and development, production planning, and support planning (Ref. 57).

It is, therefore, essential that a purposeful dialogue be established between prime and support system design personnel early in the program. Test requirements and equipment features *must* influence prime system design if a cost-effective total system design is to be achieved. Without such a dialogue, design considerations such as test points, adjustments, sensors, displays, and built-in test features will be seriously compromised, either by omission or by proliferation, either of which will be detrimental to system effectiveness. By establishing an early dialogue, maintainability and test equipment engineers can determine the need for testing, the nature of the tests to be performed, and their criticality and contribution to system effectiveness, as well as providing for the inclusion of proper test features.

The test equipment engineer, and indeed the maintainability engineer, must always question the need for each proposed test requirement and limit the amount of testing and the number of test points to those which are essential at each appropriate maintenance level. This requires the early establishment of a suitable test and checkout philosophy as part of the overall Maintenance Support Plan and rigid adherence to this philosophy so that the exceptions do not become the rule. The high cost and low effectiveness of many, systems can be traced to the over-complexity of both prime and support systems resulting from the failure to establish and adhere to an effective dialogue between prime and support system designers starting during the early system planning phases.

There are several additional reasons why intensified management of the development and use of test equipment would be fruitful. First, while Integrated Logistics Support (ILS) objectives address test equipment as well as other support resources, recent experience with ILS planning indicates that detailed procedural guidance is required if the objectives of ILS are to be effectively realized. Effective management of test equipment is an area where such procedural guidance can be developed. Second, the type of test equipment built into a weapon system or used for on-line and off-line testing will have a strong influence on the Maintenance Support Plan for the system. In fact, repair level analyses have indicated that the quantity and cost of test equipment is the single most predominant economic factor in determining whether a repairable item should be discarded at failure or repaired at organizational, field, or depot levels (Ref. 57).

Finally, test and checkout equipment may account for a considerable portion of system cost. It has been estimated that the military department spend more than $500 million annually on the development and procurement of test equipment in addition to those test features which are designed into the prime equipments. Special tools and test equipment in the DoD inventory may well exceed $3 billion.

5-8.6.1 Elements of Test and Checkout

Certain basic functions are associated with the testing process regardless of the nature of the unit under test (UUT). These functions can be identified in both manual and automatic testing and are:

1. Test control (programming)
2. Stimulation
3. Measurement and display.

Test control can be identified in either a manual or automatic test system. In the manual case, the control function is the responsibility of the human operator. The control subsystem will direct the operation of the test in the proper sequence, set up the stimuli required to exercise the UUT, and command the collection and processing of the test data. The operation of directing the control subsystem to perform a test sequence is referred to as programming. Programming for human control is by word of mouth, written routines contained in test manuals, checklists, and visual displays. Machine systems require a program in a language format that the machine can decipher.

The essential elements of test and checkout are shown in the block diagram of Fig. 5-16.

The programmer (manual or automatic) applies the appropriate input stimulus to the UUT. It also selects the appropriate reference standard. The output of the UUT is then fed to the comparator where it is compared with the standard and the result indicated on the output display indicator.

In its simplest form, the programmer is the equipment operator or maintenance technician. The input stimulus may be an electronic signal or voltage, a mechanical force, or displacement, a hydraulic or pneumatic pressure, or any other perturbing force. The output may be the visual, aural, or other sensory receptors of the technician himself. In a more sophisticated form, the programmer may be a computer program initiated by the operator or by a signal from the system itself. The input stimuli will then be a set of programmed inputs with magnitudes and other characteristics established by the program through an input signal conditioning unit.

The outputs from the UUT are fed into a measurement section (comparator and display) where they may be further conditioned, compared against selected reference standards, and then indicated by the display unit. The display unit is a device for presenting information to the operator concerning the state or condition of the UUT or of the test equipment itself. The display unit may consist of warning lights, audio signals, video displays, or printed information. Test output may also be stored on cards or tapes for future use, or fed into a data bank for statistical analysis.

The UUT may be any system or subsystem, assembly, subassembly, module, or component undergoing testing. The purpose of the testing process may be to provide assurance that the UUT is functioning within performance specifications as in the case of built-in test equipment (BITE), on-line test equipment, or some type of production test equipment; or the purpose may be to locate the exact fault of a UUT that has previously been determined to be in a failed condition, using off-line diagnostic test equipment. In any event, the UUT must be designed to allow access to various test points

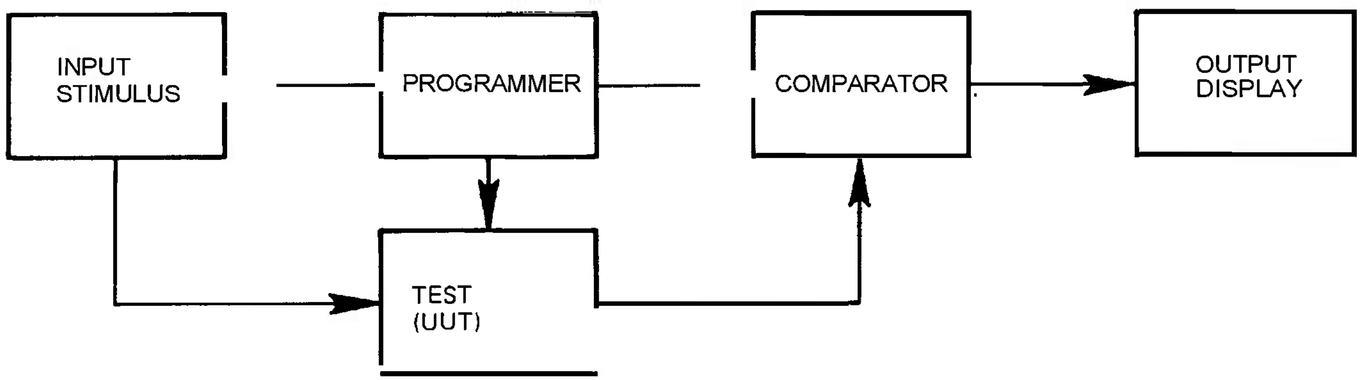


Figure 5-16. Block Diagram—Test and Checkout System

which are compatible with the test program and, to some extent, with the programmer-comparator (Ref. 57).

It should be noted that an automatic test system designed for general application of various UUT's must impose certain design requirements or constraints on the units to be tested, such as accessibility to test points. For example, the Navy has developed MIL-STD-1326 (NAVY), ***Test Points, Test Point Selection, and Interface Requirements for Equipments Monitored by Shipboard On-Line Automatic Test Equipment*** (Ref. 58).

In addition to test points, there are other design features, testing procedures, and maintainability requirements necessarily imposed on UUT's intended for use with a general purpose automatic test system. All test and checkout systems must contain these basic features. Depending upon system complexity, the ability of the test system to handle more than one prime system or equipment, the standardization of equipment and UUT design, and the phase of the system life cycle in which support system considerations are begun, additional auxiliary items may be required in order to perform an adequate test and complete checkout.

The most common of these auxiliary items are test adapters which provide the interface between the UUT and the test system. Adapters may be simple mechanical fixtures and test connectors. They may also include special features, such as special signal and power sources, signal conditioning, and test sequence switches which. are not incorporated in the test system. Although it would appear that little standardization potential exists for designing common or multipurpose adapters, it is apparent that the opportunity does exist to make design/economic trade-offs with regard to design features of the UUT, maintenance concepts to be applied, the versatility of the basic test equipment, and the method of testing. The cost associated with development, acquisition, storage, identification and modification of the adapter is but one of the factors that should be considered in such design/economic trade-offs.

5-8.6.2 Purpose and Type of Tests

There are a number of purposes involved in testing. These include:

1. Production acceptance and quality assurance
2. Calibration and alignment
3. Fault detection
4. Fault diagnosis
5. Verification
6. Prognosis.

Each of these purposes may be satisfied by one or more of the following test methods:

1. Functional (performance) testing
2. Performance monitoring
3. Marginal testing
4. Checkout
5. Self-test
6. Diagnostic testing.

Types of test techniques include:

1. Dynamic or static
2. Open-loop or closed-loop
3. On-line or off-line
4. Quantitative or qualitative
5. GO/NO-GO or interpretive.

All of the listed purposes, methods, and types of tests are concerned with maintenance and maintainability to some degree at one or more of the different maintenance levels (see par. 5-4.1).

5-8.6.2.1 Test Methods

Production acceptance and quality assurance tests are tests performed in a factory or depot and are concerned with the production .repair, rebuilding, or overhaul of systems and their components. These tests normally include functional and marginal test methods; they may use dynamic or static, open-loop or closed-loop tests; and they are usually quantitative and interpretive. Quality assurance testing additionally will include some marginal testing performed under environmental test conditions.

Calibration and alignment tests are performed not only during production and quality assurance testing, but also in the field at all echelons of maintenance and to varying degrees and accuracy levels. Calibration and alignment tests are performed as part of scheduled maintenance and as part of the verification of corrective maintenance actions. They may even constitute the repair action itself during the diagnostic period. Calibration and alignment testing uses the same test methods and techniques as described for production acceptance and quality assurance testing.

Fault detection, fault diagnosis, verification, and calibration and alignment tests are concerned with three of the four corrective maintenance downtime periods (Fig. 5-3) and can have a significant impact on minimizing downtime. They utilize all of the test methods and techniques listed.

Complex systems are frequently given readiness tests by one or both of two general methods: (1) functional tests, which simulate normal operating conditions as closely as possible; and (2) marginal tests, which attempt to isolate potential trouble areas by simulation of abnormal operating conditions.

Functional tests are commonly used for missile checkout equipment; for example, a missile is tested by supplying typical guidance data to it and observing the response of controls to computed error signals. Such tests are performed under normal temperature, mechanical, and electrical conditions.

In a functional test, the test equipment simulates significant inputs to the system being tested, evaluates final outputs against a set of predetermined standards, deduces the incidence of any probable faults, and indicates to the test equipment operator the system state of readiness.

Marginal tests, on the other hand, supply data about a system while it is being subjected to marginal and limit signals and severe environmental conditions, such as vibration, extreme heat, or lowered power-supply voltages (Ref. 42).

The system may be tested at periodic intervals or just prior to entering a mission state. This is called system checkout. *Checkout* is defined as "a sequence of functional, operational, and calibrational tests to determine the condition and status of a weapon system or element thereof." (AR 310-25). System monitoring and checkout are performed, therefore, to detect actual performance failures or to allow the prediction of incipient failures. Checkout is also performed during the verification period of corrective maintenance downtime to ascertain that the repaired system is once again operating within acceptable limits.

Prognostic tests are used for monitoring degradation in system and component performance and for predicting when nonscheduled preventive and corrective maintenance should be performed in order to inhibit system failure. During system operation, certain critical performance functions may be continuously or periodically monitored in order to ascertain that they are within acceptable limits. If these test data are recorded, the information obtained may be used for degradation analysis and prognostic evaluation. Marginal testing has its greatest value when it is used in support of fault prediction because it identifies many incipient failures resulting from abnormal environmental and operating conditions.

Prognostic testing presents a great challenge to the test equipment designer. There is no question of the importance of being able to predict the time to failure for the main failure modes of a system. In automotive and aircraft vehicles, for example, the most promising possibilities appear to be through oil analysis, onboard recording, vibration analysis, and seeding of critical components. A component is seeded so that when it wears beyond a certain limit an easily detectable chemical substance is released (Ref. 59).

One of the most promising approaches to predicting impending mechanical failure is analysis of oil. Industrial programs and the present US tri-service SOAP (Spectrometric Oil Analysis Program) operate by periodically recording the percent concentration of key elements such as copper, nickel, iron, lead, and other metals found in oil taken from lubricating systems, comparing this measurement with data compiled from tests which correlate actual wear of parts with the amount of metal in the oil, and evaluating the comparison to predict failures of system parts. Abnormal rates of increase in these elements are flagged and then correlated with historical data for final determination.

The instrumentation used is either a militarized emission spectrometer or a commercial atomic absorption spectrometer. Both types of instrument measure the percent concentration of impurities in a batch testing process on samples taken from engine oil, hydraulic lines, gearboxes, and other oil reservoirs. This percentage is converted to either Wear Metal Units (WMU) or parts per million (ppm) by comparison with oil standards having known amounts of contaminants (Ref. 59).

Spectrometric analysis utilized as a separate detection or diagnostic technique for wear metals or in conjunction with physical/chemical tests has proven effective in all military aircraft engines and gearboxes. Comparisons and evaluations of increases of certain elements in the oil-wetted areas diagnosed categorize an impending failure of the component (component fabrication already established). Physical and/or chemical tests to supplement or confirm the failure are required on occasion for support. Standardized techniques, training, and communication exchanges between military and industry have stimulated worldwide interest in maintenance efforts and ultimately in aircraft safety and savings.

Military agencies are now able to analyze as many as twelve elements spectrometrically with response time reduced to several hours on routine samples. Frequency distributions plotted with confirmed diagnosis (hits) are used to establish evaluation criteria for maintenance recommendations. Wear trend data compiled over a period of time demonstrate repair/replace/overhaul type of maintenance.

The potential for spectrometric analysis has increased, especially in areas of large equipment.

Performance checking is, in essence, a verification of UUT operation against its performance specification and can be a go/no-go test in its simplest form. This type of testing is usually associated with product acceptance or product certification, and with system checkout and maintenance for fault detection. For example, a determination of maximum engine horsepower by a dynamometer will reveal quickly whether or not the engine is meeting the manufacturer's specification. On the other hand, if the engine fails to deliver the specified horsepower, the dynamometer test will not diagnose or pinpoint the source of the problem (Ref. **59).**

Perhaps one of the most significant, as well as most difficult, types of testing is ***diagnostic testing.*** Studies have shown that this downtime period is by far the largest contributor to corrective maintenance downtime. The purpose of diagnostic testing is to first localize the failure to a specific subsystem, equipment, or assembly, and then to isolate the replaceable or repairable end item which has failed.

In many cases, fault detection tests are run initially but, instead of merely stopping or recording data on an out-of-tolerance test, recourse to a fault isolation sequence is used to identify a specific maintenance action. The test sequence can be designed to provide fault isolation or diagnosis to a predetermined depth of detail. The diagnostic and performance tests often overlap and a performance test is frequently used together with another test to diagnose a particular problem. For example, if a peak horsepower test shows a low performance, then a check of ignition timing against specified limits may pinpoint the cause of the problem as incorrect timing. On the other hand, the success of the diagnostic analysis may rely on the processing of a vibration waveform which has no connection with the manufacturer's specification (Ref. **59).** Diagnostic testing is discussed in par. **5-8.6.5.**

Regardless of the type of test equipment needed, a test should be sufficiently comprehensive to determine whether or not a system is operating properly within its design specifications. In the interests of economy, it may be necessary to limit a test to the components of a system that have the most failures. Other principles of exclusion specify components that are not really vital to system operation, those that are only marginally significant, components that cannot be tested, and single-shot power sources (Ref. 42).

5-8.6.2.2 Types of Tests

5-8.6.2.2.1 System vs Component Tests

System tests are performed on systems as single entities; in each such test a set of stimuli is presented as system inputs, and the responses are recorded. To the greatest possible extent, the test is performed as a continuous series.

Component tests, by which every individual subsystem or component of a system is tested, are made on the assumption that if all subsystems and components function satisfactorily, the system is in satisfactory operating condition.

The differences between these two types of test can be illustrated by the example of a simple amplifier. A system test of such an item would consist of providing a representative signal as an input, and measuring the response at the output for gain, phase shift, noise, transient response, etc. A component test, however, would measure power-supply voltages, resistances, capacitances, continuity, or some other performance characteristic to find potential difficulties.

Usually, the system test is the more desirable of the two types, because it is more comprehensive and represents ***functional*** performance, and therefore is more useful. No amount of component testing will definitely ensure, on the basis of test findings, that a system will actually operate as specified.

Component testing is sometimes used as an adjunct to system testing, to facilitate troubleshooting. It is especially useful for isolated networks of passive components, cable and harness assemblies, etc. It should rarely be used as a prime test technique (Ref. 42).

5-8.6.2.2.2 Static vs Dynamic Tests

System tests may be run according to either static or dynamic principles. The selection is determined by the type of system to be tested, which may dictate either one or a combination of both. Whichever technique is adopted, the ultimate aim is to exercise every subsystem as a means of determining the cumulative effect of all the subsystems on overall system performance.

A static system test is one in which a series of steady-state input signals is fed in, and the output indications are monitored to measure system operation. Although this kind of test can provide information on the transient behavior of a system through time synthesis, in practice the transient response is often ignored in favor of the simpler steady-state response.

In a dynamic system test, a transient or varying input signal is applied to a system, and the output signals are noted and analyzed to determine whether or not system requirements are being met. A test of this type more nearly simulates a typical mission of the system being tested; thus every major subsystem is checked in this test.

A static test, of course, is the simpler of the two types, and is therefore generally easier to perform. It yields results that establish a confidence factor for a system, but cannot go beyond this. A dynamic test, on the other hand, produces much additional information, such as integration rates, phase and delay characteristics, and frequency and transient response of a system to either typical or boundary-value inputs. For these reasons, a dynamic test is recommended wherever it can be used (Ref. 42).

5-8.6.2.2.3 Open-loop vs Closed-loop Tests

A further refinement of either a static or a dynamic system test is the application of open-loop or closed-loop testing techniques. The open-loop test provides for zero intelligence feedback to a system being tested. In a test of this type, system response is evaluated as the system end product, and no adjustment of the stimulus is made; i.e., a preset stimulus is fed into a system and the response is independently evaluated.

In a closed-loop test, input stimulus is continuously adjusted as a function of the response. Adjustment is computed on the basis of system behavior under test conditions; for example, in a guided missile test, the flight characteristics of the missile will control the input stimulus at all times.

Open-loop tests sometimes provide more useful maintenance information, because they make possible direct observation of system transfer functions, free of the modifying influence of feedback; this makes possible ready measurement of any degradation of system performance. Closed-loop tests provide much information for use in evaluation of system performance, effectiveness of design, adequacy of tolerances, and related characteristics (Ref. 42).

5-8.6.2.2.4 On-line vs Off-line

On-line testing is performed on the UUT while it is installed as part of the system in its operating environment. It may consist of ***performance monitoring*** utilizing built-in test features or test points and connectors for the use of external test equipment. It may use the normal operating signals and conditions of the system rather than external stimuli. Or, it may consist of ***periodic checkout*** or ***fault diagnosis*** using internal or external test equipment and standardized stimuli. When used, on-line testing is normally performed at organizational and direct support maintenance levels.

Off-line testing is generally performed at direct support, general support, and depot maintenance levels; occasionally, it may also be performed at the organizational level. Off-line testing consists of removing the UUT from its operational environment and testing it on a general-purpose or special-purpose tester. It allows tests to be performed to greater depth and detail, as well as to higher accuracy and precision. It also generally requires higher skill levels than on-line testing.

5-8.6.2.2.5 Quantitative vs Qualitative

Tests may be quantitative, i.e., the measurement of a specific performance parameter within prescribed limits of magnitude, time, and/or shape; or tests may be qualitative, the indication of the presence or absence of some system attribute. Quantitative and qualitative tests thus differ in their information content. Qualitative tests are usually simpler to make and to indicate, but do not permit prognostic evaluation. They are more desirable for organizational level and quick-check tests than for diagnostic, detailed testing.

Quantitative tests contain more information about the system or component and, therefore, can be used for more detailed examination of the performance of the UUT. Certain quantitative tests are difficult to make in an operational environment and in an on-line mode of test. For example, tests of high energy, high frequency, electromagnetic emission, and armament firing signals are difficult to perform in the operational system, or may not be permitted for safety or security reasons. However, quantitative tests rather than qualitative tests are required for calibration and alignment.

5-8.6.2.2.6 Go/no-go vs Interpretive

Go/no-go testing does not imply qualitative testing rather than quantitative. Go/no-go and interpretive tests relate more to the method of display and the degree of interpretation of results which is desired or needed by the operator or maintenance technician. Go/no-go tests may be specified because of limitations as to technician skill-level or because the nature of the test or operational situation is such that a noninterpretive type of readout is more effective. Go/no-go type displays are usually preferred for organizational level checkout tests.

There are many instances, however, in which interpretive displays are desirable. For example, a cathode-ray tube display may give the technician information about transients or the waveshape or marginal tolerances which will allow him to take proper corrective measures. Interpretive displays are generally preferred for diagnosis and adjustments at maintenance levels other than organizational.

5-8.6.3 Classification of Test Equipment

Test equipment may be classified in a number of ways (Ref. 57). Among these are:

1. Method of operation (degree of automation)
2. Point of application (maintenance level, production)
3. Design origin (military, commercial)

4. Versatility (general-purpose, special-purpose)
5. Interface with prime equipment (built-in, on-line, off-line).

These categories of test equipment are discussed in the paragraphs that follow.

5-8.6.3.1 Method of Operation

The method of operation of the test equipment is a function of the degree of automation. These range from fully automatic to no test equipment as follows:

1. Fully automatic
2. Semi-automatic (pre-programmed)
3. Semi-automatic (manually programmed)
4. Manual
5. No test equipment.

Automatic test equipment (ATE) may be defined as equipment designed to automatically test functional parameters, to evaluate degree of performance degradation, and to perform fault isolation to the replaceable or repairable end item. The decision-making, control, and evaluate functions are conducted without reliance on human intervention. No *fully* automatic test system, in the sense that a human operator is not required either to initiate the test sequence or to interrupt the test, if necessary, exists today (Ref. 57). Some automatic checkout equipments do exist which require only operator initiation, particularly for on-line checkout and diagnostic applications, or which instruct the operator to perform certain other initiative, but noninterpretive, tasks during the test operation.

Semi-automatic test equipment is "any automatic testing device which requires human participation in the decision-making, control, or evaluation functions" (Ref. 60).

Semi-automatic test equipment may be either pre-programmed with the technician following previously prepared and sequenced instructions, as directed by the test equipment, or manually programmed in which the technician follows a printed test procedure or checklist and performs the control and switching function.

Manual *test equipment* is defined in MIL-STD-1309 as "test equipment that requires separate manipulations for each task (e.g., connection to signal to be measured, selection of suitable range, insertion of stimuli, and measurement and evaluation of results)" (Ref. 60).

Many of the advantages of automatic testing have been demonstrated time and again with regard to specific test systems. Generally speaking, they include more precise measurements, greater reliability in test results, more continuous surveillance, decreased human errors, decreased testing time, decreased maintenance training time, and lower cost for testing and maintenance training. While such advantages are normally associated with automatic testing, it does not necessarily follow that all ATE is more cost-effective to apply than manual test equipment. Development and acquisition costs of ATE are high. The development of ATE test procedures also requires considerable investment. It has been estimated that the cost of software generally is three or four times that of the cost of the ATE, itself. Moreover, there are numerous other design and economic considerations which should be addressed before determining the most cost-effective test equipment to utilize in a given situation (Ref. 57). Because ATE has become of increasing importance in system and equipment test and checkout in recent years, a more detailed discussion of it is given in par. 5-8.6.4.

No test equipment. In some cases, it has been shown that system effectiveness is optimized by adopting the test philosophy of no testing at the organizational level. This has been applied to some small missiles, and is called the "wooden round" concept.

5-8.6.3.2 Point of Application

Another way of classifying test equipment is by the point or level at which the equipment is applied. This may be at the factory production line, at the depot or repair facility, at general and direct support maintenance levels, or at the using organization.

At the factory level, use is made of commercial test equipment for production acceptance and quality assurance, including calibration and alignment and environmental tests, as described in par. 5-8.6.2. Similar or identical equipment can be used at the depot level of maintenance. In addition, special- or general-purpose test equipment can be utilized at the depot level (see par. 5-8.6.3.4).

At the general support and direct support levels, test equipment may be diagnostic or checkout equipment, and the requirements for the use of automatic test techniques increase (see par. 5-8.6.4). At the organizational level, test equipment requirements increasingly emphasize checkout and built-in test for rapid test, fault detection, and diagnosis down to the replaceable item level. At this level, detailed quantitative, interpretive test equipment is less desirable than a quick, goho-go checkout set.

5-8.6.3.3 Design Origin

Design origin, as a category of test equipment, is concerned with whether military or commercially designed test equipment is used. Military designed test equipment does not necessarily imply that the test equipment is designed by a military activity but rather implies that the test equipment is designed to military specifications and for use in military environment. Thus, the development cost of military-designed test equipment is generally borne by the Government in its entirety or at least in part. Commercially designed test equipment is designed by industry generally to meet testing requirements for commercial equipment. The development cost for commercially designed test equipment is borne by industry (Ref. **57**).

Commercial test equipment is used for factory testing and is usually well suited for depot testing as well. The acquisition cost for commercial test equipment is well below that of military test equipment (see par. **5-6.8.4.6**). Military test equipment, however, is more suitable for meeting the operating environment encountered in the field.

5-8.6.3.4 Versatility

The versatility of test equipment is a function of its ability to test more than one type of system or equipment, or to be readily modified to do so. It can be classified as either special-purpose or general-purpose test equipment.

Special-purpose test equipment is designed to test a specific system, subsystem, or module. It is designed for application to a specific prime equipment where the design, operational, and environmental characteristics are generally known before design of the test equipment. Special-purpose test equipment thus performs a more efficient and effective test by providing a better interface with the prime equipment for which it is designed as well as simpler operation and readout by lower skilled personnel. It is, therefore, less flexible and generally requires modification for use with other equipments than that for which it was designed.

As a result of the proliferation of complex, costly special test and checkout equipments and their lack of flexibility, there has been a significant effort in recent years to develop versatile test equipment utilizing the "building block" concept and modern programming and digital computer techniques (see par. **5-6.8.4.7**).

On the other hand, general-purpose test equipment is designed for application to a broad class or type of prime equipment, some of which may not yet be designed when the test equipment is developed. General-purpose test equipment consists of standard commercial or military test equipment-such as voltmeters, cathode-ray oscilloscopes, frequency and time measuring devices, counters, pressure gages, and leak detectors—which are designed for general parameter measurement without regard to the system or equipment to which they may be applied. A significant number of items listed in the Army Test, Measurement, and Diagnostic Equipment (TMDE) Register (Ref. 61) are of this category.

5-8.6.3.5 Interface With Prime Equipment

Test equipment can also be categorized according to its interface with the prime equipment being tested. This categorization includes built-in test equipment (BITE), on-line test equipment, off-line test equipment, and production test equipment.

Built-in test equipment is defined in MIL-STD-1309 as "any device permanently mounted in the prime equipment, and used for the express purpose of testing the prime equipment, either independently or in association with external test equipment" (Ref. 60). BITE generally indicates a goho-go situation and can be built into a small module or into an entire complex weapon system. BITE may provide continuous performance monitoring or it may scan various test points periodically.

On-line test equipment is any testing device which is separate from the UUT but which, when connected, tests the UUT in its operational environment. On-line test equipment does not generally require the use of external stimuli. It normally measures or samples the actual operating stimuli. Test equipment that is permanently installed in the prime equipment and can test various assemblies or subsystems through switching is considered in the category of BITE and not on-line test equipment. The basic difference is that a single unit of on-line test equipment can be used to test multiple copies of the UUT to which it is applicable, whereas BITE is physically constrained to test only one copy of the UUT for which it was designed (Ref. *57*).

Off-line test equipment is any testing device which tests a UUT after it has been removed from its operational environment. Thus, off-line testing indicates that some form of fault indication (perhaps resulting from BITE or on-line test) has already taken place since the UUT has been removed from its parent operating equipment. Since removal of the UUT is the only criterion, off-line testing can occur at any level of checkout or maintenance, but is generally used at direct support, general support, and depot levels rather than at the organizational level.

Production test equipment is off-line test equipment wherein the UUT is not tested in its operational environment. Production test equipment is used primarily to facilitate production line testing and quality control. In most cases, production test equipment can perform the same test functions required by off-line test equipment used for maintenance, as well as the more numerous, detailed tests not required in the field.

5-8.6.4 Automatic Test and Checkout Equipment

Automatic test and checkout equipment was originally developed in the early missile era to check a system prior to operation rather than to monitor it during operation. It was usually tailored for a particular system or equipment group. Since those days, automatic test equipment (ATE) has become more versatile and has been applied to on-line performance monitoring, off-line diagnosis, and rear echelon maintenance, in addition to specialized checkout applications. Each of the military services has its own lists of versatile, automatic test equipment (Ref. 57).

Employment of automatic test and checkout equipment not only reduces testing time and requirements for skilled manpower, but also makes possible the establishment of uniform, controlled, and reliable checkout procedures for determining operational readiness. In addition, such equipment is not restricted to the checkout of individual components, but can be employed efficiently for entire systems and subsystems.

For automatic test equipment to be effective it must have the following characteristics:

1. Automatic sequencing of test operations
2. Control stimuli for the system to be tested, if such stimuli are external to the system undergoing test
3. Capability of evaluating signals from a system in terms of acceptable tolerance values
4. Capability of making decisions (responses) on a positive, objective basis
5. Self-checking of its checkout features
6. Monitoring displays for operator use, where required
7. Production of a permanent record of test results, where required
8. Controls for rechecking sequences and bypassing portions of the test program
9. Automatic fault isolation to the throw-away level of the system under test
10. Simplicity of operation and maintenance
11. Minimal calibration and support requirements
12. Fail-safe circuitry (Ref. 42).

In a recent report entitled *Use of Automatic Test Equipment for Maintenance—A Reconnaissance* (Ref. 57), the Logistics Management Institute has noted the trends that follow in the development and use of ATE.

5-8.6.4.1 Rapidly Advancing Technology

Technological innovations are occurring at a rapid pace, both with respect to operating equipment and to automatic test equipment. Most electronic equipment designed today is constructed of micro-electronic circuits and with solid-state components. ATE not constructed to be capable of measuring performance parameters associated with such design characteristics is obsolete for most future applications. Technological advance in solid-state components is having a significant effect on the development of ATE. For example, the recent use of field-effect transistors provides a high sensitivity, low loading capability. Field-effect transistors, if used in ATE, allow for a natural sensing function of parameter measuring equipment and avoid the need for specially designed transistor circuits.

The technology with respect to UUT is changing significantly and thus, the design of future ATE must change to keep pace with UUT technology. Rapidly advancing technology in the field of electronics suggests that there is little to be gained by attempting to modify or standardize existing ATE, but much is to be gained by establishing more effective control and visibility over the development, acquisition, and use of future ATE.

5-8.6.4.2 Broader Scope of Application

Most ATE has been developed to test electronic modules and systems. However, technology now makes it possible to apply automatic testing techniques to mechanical, hydraulic, and electro-optical areas heretofore precluded. New measuring devices are being developed which will allow mechanical, hydraulic, and electro-optical characteristics to be translated into electronic signals which can then be transmitted to a programmer-comparator or a digital computer. The Army's automotive diagnostic test equipment development program is evidence of this trend.

This broader scope of application for ATE is not only significant with regard to off-line testing, but is also significant when considering a built-in test system. Built-in test equipment with a broad scope of application has the advantage of utilizing a single, dedicated digital computer to monitor an entire system. For example, the Malfunction Detection, Analysis and Recording Sub-system (MADARS) installed on the C5A Aircraft monitors about 1800 line-replaceable units, including both electronic and mechanical types.

5-8.6.4.3 Greater Emphasis on Equipment Degradation

The development of more sophisticated measuring devices has enabled test equipment designers to develop test equipment which will indicate degradation trends of system components in addition to goho-go test results. This capability can have a significant impact on maintenance concepts for future equipment and on the design of future ATE. Information regarding the degradation of system or component performance enables a scheduled overhaul or remove-and-replace operation to take place prior to actual failure, and can be useful in determining the extent of the repair operation and in facilitating preventive maintenance. For example, the Army has developed a need for Computer-controlled Automatic Test Equipment (CATE) for use during the 1975-1990 time period. This requirement describes CATE as a family of automatic test equipments that will provide all test capability necessary for direct, general; and depot support of all US Army electrical, electronic, electro-mechanical, and electro-optical materiel. A cost-effectiveness study compares the CATE concept with the use of present test equipment and methods and shows potential life cycle savings of approximately one billion dollars over a 15-year period by applying CATE in support of a two-corps field army (Ref. '62).

5-8.6.4.4 More Effective Self-testing Capabilities

The development of more complex ATE has required that a self-testing capability be built into the test equipment itself. A number of improvements have been made recently along those lines and most current generation ATE has self-testing capability. Future ATE design will automatically indicate to the operator when the test equipment is malfunctioning or requires calibration. In fact, some designers indicate that ATE of the future can be self-calibrating. In any event, ATE of the future should require less maintenance and provide greater confidence in the test results.

5-8.6.4.5 ATE Designs Compressed/ATE Application Broadened

While the latest ATE is being designed for broader test applications, the equipment itself is being compressed into smaller units. The decrease in test equipment size, weight, and power consumption is caused primarily by the use of micro-electronic circuits and transistorization of the test equipment, and by the use of smaller digital computers as programmer-comparators. Certain ATE under development today is aimed at providing testing capabilities for 85% to 95% of all equipment within broad categories, such as avionics and communications.

5-8.6.4.6 ATE Cost

There are indications that future ATE cost per UUT tested can be significantly decreased for a number of reasons. First, commercially developed ATE could be more widely used, particularly for off-line testing. In recent years commercial manufacturers have developed and built ATE systems for both commercial and military applications. The result is that a great deal of commercial off-the-shelf equipment is currently available to satisfy many military needs where the operational environment permits it, such as at depots or industrial facilities. The acquisition costs of such commercially available equipment are estimated to be somewhere between 10% and 20% of the acquisition costs for ATE designed to military specifications.

A second factor which would reduce the cost of future ATE is the decreased cost of digital computers which affects built-in test equipment as well as off-line ATE. With lower cost digital computers becoming available, a broader application of BITE is economically justified. Providing greater built-in test capability reduces the requirement for off-line testing.

A third reason for reduction of future ATE costs is the standardization of software. Much standardization has already been achieved with respect to commercial equipment; for example, most of the major commercial airlines have adopted a common computer language for automatic test equipment applications. This allows test programs for common equipment to be shared by all airlines. Standardization of test procedures and computer programs has not yet met with wide success in military applications, but some efforts have been directed toward standardizing a common computer language.

5-8.6.4.7 Standardization of ATE

A number of studies have been undertaken by the military departments concerning the standardization of ATE. There are two areas of standardization that would contribute to the reduction of the number of future ATE's. The first is the standardization of ATE, itself, and the second is the standardization of the UUT which the ATE is intended to test. A recent Air Force study which examined 41 different ATE's for avionics indicated that many test parameters built into a large scale ATE are not being utilized at all or are utilized very little (Ref. 63). This Air Force study proposes the design of 62 standard "building blocks" from which most specific test requirements for avionics can be satisfied. The building block concept is to design the test system so that it consits of a number of functional modules, any of which, can be eliminated when not required, without affecting the operation of the system as a whole. This concept allows an automatic test system at any specific installation to be configured to meet the unique requirements of that installation. As the state of the art advances, new modules can be designed and incorporated without redesigning the basic test system. The building block concept has been applied to some extent by all three military departmegts and to an even greater extent by certain commercial manufacturers. Standardization potentials are also achievable for off-line special-purpose ATE. An example in this area is the standardization of "Spectrometric Oil Analysis Equipment" used by all three military departments (see par. 5-8.6.2). By standardization of oil analysis equipment, a reduction in the number of oil analysis laboratories, and the establishment of a new oil analysis program management, the Department of Defense estimates savings of some $5.3 million in planned equipment costs and $18.1 million a year in operating costs.

The Navy has made some significant progress with regard to the standardization of electronic circuit modules. It was found that the increasing use of microelectronic circuits made it possible to develop basic, functionally oriented circuits as modular assemblies which could be plugged in as needed to satisfy a variety of requirements for larger electronic assemblies. Thus, the Navy established the Standard Hardware Program (SHP) to develop standard electronic modules at the functional circuit card level. There are approximately 105 different standard modules in 19 general categories. In addition, there are approximately 86 special modules which could become candidates for standard modules. The SHP program is a dynamic one, developing new modules as the state of the art advances and deleting those which become obsolete. The Navy reports that there are over thirty systems which have voluntarily used SHP modules in their design. The Poseidon Fire Control System, for example, uses approximately 10,000 SHP modules, 85% of which consist of only 12 different modules.

The interface between standardization of ATE and standardization of prime equipment served by ATE is important to recognize. This interface highlights the need to consider test equipment requirements during the design process of prime equipment, and conversely to consider design features of prime equipment during the test equipment design process.

5-8.6.4.8 ATE Policies

There are currently no overall Department of Defense policies aimed specifically at the development and use of ATE. The Army's principal policy, set forth in AR 750-43, ***Maintenance of Supplies and Equipment: Test, Measurement, and Diagnostic Equipment (TMDE)*** (Ref. 64), establishes broad policies covering the development, purchase, acquisition, and use of all test, measurement, diagnostic, prognostic, and calibration equipment. In addition to prescribing policies, Ref. 64 establishes objectives and priorities, and assigns responsibilities for life cycle management of TMDE. The scope of application is broad inasmuch as it spans all TMDE, all Army programs for maintenance, research and development, procurement/production, investment, operating expense, and standardization.

The Army's TMDE Register (Ref. 61) lists both equipments which are operational and those under development. Most of the equipments, both operational and under development, are manual rather than ATE. In addition to the test equipments documented in the Army's TMDE Register, there are a number of items of ATE for which the Army has developed requirements. An example of a multisystem application of automatic test equipment is the Land Combat Support System (LCSS) (Refs. 65 and 66). LCSS is an Army diagnostic field maintenance system designed to support the SHILLELAGH, TOW, LANCE, and DRAGON weapon systems. The Army has also developed depot test equipment for large families of high population communication equipment. These equipments, DEE (Digital Evaluation Equipment) and DIMATE (Depot Installed Maintenance Automatic Test Equipment), (Refs. 67 and 68), although designed for full diagnostic capability, are primarily used for performance testing because of the high workload at the depots and the dramatic time savings and increased confidence in testing realized. Over 300 programs have been prepared for these equipments which have been used for tens of thousands of tests in depots and the field.

In addition to these Army programs, the US Navy and the US Air Force have made heavy commitments to their VAST (Versatile Avionic Shop Tester) and GPATS (General Purpose Automatic Test System). The overall result of the military services automated electronic test project has been a technology base which has not only generated interest in industrial electronics, but in other commodity areas in the Government. For example, all of the military services currently have development programs underway or under consideration for automatic test of gas turbine engines and turbine engine components.

5-8.6.5 Diagnostic Techniques

The term "diagnostics" refers to actions required for actual location of a fault in an operational system; it is better known as "troubleshooting." The primary objective of the maintainability engineer in the field of diagnostics is an overall reduction of system downtime by providing for the rapid location and isolation of faults. Different diagnostic techniques are available as follows:

1. ***Manual techniques*** (which are the type most frequently referred to as troubleshooting) are basically trial-and-error efforts by skilled technicians who use meters, oscilloscopes, and other devices, as well as detailed procedures and schematics, to isolate a malfunctioning component by progressively testing all components and eliminating those that are still functioning. These techniques require that external test equipment be used, and, in the great majority of instances, that decisions be made by the technicians.

2. ***Semi-automatic techniques,*** as the term suggests, represent one or more steps toward automation of the isolating of faults. In all instances, they fall short of complete elimination of dependence on direct participation by technicians. Techniques of this type involve a sufficient number of internal test units and indicators to make any decision by the technician unnecessary. The indicators either identify the subsystem, cabinet, chassis, or other component in which a malfunction exists, or they direct the technician to the next action to be taken. A semi-automatic technique is characterized by automated decision-making and by partial automation of test units and procedures.

3. ***Automatic techniques*** completely eliminate the need for technician participation in locating a fault. Upon failure of a component, a system fitted with automatic techniques switches to a diagnostic mode, and by means of internal circuitry, isolates and identifies the malfunctioning item to the repair-by-replacement level.

Recent trends toward both standardization and modularization have accelerated the development and employment of automatic and semi-automatic diagnostic techniques. Inasmuch as both types are used to allow systems to be maintained by unskilled technicians, they materially reduce overall system requirements (Ref. 42).

In the absence of automatic diagnostic equipment in the Army maintenance system, the practice of "diagnosis through part replacement" has become a problem. The result of this practice is a high rate of incorrect diagnosis. For example, an Army study in 1966 found that faulty engine diagnosis on tracked vehicles occurred 34% of the time, on wheeled vehicles 20% of the time, and on fuel and electrical components 47% of the time. It has also been noted that with the present maintenance tools, 60% to 90% of the maintenance repair time is usually spent in determining the cause and location of a malfunction (Ref. 69).

Examples of the application of diagnostic techniques to Army systems and equipments, specifically of automotive diagnostic equipment, is given in Refs. 69, 70, and 71.

The diagnostic equipment for the depot or factory rebuild vehicle operation is referred to as the depot Multipurpose Automatic Inspection Diagnostic Systems (MAIDS). This equipment differs from that of the other maintenance levels previously described in that it is not used on the complete vehicle. Its current purpose is to diagnose and checkout engines and transmissions which have been removed from the vehicle and returned to the depot for overhaul. It is used in conjunction with conventional engine dynamometer test facilities. The depot MAIDS represents a pioneering effort based on the approach that diagnosing high cost engines and transmissions ***before*** repair and then correcting only those items needing repair is economical. Depot MAIDS identifies and isolates individual engine/transmission conditions through a process of automatic analysis. Resulting overhaul work is directed to the correction of specific malfunctions with a consequent increase in productivity and cost saving (Refs. 70 and 71).

5-8.6.6 Test and Checkout Design Considerations

Maintainability design considerations with respect to test and checkout fall into two categories. The first is concerned with test philosophy, the second with specific test equipment design characteristics.

5-8.6.6.1 Test Philosophy

The considerations that follow constitute the general philosophy of test and checkout. They should be kept in mind by system designers, and maintainability and test equipment engineers as design guidelines.

1. At the onset of a program, the test philosophy should be to hold system checkout tests to a minimum, preferably to test nothing. System parameters should be carefully examined for criticality, and a justification for each test proposed should be required rather than the early adoption of a policy to test everything unless it can later be shown which tests should be eliminated. The "later" seldom happens, and the resulting test equipment usually turns out to be overly complex with consequent poor reliability and maintainability.

2. Test tolerances should get broader as testing proceeds from rear areas to forward using areas (factory to depot to field). If close test tolerances are re-

quired at forward areas, then there is a very strong implication that the system design is marginal and that the system will not achieve the required operational availability.

3. In order to prevent the UUT from circulating back and forth from using activity to repair facility because the user test says the UUT is faulty and the maintenance facility says that it is not, it is desirable that each successive rear echelon maintenance level be equipped to perform the same tests to the same tolerances and with the same test methods as its adjacent forward level, preferably using the same test equipment. If possible, this should be in addition to the more detailed, tighter tolerance testing required at the rear echelon concerned.

4. Testing at forward areas should be limited to those parameters which are essential to checking overall system or equipment performance, preferably on a quantitative, noninterpretive basis. This does not necessarily mean the use of go/no-go lights. If a different type of indication, such as LO-GO-HI meters and CRT with tolerance masks, gives both go/no-go indications as well as indicating trend levels, it will assist the operator or maintenance technician in monitoring system degradation or noting anomalies which might help in fault diagnosis.

5. To optimize test and maintenance, the system design should use "line replaceable units" (LRU's) which can be replaced at the organizational and direct support levels using only rudimentary fault isolation techniques. An LRU is a functional module or set of functional modules which, when replaced as an entity, require no adjustment of other system modules, although fine adjustment of the newly inserted LRU is allowable to optimize its performance in the system. If other modules must be adjusted, they should be considered as part of the LRU.

6. To obtain the highest operational availability, simplicity in both prime system and test equipment design and performance must be considered as a governing criterion. It is a fact that "simplicity means reliability".

7. ATE should not be usedjust for sake of automation. Its use must be justified on system cost-effectiveness considerations only. Experience with early use of ATE has shown that, in a high percentage of cases, ATE is more complex and costly and is often unreliable and hard to maintain, resulting in requirements for highly skilled personnel and lower system availability.

8. For system checkout, it is desirable to let the system being tested introduce as many of its own test stimuli as possible. The system will then be exercised in a manner which represents as closely as possible its operational parameter values, and the test outcomes will be more indicative of true system performance. Although this is often difficult to implement, it should be given consideration during the design of system test and checkout. The use of substitute standard black boxes or simulators should be minimized.

9. The confidence of the operator in the test equipment cannot be overlooked. If he has no confidence in it, he will not use it. If it turns out that indications of failed performance are found to be due to the test equipment rather than the UUT, he will not use it. If the technician becomes the servant of the test equipment rather than the test equipment the tool of the technician, he will avoid its use. Although often said in jest, the question of whether the test equipment is testing the system or vice versa is vitally important.

10. The cost of test equipment may become comparable with or may even exceed the cost of the system which it is to test, especially if ATE is used. To justify such cost, the operational availability of the system must be commensurately better.

11. A system which contains an undetected failure may be more harmful to mission success than one which is down for maintenance.

12. The ability of the prime system to perform secondary or alternative missions with reduced performance must be considered. The test philosophy and test equipment design should be able to give such status information to the operational commander.

13. The ability of the test and checkout equipment to provide feedback data for analysis in order to allow product improvement, trend analysis, and improved maintenance and logistic planning should be considered.

14. The test and checkout philosophy must be consistent with and derived from overall system operational and logistic policy and must be part of the maintenance concept. It must provide for fault isolation and repair/replace to the proper system level.

15. The test equipment must not "wear out" the functional system.

Checkout of a system only indicates that the system was operating properly at the time the test was made. It does not guarantee that the system will operate immediately thereafter; rather it gives a degree of confidence that the system will operate properly.

5-8.6.6.2 Test Equipment Design Characteristics

It is essential that extreme care be given to the maintainability design features in the design of both the prime system and its test and support equipment, as indicated in par. 5-8.6. Such design considerations must start as soon as the prime system design is started. Test philisophy considerations and test equipment design policies must be rigorously monitored and enforced by Government technical and contract personnel. It may well turn out that the system overall effectiveness will be influenced as much, and in some cases even more, by giving adequate consideration to test and maintainability design features than by improvement in component and equipment reliability.

Some of the maintainability design considerations with respect to test and checkout are:

1. Use modular design. Design the prime equipment and test equipment using Line Replaceable Units (LRU's).

2. Use standard and proven circuits and components.

3. Provide for built-in test points and test features, including self-check, to the greatest extent feasible. Self-check features should promote a high degree of confidence in the capability of the equipment to indicate a fault in the system being tested, rather than in the checkout equipment itself.

4. Where BITE is not feasible, use standard, general-purpose rather than special-purpose test equipment.

5. Require logically presented maintenance instructions and diagrams in technical manuals and maintenance handbooks. Use of techniques such as Design Disclosure Format (DDF), logic diagrams, fault trees, and half-split should be investigated.

6. Provide test features and procedures which are consistent with the maintenance level and environmental requirements.

7. The man-machine relationship is important in the design and utilization of test and checkout equipment. The man is part of the loop.

8. Safety precautions and safety design features must be emphasized. Fail-safe design should be employed. Adequate protection should be given to both operators and equipment; this should include shielded interlocks and properly placed warning signs.

9. Test design should consider flexibility in order to accommodate changes in the prime system and equipment design as well as the ability to handle more than one prime system or equipment. The building block technique should be used wherever possible, particularly for ATE.

10. Particular attention should be paid to displays. Go/no-go indicators should be used when noninterpretive readout is desired, as at the using level. Templates, overlays, meters, CRT's, and other interpretive displays should be provided for maintenance personnel when determination of system degradation, marginal performance, and prognostic information is desired.

Additional design information and checklists applicable to test and checkout are given in Ref. II.

5-8.7 Trade-offs

A trade-off is a rational selection among alternatives in order to optimize some system parameter that is a function of two or more variables which are being compared (traded *off*). Trade-offs involve performance, cost, schedule, and risk. A trade-off may be quantitative or qualitative. Insofar as possible, it is desirable that trade-offs be based on quantifiable, analytic, or empirical relationships. Where this is not possible, then semi-quantitative methods using ordinal rankings or weighting factors are often used.

The trade-off concept is not new. Each of us makes use of trade-offs in the daily routine of living. Every decision we make is based on judgment, which in its turn represents the weighing of known facts. The greater the number of facts, provided they are dealt with rationally and systematically, the greater is the probability of arriving at a correct decision.

Trade-off techniques, by providing the objectivity and systematic approach required, contribute greatly to the validity of maintainability design decisions.

Formally developed trade-off studies are needed at every stage of the design and development of new systems. In the planning phases, they determine the feasibility of a program. Military system requirements must be analyzed and weighed in terms of such factors as state of the art, development time required, total cost, and extent to which off-the-shelf hardware can be used; for civilian systems, such factors as potential consumer demand, company capabilities, and profit margin are analyzed and weighed. After the feasibility study has produced a positive finding and design work begins, trade-off techniques are applied to such problems as determining the relative advantage of various system concepts, throw-away-at-failure vs piece-part repair, different packaging concepts, and, at every level, variant specific design features. Trade-offs also play a primary role in decision making during design review, when the diverse interests of several project objectives must be reconciled (Ref. 42).

The methodology for structuring and performing trade-off analyses is found in the system design decision process described in par. 5-6. The design decision model is shown in Fig. 5-5. The basic steps, summarized here, are:

1. Define the trade-off problem and establish the trade-off criteria and constraints.

2. Synthesize alternative design configurations.

3. Analyze these alternative configurations.

4. Evaluate the results of the analyses with respect to the criteria, eliminating those which violate constraint boundaries.

5. Select the alternative which best meets criteria and constraint boundaries, or iterate the design alternatives, repeating Steps 2 through 5 to obtain improved solutions.

These steps are described in more detail in pars. 5-6 and 7-6.2.1.

Trade-offs concerned with maintainability range from the relatively simple reliability-maintainability-availability trade-off to complex issues which concern such items as the level of automation of test and checkout, selection of module size, repair level analysis, and trade-offs concerning skill level, performance and packaging requirements, human factors, and the operational and support environment.

Among the list of possible trade-off considerations which concern maintainability are the following:

1. ***Reliability vs maintainability.*** This trade-off is concerned with whether a given availability requirement can better be met through increased emphasis on reliadility features or on maintainability features, and to obtain a cost-effective balance between them (see pars. 5-8.7.1 and 7-6.1).

2. ***Repair level.*** This trade-off is concerned with the determination of the cost-effective allocation of maintenance tasks to the various maintenance levels---organizational, direct support, general support, and depot.

3. ***Repair/replace/discard.*** This trade-off is concerned with determining and applying economic decision criteria and rules for making decisions as to whether a failed item should be repaired or discarded. It is also concerned with whether, if repairable, the failed item should be repaired in place at the equipment site or removed from the equipment, replaced by a good item, and then repaired at the same or a rear level (see pars. 5-8.7.4 and 7-6.2.5).

4. ***System level.*** This trade-off is concerned with the system level—e.g., equipment, unit, module, component, or part—at which repair or maintenance should be performed.

5. ***Corrective vs preventive maintenance.*** This trade-off is concerned with the decision as to whether a particular failure should be inhibited by means of preventive maintenance action or whether it should be left as a corrective maintenance action upon failure. Among other considerations, this trade-off is a function of failure modes, effects, and criticality, failure frequency, failure distribution, and mission and operational requirements.

6. ***Level of automation.*** This trade-off decision is one concerned with the level to which built-in and automatic vs general-purpose and manual test and checkout features should be incorporated into system and equipment design in order to meet system performance and effectiveness requirements. It concerns availability, skill levels, complexity, criticality, monitoring, number of tests, test time, and economics, among others (see Ref. 11, Chapter 5).

7. ***Packaging.*** This is concerned with trade-off, between standardized and nonstandardized components and modules, and the extent to which modularization will be utilized. It affects provisioning and repair parts (see par. 5-8.2).

8. ***Human Factors.*** These are man vs machine trade-offs dealing with which tasks would be better performed by technicians and which by machine (see par. 5-8.4).

9. ***Downtime Allocations.*** These are trade-offs made among maintainability design factors which will allow the apportionment of downtime to the various repair time elements on a cost-effective basis. Such trade-offs are concerned with maintainability prediction and allocation as well as decisions regarding the various maintainability design factors.

10. ***Cost.*** Almost all trade-offs are based on economic criteria in addition to performance and effectiveness. Distinct trade-offs may be made between acquisition costs and operation and support costs as part of total system life-cycle costs. This may include trade-offs made during the early system life cycle planning and development phases in determining maintenance and support concepts, reliability and maintainability features, and repair level and repair/discard criteria. It also includes trade-offs made during the operation and support phases as part of maintenance support decisions, such as whether a normally repairable item is still worth repairing after sustaining damage or after a significant part of its useful life has expired (see Chapter 7).

There are other trade-off considerations which impact maintainability decisions. Those listed, however, are representative and important to system maintainability, and design engineers. All of the design consideration discussed in this chapter are part of the maintainability engineer's concern with respect to trade-offs. Some techniques and examples of trade-offs are given in the remaining paragraphs of this chapter and in Chapter 7.

5-8.7.1 Reliability-Maintainability-Availability Trade-off

The reliability-maintainability-availability trade-off is an example of an analytic, quantitative trade-off. The system parameter availability A is a function of the variables of reliability (*MTBF*) and maintainability (MTTR), as given in Eq. 1-23 and repeated here

$$A = MTBF/(MTBF + MTTR) \quad (5\text{-}2)$$

Since MTBF $= 1/\lambda$ where h is the failure rate and MTTR $= 1/\mu$ where μ is the repair rate, Eq. 5-2 may be rewritten as

$$A = \mu/(\mu + \lambda) \quad (5\text{-}3)$$

or

$$A = 1/[1 + \lambda/\mu] \quad (5\text{-}4)$$

A generalized plot of Eqs. 5-2 and 5-3 is given in Fig. 5-17. A plot of Eq. 5-4 A vs λ/μ, is given in Fig. 5-18. These equations and graphs show that, in order to optimize availability, it is desirable to make the ratio of *MTBF/MTTR* as high as possible.

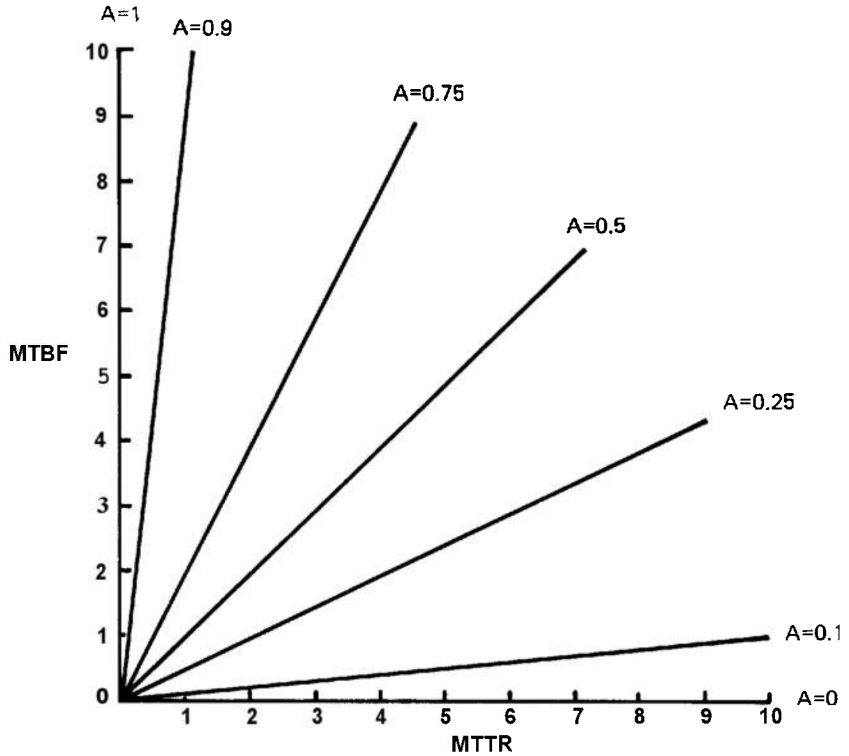
Since increasing MTBF and decreasing MTTR is desirable, the equation for availability can be plotted in terms of MTBF and $1/MTTR$ (or μ) as shown in Fig. 5-19. Each of the curves representing the same availability in Fig. 5-19, just as each of the lines in Fig. 5-17 is called ***isoavailability*** contours; corresponding values of MTBF and MTTR give the same value of A, all other things being equal. A more complete discussion of this is given in par. 7-6.1 and in Ref. 12. Availability and dependability nomographs useful for reliability and maintainability trade-offs are given in Chapter 2 and Refs. 32 and 72.

There are obvious practical limits which must be considered in trade-off optimization. These are called ***constraints,*** and all purposeful optimization must be bounded by constraints into ***feasible regions*** For example, there are practical limits as to how high a value for *MTBF* can be achieved or how low MTTR can be made. In the one case, the reliability of system components or the required redundancy might be so high that the desired reliability could not be realistically achieved within the state-of-the-art or would be so expensive as to violate cost constraints. Similarly, MTTRs close to zero would require extreme maintainability design features, such as completely built-in test features or automatic test and checkout to allow fault isolation to each individual replaceable module, with perhaps automatic switchover from a failed item to a standby item. This also could easily violate state-of-the-art or cost constraints.

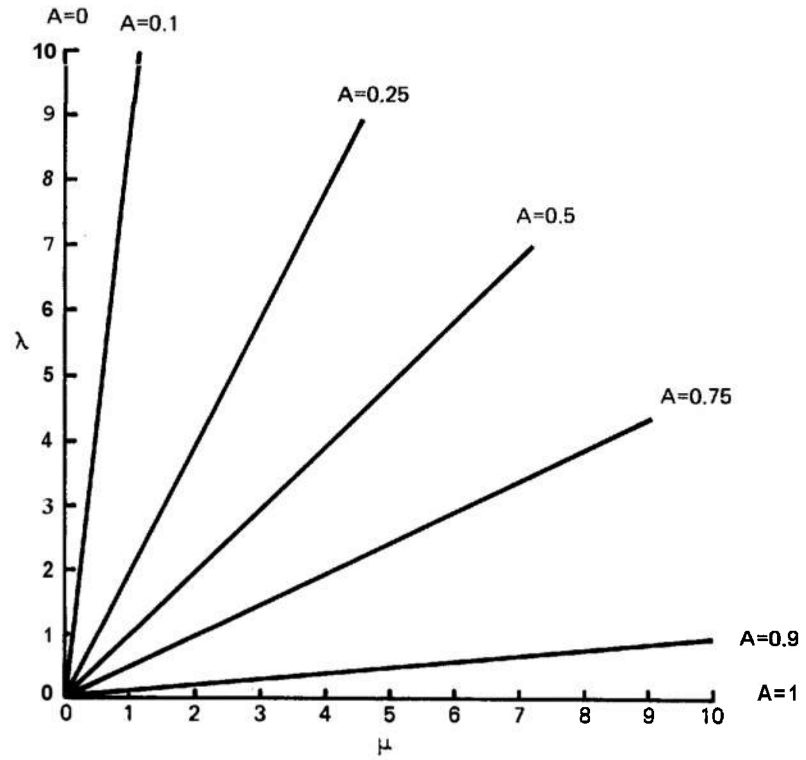
It follows, then, that trade-offs not only involve relationships among system parameters and variables, but they are bounded by both technical and economic constraints. In a sense, all trade-offs are economic ones, requiring cost-benefit analysis (not necessarily in terms of dollar costs, but rather in terms of the availability and consumption of resources, of which dollars are often the most convenient measure). Resource constraints may also include manpower and skill levels, schedule or time availability, and the technical state-of-the-art capability. Chapter 7 deals with the cost problem.

There are two general classes of trade-offs. In the first, the contributing system variables are traded off against one another without increasing the value of the higher level system parameter, for example, trading off reliability and maintainability along an isoavailability contour (no change in availability). This might be done for reasons of standardization or safety, or for operational reasons such as the level at which the system and its equipments will be maintained. The other class of trade-off is one in which the system variables are varied in order to obtain the highest value of the related system parameters within cost or other constraints. For example, reliability and maintainability might be traded off in order to achieve a higher availability. This could result in moving from one isoavailability curve to another in Fig. 5-19, perhaps along an ***isocline*** (a line connecting equal slopes) (see Ref. 12 and Chapter 7). Chapter 5 of Ref. 11 contains another system availability trade-off example.

An example of a reliability-maintainability-availability trade-off illustrating the preceding concepts is given. It is taken from Example 6 of Ref. 73. The design problem is as follows: A requirement exists to design a radar receiver which will meet an inherent availability of 0.990, a minimum MTBF of 200 hr, and an MTTR not to exceed 4.0 hr. Existing design with the use of



(A) AVAILABILITY AS A FUNCTION OF MTBF AND MTTR



(B) AVAILABILITY AS A FUNCTION OF λ AND μ

Figure 5-17. Reliability-Maintainability-Availability Relationships

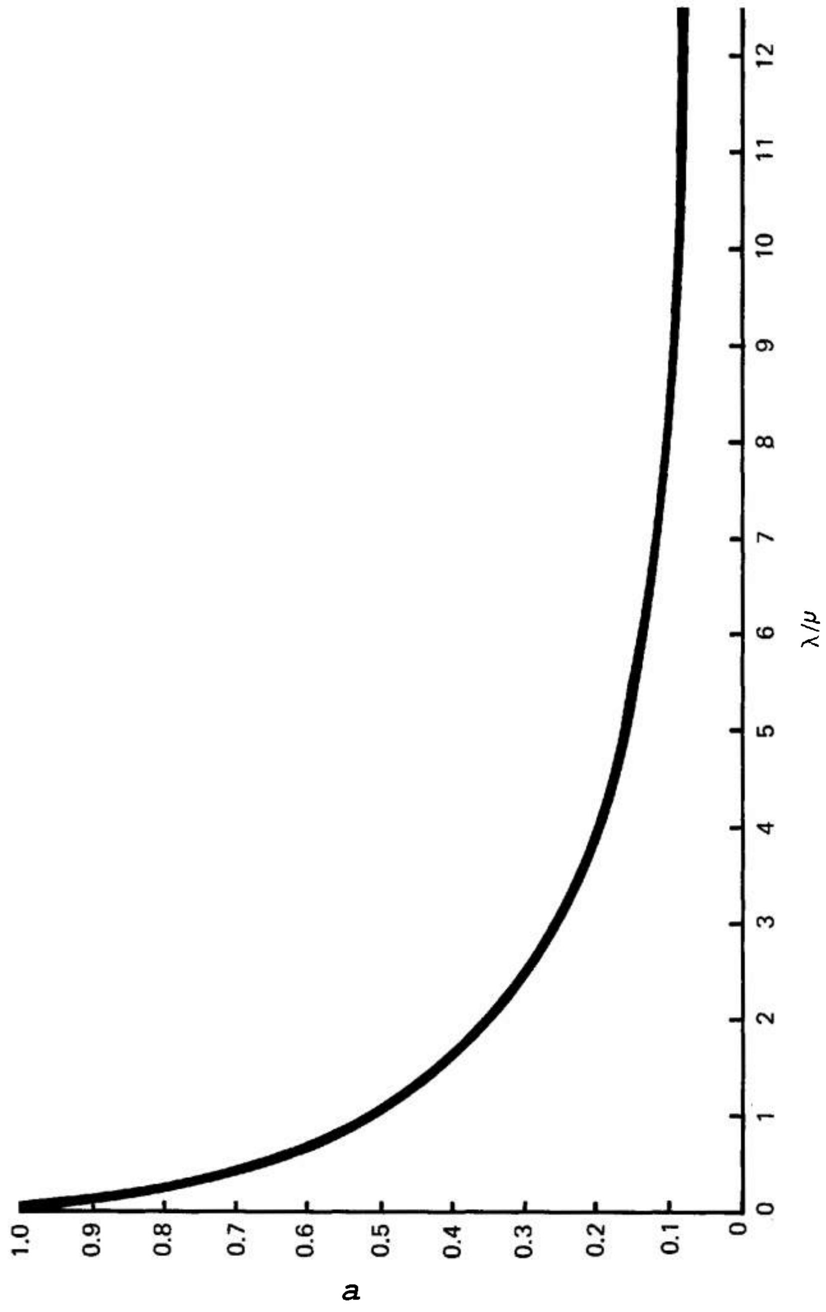


Figure 5-18. Availability as a Function of λ/μ

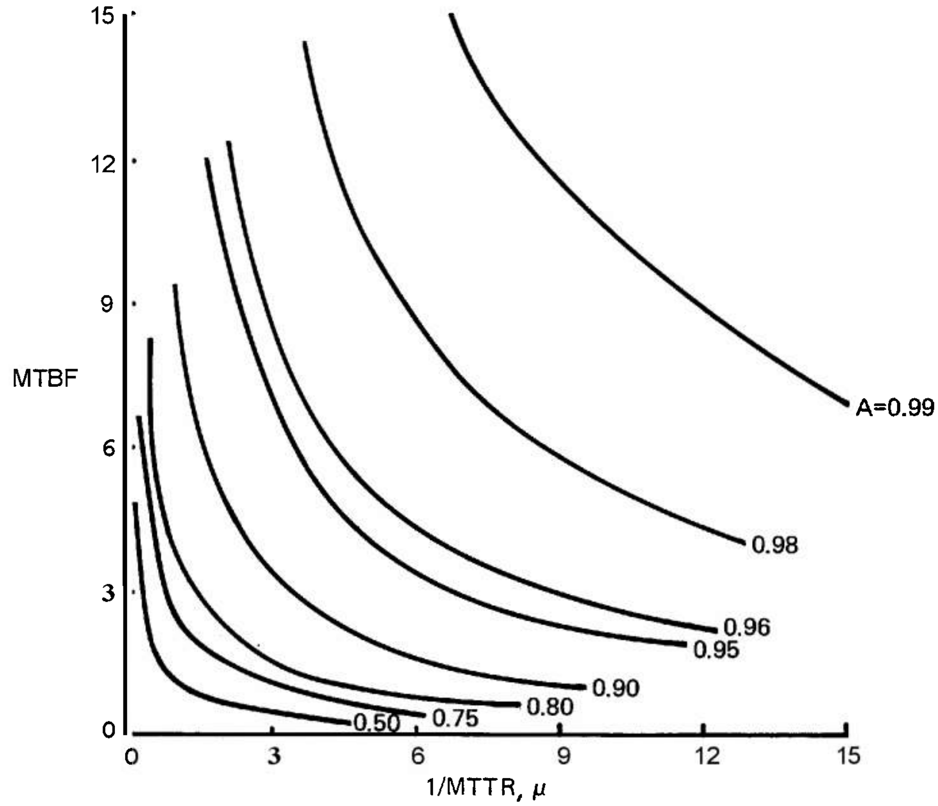


Figure 5-19. Availability as a Function of *MTBF* and 1/*MTTR*

TABLE 5-11.
ALTERNATIVE DESIGN TRADE-OFF CONFIGURATIONS

Design Configuration	*A*	MTBF, hr	MTTR, hr
1. R — derating of military standard parts. *M* — modularization and automatic testing.	0.990	200	*2.0*
2. R — design includes high reliability parts/ components. *M* — limited modularization and semi-automatic testing.	0.990	300	3.0
3. R — design includes partial redundancy. *M* — manual testing and limited modularization.	0.990	350	3.5
4. *R* — design includes high reliability parts/ components. *M* — modularization and automatic testing.	0.993	300	2.0

Military Standard parts meets an availability of 0.97, an MTBF of 150 hr, and an MTTR of **4.64** hr.

Using *Eq.* 5-2 and an expanded version of Fig. 5-15, the area within which the allowable trade-off may be made is shown by the cross-hatched portion of Fig. 2-15. The capability of the present system is also shown in Fig. 2-15. As indicated in the previous paragraph, there are two approaches which can be used for the trade-off. One is to fix the availability at 0.990. This means that any combination of MTBF and MTTR between the two allowable end points on the 0.990 isoavailability line may be chosen. These lie between an MTBF of 200 hr with an MTTR of 2 hr, and an MTBF of 400 hr with an MTTR of **4** hr. The other approach is to allow availability to be larger than 0.990, and thus allow any combination of MTBF and MTTR within the feasible region.

It is clearly seen that, without any additional constraints, the designer has a limitless number of combinations from which to choose. Assume that the following four alternative design configurations have been selected for trade-off as shown in Table 5-1**1.**

Design Configuration Nos. 1, 2, and 3 all have the required availability of 0.990. Design Configuration No. 1 emphasizes the maintainability aspects in design while Design Configuration No. 3 stresses reliability improvement. Design Configuration No. 2 is between Nos. 1 and 3 for the same availability. Design Configuration No. **4** is a combination of Nos. 1 and 2, and yields a higher availability. The values are plotted on Fig. 2-15.

Since all of these alternatives are within the feasible region shown in Fig. 2-15, some other criterion must be used for selection of the desired configuration. In this case, we will use the least cost alternative, or the one showing the greatest life cycle cost savings over the present configuration as the basis for trade-off decision. A cost comparison of the different alternatives is shown in Table 5-12.

The cost table shows that Configuration No. 2 is the lowest cost alternative among those with equal availabilities. It also shows that Configuration No. **4,** with a higher acquisition cost, has a significantly better 10-year life-cycle support cost, and lowest overall cost, as well as a higher availability. Thus Configuration No. **4** is the optimum trade-off, containing both improved reliability and maintainability features.

5-8.7.2 The NSIA Trade-off Technique

Another example of the semi-quantitative Trade-off technique similar to the one in the previous paragraph is the National Security Industrial Association (NSIA) maintainability trade-off technique (Refs. **42, 73,** and **74).**

The NSIA technique enables designers to determine quickly, and with reasonable accuracy, which of several alternative maintainability considerations should be adopted to govern the design work to be done. This technique has a major limitation in that it does not require all low-order elements to be analyzed and evaluated before a design decision is arrived at; from the viewpoint of the analyst, this reduces somewhat the accuracy of its findings. Despite this weakness, the NSIA technique is superior to a qualitative estimate of the relative desirability of design features. It can be used most advantageously when either time or manpower limitations preclude the application of more sophisticated techniques.

The NSIA technique considers each proposed variant of each individual design feature of a system or subassembly and then evaluates it in terms of its effect on any and all of the characteristics and other design features of the system being designed. Each such evaluation, which produces a finding expressed as a numerical value, is made by the individuals or groups directly concerned with the various characteristics of the system that will be affected by the feature under consideration. For example, a reliability engineering group would evaluate the effect of a proposed maintainability design feature on the reliability of the subassembly to which it would be applied or on the overall reliability of the system itself.

The individual or group who evaluates the effect of a particular design feature can and should extend the evaluation to all aspects of the problem that would possibly be involved. For example, it may be desirable, when evaluating a certain design feature, to investigate such improvements in operational reliability that may result from more frequent inspections, adjustments, cleaning, etc., than normally would be made. When this is done by trade-off, it is possible to refine the balance of the favorable and unfavorable effects of the proposed design on a major characteristic of the system being planned—in this case, its reliability. The total effect on system reliability thus determined is expressed as a numerical value. This value is combined with similar measures of the effect of other features of the

TABLE 5-12.
COST COMPARISON OF ALTERNATIVE DESIGN CONFIGURATIONS

ITEM	CONFIGURATION Existing	1	2	3	4
Acquisition Costs **(Thousands of Dollars)**					
RDT&E	300	325	319	322	330
Production	4,500	4,534	4,525	4,530	4,542
Total	4,800	4,859	4,844	4,852	4,872
10-Year Support Costs **(Thousands of Dollars)**					
Spares	210	151	105	90	105
Repair	1,297	346	382	405	346
Trng & Manuals	20	14	16	18	14
Provisioning & Handling	475	525	503	505	503
Total	2,002	1,036	1,006	1,018	968
Life-Cycle Cost	6,802	5,895	5,850	5,870	5,840

proposed design. The final result obtained becomes an objective basis for judging the desirability of adopting the design feature that has been so analyzed.

This technique obtains positive or negative numerical values for various parameters and aspects of influence—based upon a determination by a qualified individual or a group—as to whether a particular design change under consideration shall have a favorable or unfavorable effect upon the total equipment or system, as viewed from one particular area of interest. The evaluator is also required to associate a numerical value between 0 and 100 depending upon the degree to which it is considered favorable or unfavorable, indicated by positive and negative values, respectively. The value of − 100 **is** applied when a change is considered completely unacceptable, and + 100 is the value applied when it is considered absolutely necessary. Zero is the value applied when it is considered that the advantages resulting from the change balance out the disadvantages. All other values fall somewhere between these extremes. The basic rating scale is shown in Fig. 5-20. All evaluations are then combined algebraically to arrive at an overall value to be used in making a final decision.

There are certain areas that require special consideration because of the fact that the numerical values established are based on personal opinion and judgment. Any method or technique in which personal opinion or experience is used in evaluating the relative merits or demerits of an item under consideration inherently includes the possibility of subjective bias being incorporated into the final decision. While this can be considered as an inherent deficiency, it need not invalidate the method if certain precautions are taken to minimize possible adverse effects.

To reduce or minimize any bias that may be introduced through subjective evaluation, the following precautions should be considered when making a maintainability evaluation:

1. Evaluations should only be made by individuals qualified to do so, i.e., experts.

2. Evaluations of a single area of consideration should be made, whenever possible, by more than one expert on an independent basis, and the algebraic average of all evaluations used. If evaluations cannot **be** made independently, they should be made on a group basis. The larger the number of qualified evaluators that comprise the group, the more accurate and unbiased the final evaluation should be.

3. Any bias that might be introduced by the opinion of an individual, or a group, is modified in its effect upon the final value because it is only one of several other factors. It is necessary therefore that all possible areas of influence be listed as parameters for consideration and evaluation, and that each parameter be evaluated with respect to all possible areas of influence to maximize this effect (Ref. 11).

It should be noted that the results of this type of analysis can usually be expected to give a reasonably clear and conclusive indication of which particular design feature of several under consideration **is** most desirable, and to what degree. If the results of applying this technique do not provide this clear indication, then personal subjectivity may possibly be a significant factor. In such cases, all values developed during the initial evaluation should be carefully reviewed before making a final decision to be sure that it is based upon reasonable objectivity.

One other area in which personal opinion could possibly contribute to inaccuracy in the final result is in the determination of weighting factors. Here again the opinion of a group of interested and informed people is most desirable since it is less subject to personal bias than the opinion of a single individual. In general this part of the technique should be approached with the same general precautions recommended for the evaluation of the various aspects and parameters. It also stands to reason that every effort should be made to clearly describe the change so that a uniform and accurate understanding is conveyed to all evaluators.

In recent years, a number of studies have been made in the area of quantifying subjective judgment. These studies have included the application of probability and statistics, Bayesian analysis, psychophysics, and modern decision and utility theory. The Delphi technique developed by the **RAND** Corporation is one of these whose purpose is to quantify expert judgment. It is a means of quantifying values obtained from group opinion (Refs. 75 and 76).

Weighting schemes as discussed and **as** used in the NSIA trade-off technique usually assume a linear relationship between some design or other system attribute and its value or contribution to system worth. In practice, this is seldom linear. For example, the concept of diminishing marginal utility in economics indicates that the more one has of a resource the less the next increment is worth. More sophisticated weighting schemes are based on this concept. Ref. **77** describes how this can be applied to trade-off analysis. **A** step-by-step description of the application of the NSIA trade-off technique is given in Ref. **42.**

The example of the NSIA trade-off technique, taken from Refs. **73** and **74,** is presented.

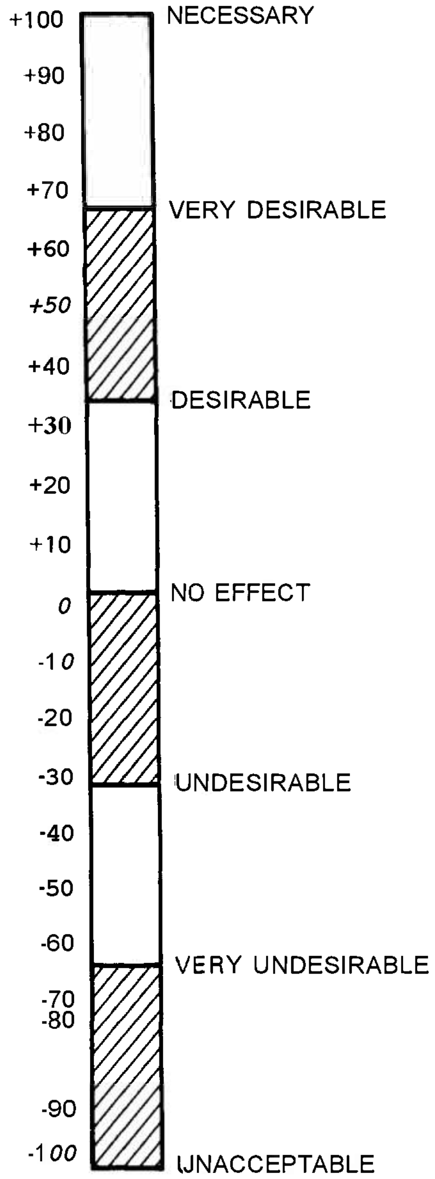


Figure 5-20. Basic Rating Scale

Design Problem. An airborne navigation system requires precise voltage regulation in its power supply for satisfactory operation. Because of rigid weight and space requirements placed upon the designer, it is necessary to stress the power transformer beyond recommended limits, thereby increasing the failure rate potential of the transformer. Replacement is anticipated approximately every 500 hr of operation. Turn-around time of the airplane is 18 min and cannot be increased. Replacement of the power transformer in present design will require approximately 35 min on the average.

The question arises as to how the power supply should be redesigned to meet aircraft turn-around time requirements most effectively.

Possible Solution. Two possible solutions are considered for application:

1. Redesign the power supply as a completely replaceable unit for rapid replacement of the entire assembly.

2. Redesign the power transformer as a quick-disconnect (plug-in) part for rapid replacement of failed transformer within the power supply assembly.

The data sheets for each of these alternative solutions are given in Tables 5-13 and 5-14 and shown in the graphs of Figs. 5-21 and 5-22. From these, it can be seen that the replaceable power supply is the preferred solution.

The NSIA Trade-off Technique was used in a maintainability study for an Army diesel-driven tractor performed by the US Army Engineer Research and Development Laboratories, Fort Belvoir, Va. (Ref. 78). The results are described in Ref. 11.

5-8.7.3 Repair/Discard Trade-off Decisions

One of the important maintainability design trade-off decisions is whether to design a particular item for repair or discard at failure. This decision impacts support resources (technician numbers and skill levels, repair parts, facilities, test and support equipment, and maintenance information), as well as the specific maintenance actions to be taken and the levels at which repairs are to be made. The decision also affects such system and equipment design attributes as safety, reliability, accessibility, test points, controls, displays, and human factors. The repair/discard trade-off decision is primarily an economic one, and it significantly influences system life-cycle costs (see Chap. 7).

A number of repair/discard trade-off models have been developed since the mid-1950's, primarily for electronic systems and equipments. At the time of some of the earlier studies, most electronic equipment consisted of vacuum tubes and discrete components mounted on terminal boards. Printed circuit boards and transistors were just coming into use, and early development versions of micro-electronic packaging techniques existed. In more recent years, with the extensive development of solid-state devices and integrated circuits which allow high package densities, the emphasis has shifted to modular packaging and new maintenance concepts and techniques. These newer techniques have often resulted in order-of-magnitude improvements in both reliability and maintainability and, together with production efficiencies and quantity cost reduction, have made discard-at-failure an economic reality today.

Repair/discard trade-off decisions may be classified into two types. The first type is primarily a ***repair level*** decision, useful for developing maintenance support concepts during the system planning phases and during deployment phases after design has been completed. The decision to be made is concerned with optimizing the maintenance and support levels at which repairs are most economical to effect; for example, whether it is more economical to repair a repairable item at direct support or at general support. The second type of repair/discard decision is design-oriented for application during the late planning and the design phases of the system life cycle. Some models can be used for design and repair level decisions. A flow chart of the replace/repair/discard decision problem is shown in Fig. 5-23.

In a study entitled ***Criteria for Repair vs Discard Decisions*** (Ref. 79) the Logistics Management Institute (LMI) noted the high interaction between repair/discard decisions and other system design and support economic (life-cycle cost) decisions. It identified five major decision points in the system life cycle where repair/discard decisions might logically be made (Fig. 5-24). The first of these decision points, Development of Design Specifications, occurs during the Concept Formulation and System Definition Phases. It depends upon operational, maintenance, and logistic support policies as well as cost-effectiveness and other economic criteria established during concept and system studies. At this level, repair/discard decisions are primarily broad policy decisions which become part of the overall maintenance and logistic support concept. They result in the establishment of both qualitative and quantitative criteria in system development specifications to guide system/equipment design engineers during the development and design phases.

The second point in the LMI model, Initial Design or Item Selection, occurs during full-scale development. The policies and criteria previously established are now applied to assemblies, subassemblies, and

TABLE 5-13.
DATA SHEET—REPLACEABLE POWER SUPPLY (Solution No. 1)

	Parameters	Considerations	Relative Weighting	Basic Rating		Adjusted Values	
				Undesirable	Desirable	Undesirable	Desirable
1.	Production Schedule	Will be delayed 2 wk. Undesirable but can be tolerated.	3	−30		−90	
2.	Support Requirements	Stocking of complete power supplies as spares increases storage space requirements.	3	−20		−60	
3.	Maintenance Costs	No net effect.	1	0	0	0	0
4.	Environmental Influence	No net effect.	1	0	0	0	0
5.	Reliability	Inspection and rapid replacement of power supply when operation is marginal should improve system reliability.	3		+40		+120
6.	Safety	Improved operation of navigation system improves safety of aircraft.	4		+30		+120
7.	Human Factors	Power supply failures can be repaired under more favorable depot conditions. Ease of replacement improves maintenance.	1		+30		+30
8.	Fabrication Costs	Increases cost approximately $200/unit.	2	−40		−80	
9.	Maintenance Time	Replacement time reduced to approximately 10 to 14 min/replacement.	4		+60		+240
10.	Performance	No net effect.	3	0	0	0	0
11.	Maintenance Personnel	Reduction in maintenance time reduces overall manpower requirements.	2		+20		+ 40
12.	Weight and Space	No net effect.	0	0	0	0	0
		TOTAL	27	−90	+180	−230	+550

Calculations: Net Value: +550 − 230 = + 320
Average Net Value: +320 ÷ 27 = + 11.8

Desirable [X] Undesirable [ ] Average Net Value +11.8

TABLE 5-14.
DATA SHEET—PLUG-IN TRANSFORMER (Solution No. 2)

	Parameters	Considerations	Relative Weighting	Basic Rating Undesirable	Basic Rating Desirable	Adjusted Values Undesirable	Adjusted Values Desirable
1.	Production Schedule	Will be delayed 3 wk. Undesirable, but can be tolerated.	3	–40		–120	
2.	Support Requirements	Requires that replacement transformers be stocked as spares. No problem.	3	0	0	0	0
3.	Maintenance Costs	Maintenance cost slightly lower due to reduced replacement time.	1		+10		+10
4.	Environmental Influence	No net effect.	1	0	0	0	0
5.	Reliability	Introduces possibility of minor increase in connector failures.	3	–10		–30	
6.	Safety	Possible increase in connector failures decreases aircraft safety when navigation system is inoperative.	4	–20		–80	
7.	Human Factors	Plug-in units reduce work load on maintenance personnel by simplifying replacement.	1		+20		+20
8.	Fabrication Cost	Increases cost approximately $50/unit.	2	–10		–20	
9.	Maintenance Time	Replacement time could range from 13 to 22 min. Overall repair time reduced; expected to exceed turnaround time approx. 50% of the time.	4	–30		–120	
10.	Performance	Possible connector difficulties could reduce system performance slightly.	3	–10		–30	
11.	Maintenance Personnel	Reduction in maintenance time reduces manpower requirements.	2		+10		+20
12.	Weight and Space	No net effect.	0	0	0	0	0
		TOTAL	27	-120	+40	–400	+50

Calculations: Net Value: + 50 – 400 = –350
Average Net Value: – 350 ÷ 27 = –12.9

Desirable ☐ Undesirable ☐ Average Net Value –12.9

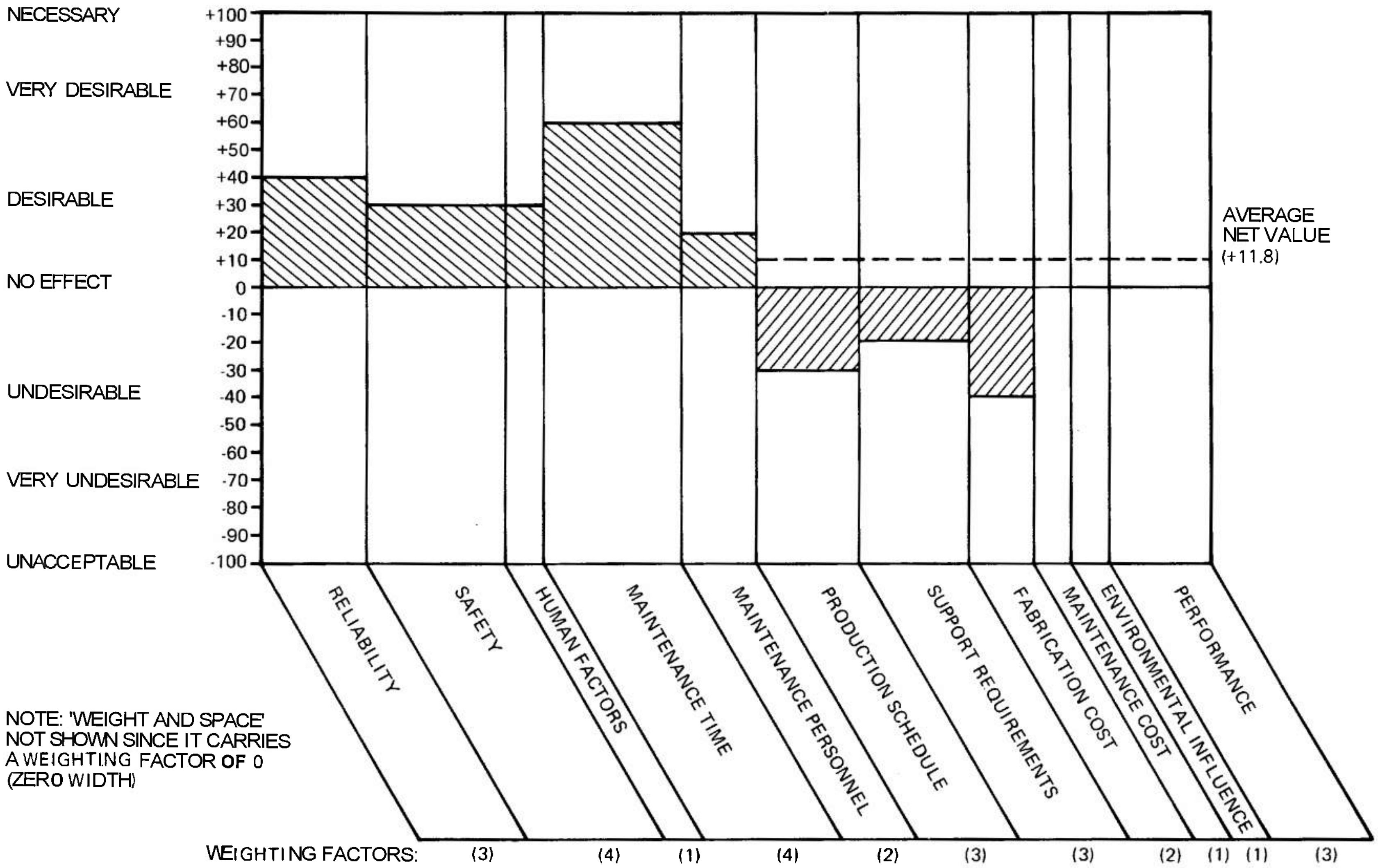


Figure 5-21. Replaceable Power Supply Evaluation

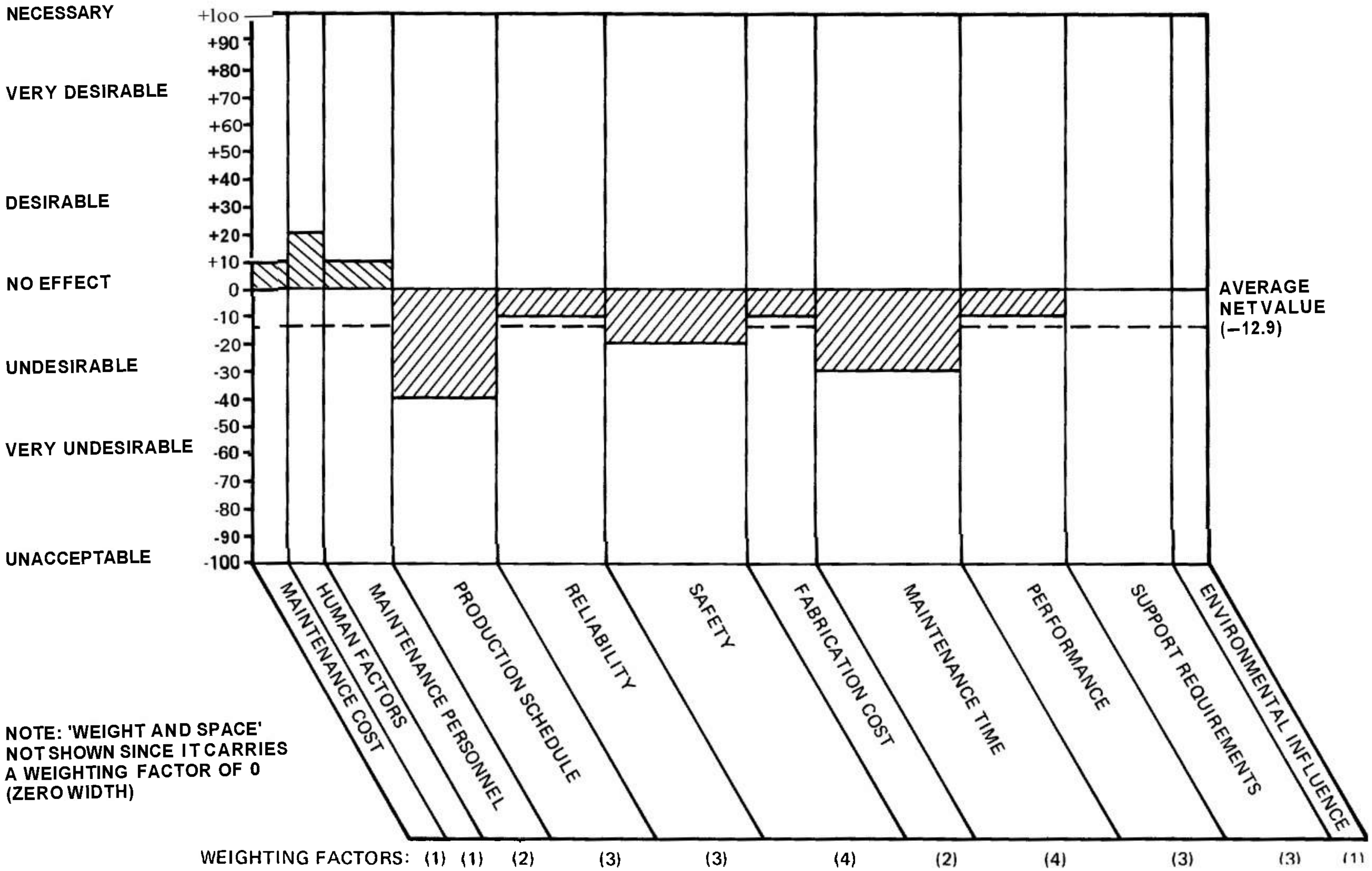
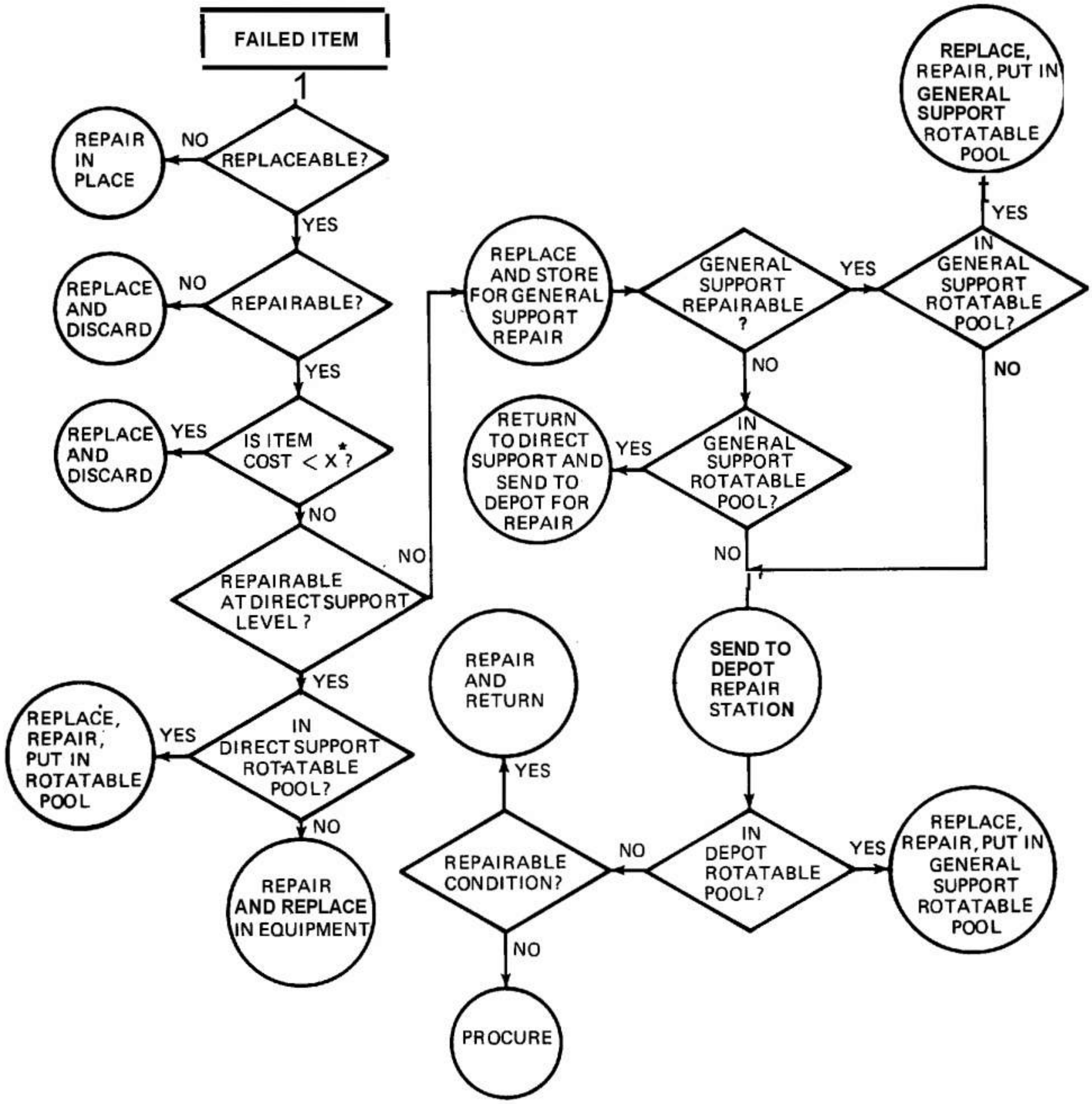


Figure 5-22. Plug-in Transformer Evaluation



* X= REPAIR-DISCARD CRITERION

Figure 5-23. Flow Chart of Replace/Repair/Discard Decisions

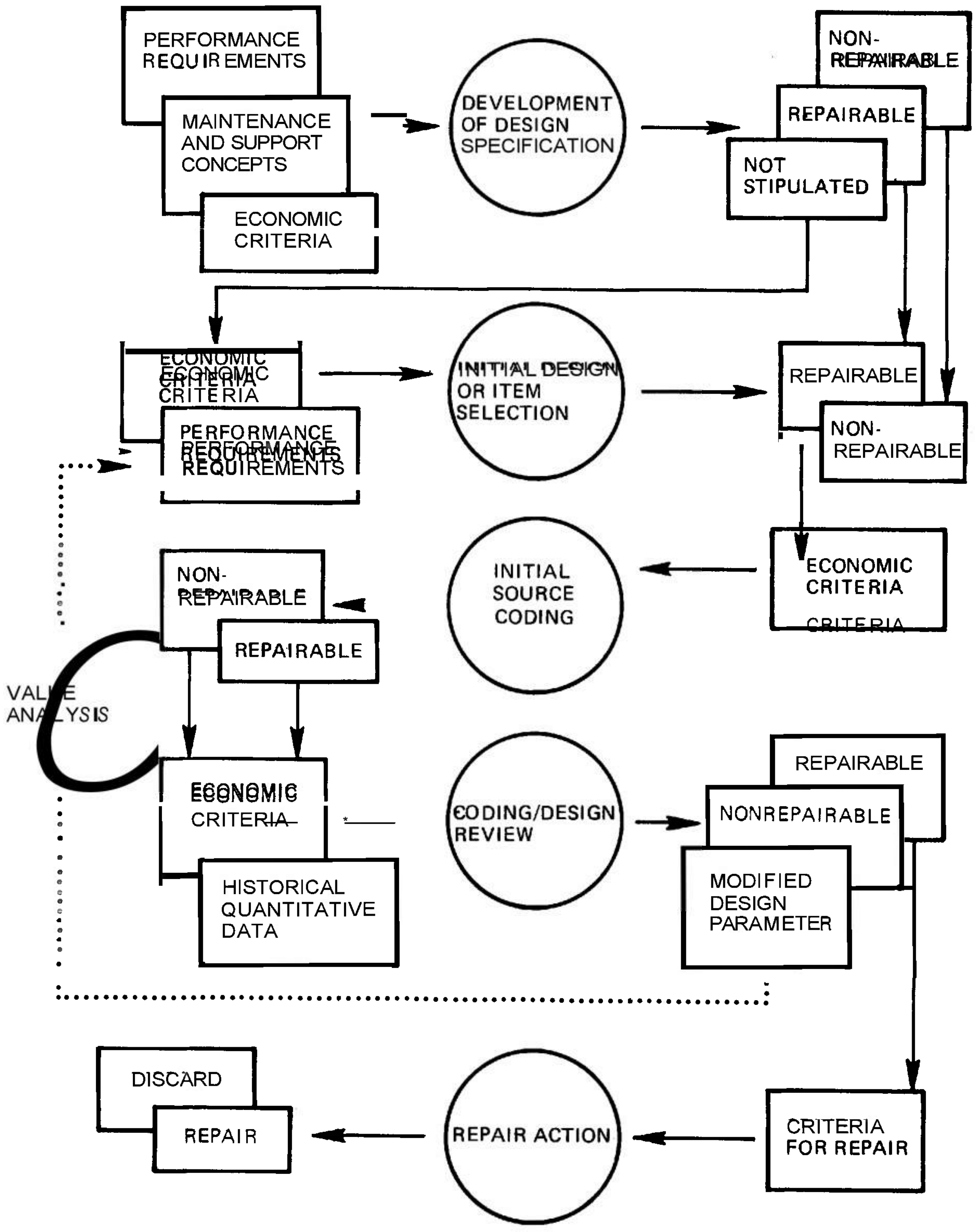


Figure 5-24. Repair vs Discard Decision Points
Reprinted with permission from *Criteriafor Repair vs Discard Decisions*, Task 65-15, Logistics Management Institute, May 1966.

modules based upon analyses using quantitative repair/discard cost models.

The third decision point, Initial Source Coding for Provisioning, occurs during the late design and early production phases. At this point, the major design decisions with regard to repair/discard have already been made. The decisions at this point, therefore, are primarily logistic support decisions, such as range and depth of spares, the effect on operational readiness of maintenance and supply delays, transportation and pipeline effects, numbers and locations of test and repair stations, operational readiness float, and similar logistic support decisions.

The fourth decision point in the LMI model, Coding /Design Review, occurs during the operation and support phases of system deployment. It is the point where previously established repair/discard decisions and criteria may be reviewed for validity based upon historical operation and support data collected from the field use of the system/equipment. Such reviews may result in a change in the repair/discard decision or in a modification of the design through the use of value engineering.

The fifth decision point, Repair Action, is concerned with whether a reparable item is still economically worth repairing after a failure has occurred or upon a scheduled maintenance inspection, in view of damage, age, wear, or other condition of the item.

Of these five decision points, the first two have significant impact upon system/equipment design for maintainability. The LMI report points out that the quantitative values of most decision criteria are dependent upon the results of a variety of design and support decisions other than the repair/discard decision.

5-8.7.3.1 Repair/Discard Models

Repair/discard models are economic decision models in which the cost of repairing an item is compared with the cost of discarding the item and replacing it with a new one, and the lower cost option chosen. The costs considered are usually total life-cycle costs. The comparison is often displayed in terms of a difference equation, such as

$$\Delta C = C_r - C_d \qquad (5\text{-}5)$$

where

ΔC = cost difference
C_r = cost of repair
C_d = cost of discard and replacement.

Thus, if ΔC is positive, the discard option is selected; if it is negative, the repair option is selected. The cost equations are usually quite complex, but the models are readily computerized. The repair and discard cost equations may contain identical terms for some cost factors, or terms which are relatively insensitive and can thus be treated as constants. These terms can then be eliminated from the difference equations. The difference equations are sometimes simplified into limited "delta cost models", and simpler screening rules are sometimes used to minimize the need to use the more complete cost models. In some cases, especially during early life-cycle phases, all data required by the complete model may not be available. The screening rules are, therefore, useful for early decisions. An example of graphic screening techniques is given in Ref. 80. Refs. 79 and 81-90 describe a number of repair/discard models which have been developed and applied. These models contain various total cost and delta cost mathematical models, screening rules, and graphical decision techniques.

The data required for use in repair/discard models fall into the following general categories:

1. Dollar cost factors
2. Maintenance and supply resource quantities and their associated costs
3. Maintenance and supply activities and their associated costs
4. Other relevant factors and their associated costs.

Typical data elements in these categories required by some or all repair/discard models are shown in Table 5-15.

Repair/discard models must be carefully applied to the particular maintainability design problem if they are to be of real benefit. The Logistics Management Institute report (Ref. 79) suggests the following guidelines:

1. Assume the item will be repaired unless a discard decision has been justified.
2. Direct the analysis initially toward the highest level of assembly and proceed to lower levels until a discard decision is justified.
3. Apply screening rules before subjecting items to exhaustive economic analysis.
4. Identify and analyze significant related decisions prior to repair/discard analysis.
5. Conduct the analysis, weighted by technological and military constraints.

TABLE 5-15.
QUANTITATIVE FACTORS IN REPAIR/DISCARD MODELS

I. Dollar Costs
   - A. Calculated Values
     1. Total Life Cycle Cost
     2. Discard Cost
     3. Repair Cost
     4. Cost Difference (Δ Cost)
   - B. Input Costs
     1. Design Cost
     2. Production Cost
     3. Procurement Costs
     4. Logistic Costs
     5. Salvage Value
     6. Constants

II. Resources
   - A. Personnel
     1. Direct Labor
     2. Indirect Labor
     3. Training
   - B. Equipment
     1. Prime Equipment
     2. Support Equipment
     3. Test Equipment
     4. Tools and Fixtures
   - C. Materials
     1. Expendables
     2. Spares
     3. Repair Parts
   - D. Facilities
     1. Buildings
     2. Maintenance Areas
     3. Supply Areas
   - E. Information
     1. Maintenance Manuals
     2. Logistic Data
     3. Provisioning Data
     4. Maintenance Engineering Analysis Data
     5. Drawings and Specifications
     6. Test Programs (Software)

III. Activities
   - A. Maintenance
     1. Preventive Maintenance
     2. Corrective Maintenance
     3. Repair Levels
   - B. Supply
     1. Inventory and Inventory Control
     2. Entering and Retaining a New Line Item
     3. Logistic Processing
     4. Procurement

TABLE 5-15.
QUANTITATIVE FACTORS IN REPAIR/DISCARD MODELS (Cont.)

- C. Transporation and Handling
  - 1. Packaging and Preparation
  - 2. Storing and Handling
  - 3. Transportation
- D. Miscellaneous
  - 1. Technical Services
  - 2. Administration

IV. Other Factors

- A. Quantity
  - 1. Item Population
  - 2. Items per Module
  - 3. Number of Parts Peculiar
- B. Reliability
  - 1. Failure Rate
  - 2. Total Number of Failures
- C. Maintainability
  - 1. MTTR or Repair Rate
  - 2. Repair Cycle Time
- D. Time
  - 1. Item Life
  - 2. System/Equipment Life
  - 3. Operating Time
  - 4. Utilization Rate
  - 5. Lead Time
  - 6. Waiting Times
- E. Effectiveness
  - 1. System Availability
  - 2. Logistic Availability
  - 3. Operational Readiness

The application of repair/discard models requires that careful consideration be given to each of the cost terms with regard to their applicability to the specific problem. In some cases, the model may have to be modified. In most cases, the data base, constants, and weighting factors will have to be updated to currently applicable figures.

5-8.7.3.2 Impact of Repair/Discard Decisions on Maintainability

To be adaptable to discard-at-failure maintenance, equipment must be so designed. The discard-at-failure policy has a number of implications for maintainability design. These include reduced requirements for accessibility to lower equipment levels, reduced need for skilled personnel and for complex test equipment, reduced number of test points, simpler displays, and even the possibility of elimination of an entire level of maintenance. It also reduces the support time necessary and the need for inventory and stocking of a large number of detailed repair parts.

On the other hand, depending upon module size and the number of items per module, discard-at-failure maintenance requires a sufficient number of perhaps larger, bulkier, heavier, and more expensive modules at forward levels and in the inventory so that operational readiness will not be compromised by lack of spares.

Table 5-16 lists trade-off effects of repair vs discard considerations on the maintainability design factors listed in Table 5-5.

TABLE 5-16.
EFFECT OF REPAIR/DISCARD CONSIDERATION ON
MAINTAINABILITY DESIGN FACTORS

Factor	Repair	Discard
Accessibility	To lowest repair level	Reduced in equipment None in modules.
Test Points	To lowest level	Fewer. Perhaps as few as one per module
Controls	More with greater interaction among modules	Within module
Labeling and Coding	To individual item	To module only
Displays	More Complex. To lowest repair level	Simpler and fewer
Manuals, aids	More complex	Fewer and simpler
Test Equipment	More specialized and manual. Down to individual item for fault isolation	BITE and more automatic. Simpler fault isolation.
Tools	More	Fewer
Connectors	Probably fewer	Probably more for plug-ins
Mounting and Fasteners	Hard wiring, fewer plug-ins and connectors	More plug-ins
Handles and handling	May require more handling	May require more handles on plug-ins. Less handling.
Safety	Greater hazard to men and equipment	Less hazard
Skill level	Higher	Lower

In addition to cost factors which enter the repair/-discard decision, there are other factors which are not quantifiable or are intangible. Goldman and Slattery (Ref. 12) list some of these as follows:

1. Factors favoring discard:

a. Probable gain in equipment reliability through elimination of repair; i.e., decrease the adverse effects of technicians accidentally injuring equipment.

b. Probable saving in development cost and manufacturing because of elimination of the need for accessibility within the throwaway module.

c. Overall reduction in unit cost of module because of the manufacture of larger quantities.

d. Compatibility of throwaway philosophy with present trend toward marginal testing and automatic failure isolation. The development effort required to include marginal testing and automatic failure isolation features in test equipment is often the same as required to make the test equipment compatible with throwaway maintenance.

e. Simplification of field modifications where plug-ins of a new unit will substitute for field rework.

f. Reduction in cost of training and supporting maintenance technicians and the repair facilities and performance aids they require.

g. Release of repair facility floor space for other uses, for example, storage of spares.

h. Use of repair facilities for priority repairs rather than indiscriminate queries for many items.

2. Factors opposing discard:

a. Emotional feelings that discarding complex units is wasteful. (Salvaging of high value parts may sometimes be a solution.)

b. Loss of reliability information when troubleshooting of the component part is eliminated.

c. Possible increased size and weight (and loss of reliability) because of the need for more connectors.

d. Storage space requirements over and above what is required under a repair policy.

e. Loss of some of the capability of the forward echelon to react to emergency situations. Shortages of spares become more critical unless emergency repair is provided for (for example, by stocking emergency parts kits).

f. Logistic cost of discarded items.

AFLCM/AFSCM 375-6 (Ref. 87) points out that repair level decisions must be made as an integral part of system design, since the investment made during the acquisition phase precludes or seriously inhibits subsequent reversal of repair level decisions during the operational phase.

5-8.8 COST CONSIDERATIONS

It should be evident from the previous discussions in this chapter that cost is a major element in maintainability design. While the accomplishment of mission and operational objectives must be a prime consideration, many opportunities arise which allow for making the necessary cost-effectiveness trade-offs that provide the proper balance between desired operational and support features in a system and the economic use of resources. Costs represent one of the constraints within which a system must be designed, produced, operated, and supported.

One of the purposes of this chapter was to discuss the system approach to maintainability design; the specific maintainability design factors which must be considered by system, design, and maintainability engineers; the interface relationships which exist between maintainability and its related disciplineg, including reliability, human factors, safety, standardization, automation, and others, along with the trade-offs between them. It has been pointed out that the competition for scarce resources required that careful consideration be given to the economic and cost implications of maintainability design decisions. In particular, it has been shown that an increased investment in acquisition costs to improve system maintainability may return many times its amount in savings in operation and support costs, and thus may minimize system life-cycle costs in terms of savings in manpower costs, cost of repairs and repair parts, storage, transportation, handling costs, and many others. For example, it has been estimated that the reduction in the need for one maintenance technician may bring savings of up to $300,000 in the life-cycle cost of a system.

Cost considerations in maintainability are further discussed in Chap. 7.

REFERENCES

1. H. S. Dordick, "Design Configuration", Chapter 4 in *Maintainability Design,* F. L. Ankenbrandt, Ed., Engineering Publishers, Elizabeth, N.J., 1963.
2. H. S. Dordick, *Maintainability—A Primer in Designingfor Profit,* RAND Corporation, Report P-2886, May 1964.
3. AR 750-1, *Army Materiel Maintenance Concepts and Policies.*
4. AMCR 750-15, *Integrated Logistic Support.*
5. F. D. Mazzola, et al., *Maintainability Technique Study,* Final Report. Phase V, RADC-TR-63-85, Vol. 1, RCA Service Company, February 1963.
6. F. D. Mazzola, et al., *Maintainability Engineering,* RADC-TR-63-85, Vol. II, RCA Service Company, Feb. 1963.
7. DoD Directive 4100.35, *Development of Integrated Logistic Support for Systems/Equipments,* October 1970.
8. TM 38-703, *Integrated Logistic Support, Management Guide.*
9. J. S. Foster, Jr. and B. J. Shillito, Memorandum subject *Revision of the Integrated Logistic Support Planning Guide, 4100.35-G,* Office of the Secretary of Defense, 25 March 1971.
10. M. B. Kline, *Maintainability Considerations in System Design,* Doctoral dissertation, Department of Engineering, UCLA, 1966, University Microfilms, Inc., Ann Arbor, Michigan.
11. AMCP 706-134, Engineering Design Handbook, *Maintainability Guide for Design.*
12. A. S. Goldman and T. B. Slattery, *Maintainability: A Major Element of System Effectiveness,* John Wiley & Sons, 1964.
13. H. H. Goode and R. E. Machol, *System Engineering,* McGraw-Hill, 1957.
14. A. B. Rosenstein and J. M. English, "Design as a Basis for a Unified Engineering Curriculum", *Proceedings of the First Conference on Engineering Design Education,* Case Institute of Technology, Cleveland, Ohio, September 1960.
15. M. Asimow, *Introduction to Design,* Prentice-Hall, 1962.
16. A. D. Hall, A *Methodology for Systems Engineering,* Van Nostrand, 1962.
17. M. B. Kline and M. W. Lifson, "Systems Engineering", Chapter 2 in *Cost-Effectiveness: Economic Evaluation of Engineered Systems,* J. M. English, Ed., John Wiley & Sons, 1968.
18. MIL-STD-471, *Maintainability Demonstration.*
19. MIL-HDBK-472, *Maintainability Prediction Handbook*
20. TM 38-703-3, *Maintenance Engineering Analysis Data System.*
21. P. Sgouros, "Life Cycle and Implications", *Proceedings of the NMSE Systems Performance Effectiveness Conference,* (SPECON *1),* US Naval Applied Science Laboratory, 1965, pp. 75-92.
22. A. S. Goldman, et al., *Economics of the Trade-off Among Reliability, Maintainability, and Supply,* G. E. TEMPO Report RM62TMP-42, June 1962.
23. J. V. Sanderson, "DDF for Design Review and Maintainability", *Proceedings of the Second System Performance Effectiveness Conference,* (SPECON *2),* US Naval Applied Science Laboratory, April 1966, pp. 117-134.
24. J. M. McKendry, et al., *Maintainability Handbook for Electronic Equipment Design,* US Naval Training Device Center Technical Report NAVTRADEVCEN 330-1-4, April 1960.
25. J. G. Wohl, "Why Design for Maintainability?", *IRE Transactions on Human Factors in Electronics,* Vol. HFE-2, September 1961, pp. 87-92.
26. J. D. Folley and J. W. Altman, *Guide to Design of Electronic Equipment for Maintainability,* US *Air* Force, Wright Air Development Center, Dayton, Ohio, TR 56-218, April 1956.
27. M. R. Munger and M. P. Willis, *Development of an Index of Electronic Maintainability,* American Institute for Research, Pittsburgh, PA., Report AIR-59-PR-207, 1959.
28. *Weapons Requirements: Integrated Maintenance Management for Aeronautical Weapons, Weapon Systems, Related Equipment,* US Navy, Bureau of Naval Weapons, Specification WR-30, May 1963.
29. J. G. Wohl, et al., *An Analysis of the Maintenance Support System,* US Naval Training Device Center, Technical Report NAVTRADEVCEN 502-1, August 1961.
30. J. W. Altman, et al., *Guide to Design of Mechanical Equipment for Maintainability,* US

Air Force, Aeronautical Systems Division, Dayton, Ohio, Technical Report ASD-TR-61-381, August 1961.
31. L. V. Rigby, Jr., et al., *Guide to Integrated System Deisgn for Maintainability, US* Air Force, Aeronautical Systems Division, Dayton, Ohio, Technical Report ASD-TR-61-424, October 1961.
32. *Maintainability Design Criteria Handbook, US* Navy, Bureau of Ships, NAVSHIPS 94324, Change No. 2, March 1965.
33. *Guide for Measurement and Reporting of Electronic Equipment Maintainability,* Electronic Industries Association, EIA Maintainability Bulletin No. 2, January 1961.
34. N. J. Buchaca and J. Brinda, Jr., *Maintainability Simulation and Demonstration Equipment Description and Application,* Technical Memo No. 1, Philco Corporation, Western Development Laboratories, Palo Alto, Calif., December 1962.
35. J. Cooper, "An Integrated Approach to System Maintainability", Chapter 14 in *Electronic Maintainability,* F. L. Ankenbrandt, Ed., Engineering Publishers, Elizabeth, N.J., 1960.
36. *Philco WDL Maintainability Program,* WDL-TR-1708B, Philco Corp., Western Development Laboratories, Palo Alto, Calif., April 1964.
37. G. Bissell and G. M. Almeida, Jr., *Maintainability and Maintenance Support Plan,* Martin Company, Orlando, Fla., Report No. OR-3822, April 1964.
38. *Human Factors Design Standards for the Fleet Ballistic Missile Weapon System,* Vol. 2—Design of Equipment, US Navy, Bureau of Naval Weapons, NAVWEPS OD-18413A, Vol. 11, August 1962.
39. B. S. Blanchard and E. E. Lowery, *Maintainability: Principles and Practices;* McGraw-Hill, 1969.
40. *Maintainability Engineering Handbook,* US Navy, Ordnance Systems Command, NAVORD OD-39223, February 1970.
41. R. E. Redfern, "Design Reviews", Chapter 6 in *Maintainability Design,* F. L. Ankenbrandt, Ed., Engineering Publishers, Elizabeth, N.J., 1963.
42. DA Pamphlet 705-1, *Maintainability Engineering.*
43. C. T. Morgan, et al., Eds., *Human Engineering Guide to Equipment Design,* McGraw-Hill, 1963.
44. R. DePaul, "Design Disclosure Format", *Proceedings of the NMSE Systems Performance Effectiveness Conference,* (SPECON I), *US* Naval Applied Science Laboratory, 1965, pp. 137-165.
45. T. G. Timberlake, "Human Engineering in Corps of Engineers Equipment Design", *Conference Report, Fourth Annual US Army Human Factors Engineering Conference,* September 1958.
46. A. Chapanis, *Man-Machine Engineering,* Wadsworth Publishing Company, Inc., Belmont, Calif., 1965.
47. Col. K. H. Ewbank, "The Application of User Guidance to Equipment Design", *Conference Report, Fourth Annual US Army Human Factors Engineering Conference,* September 1958.
48. W. E. Woodson and D. W. Conover, *Human Engineering Guide for Equipment Designers,* 2nd edition, University of California Press, Berkeley, Calif., 1966.
49. E. J. McCormick, *Human Engineering,* McGraw-Hill, 1957.
50. MIL-STD-1472, *Human Engineering Design Criteria for Military Systems, Equipments, and Facilities.*
51. MIL-STD-882, *Requirements for System Safety Program for Systems, Associated Subsystems, and Equipment.*
52. R. B. Gordon, "Defining the Safety Mission", *Annals of Reliability and Maintainability, Vol.* 6, 1967, Society of Automotive Engineers, New York, pp. 536-541.
53. K. H. Eagle, "Fault Tree and Reliability Analysis Comparison", *Proceedings* 1969 *Annual Symposium on Reliability,* Institute of Electrical and Electronics Engineers, New York, pp. 12-17.
54. W. Hammer, "Numerical Evaluation of Accident Potentials", *Annals of Reliability and Maintainability,* Vol. 5, 1966, American Institute of Aeronautics and Astronautics, New York, pp. 494-500.
55. R. J. Schroder, "Fault Trees for Reliability Analysis", *Proceedings 1970 Annual Symposium on Reliability,* Institute of Electrical and Electronics Engineers, New York, pp. 198-205.
56. D. F. Haasl, "Advanced Concepts in Fault

Tree Analysis", *System Safety Symposium,* Seattle, Washington, June, 1969.

57. *Use of Automatic Test Equipment for Maintenance – a Reconnaissance,* Task 70-11, Logistics Management Institute, Washington, D.C., May 1971.
58. MIL-STD-1326(NAVY), *Test Points, Test Point Selection, and Interface Requirements for Equipments Monitored by Shipboard On-Line Automatic Test Equipment.*
59. L. R. Hulls and J. E. Lidiak, *Vehicle Diagnostic Systems,* Aerospace Systems Division, RCA Defense Electronic Products, Burlington, Mass. (undated)
60. MIL-STD-1309, *Definition of Terms for Automatic Electronic Test and Checkout.*
61. *Test, Measurement, and Diagnostic Equipment Register, US Army* Weapons Command, June 1971.
62. *Cost and System Effectiveness Analysis of the Computer-controlled Automatic Test Equipment (CATE),* Report No. ECOM-CATE-1, Aerospace Systems Division, RCA Defense Electronic Products, Burlington, Mass., July 1969.
63. *Final Report on USAF Qualitative Requirements for Standardized Automatic Test Equipment,* Standard ATE Study Group, Directorate of Materiel Management, San Antonio Materiel Area, Kelly AFB, Texas, June 1969.
64. AR 750-43, *Maintenance of Supplies and Equipment: Test, Measurement, and Diagnostic Equipment (TMDE).*
65. W. E. Grunwald and O. T. Carver, "LCSS—an ATE Case History".
66. E. M. Stockton and D. A. Douglas, "Automatic Test Equipment for Army Field Use", *Automated Support Systems,* Aerospace Systems Division, RCA Defense Electronic Products, Burlington, Mass, 1968.
67. R. L. McCollor, "DIMATE—A Space-Age Robot", *Automated Support Systems.* Aerospace Systems Division, RCA Defense Electronic Products, Burlington, Mass., 1968.
68. R. N. Knox, "The DIMATE System—Programming and Experience", Automated Support Systems, Aerospace Systems Division, RCA Defense Electronic Products, Burlington, Mass., 1968.
69. F. Pradko, *Automotive Test Equipment,* US Army Tank-Automotive Command, Warren, Mich., 20 December 1969.
70. D. S. Sama, F. Pradko, and D. F. Ancona, *Diagnostic Equipment for Army Vehicles,* US Army Tank-Automotive Command, Warren, Mich. (undated)
71. *Multipurpose Automatic Inspection Diagnostic Systems (MAIDS) Opemtion at a US Army Depot,* ESP 7101, Hamilton Standard Division, United Aircraft Corporation, Windsor Locks, Conn., March 1971.
72. M. B. Kline, "Application of the Lognormal Distribution to Maintainability Design", *Logistics Spectrum,* Vol. 1, No. 2, Winter 1967-68.
73. *EIA Maintainability/Trade-off Techniques,* EIA Bulletin No: 8, Electronic Industries Association, Washington, D.C., July 1966.
74. *Maintainability Trade-off Techniques,* Final Report of the Trade-off Task Group, National Security Industrial Association, Federal Electric Corporation, Paramus, N.J., Report No. 217-61A, March 1962.
75. T. J. Gordon and O. Helmer, *Report on a Long-Range Forecasting Study,* Report P-2982, The RAND Corporation, Santa Monica, Calif., September 1964.
76. N. Dalkey, *The Delphi Method: An Experimental Study of Group Opinion,* Report RM-5888-PR, The RAND Corporation, Santa Monica, Calif., June 1969.
77. M. W. Lifson, "Value Theory", Chapter 6 in *Cost-Effectiveness: Economic Evaluation of Engineered Systems,* J. M. English, Ed., John Wiley & Sons, 1968.
78. *Maintainability Engineering Study for Universal Engineer Tractor No, III, – Rubber Tired,* US Army Engineer Research and Development Laboratories, Fort Belvoir, Va., 1963.
79. *Criteria for Repair vs Discard Decisions,* Task 65-15, Logistics Management Institute, May 1966.
80. J. K. Segar, "Repair or Throwaway–Graphic Screening Techniques Help Supply the Answer", *Logistics Spectrum,* Vol. 5 No. 3, Fall, 1971.
81. J. J. Andrea, *An Evaluation of Module Replacement and Disposal at Failure for Maintenance of Ground Electronic Equipment,* RADC-TR-60-5, Collins Radio Company, Cedar Rapids, Iowa, December 1958.
82. E. G. Wrieden, *Criteria for Discard-at-Failure*

Maintenance, RADC-TDR-63-140, IBM Corporation, Federal Systems Division, Oswego, N.Y., March 1963.

83. R. E. Purvis and R. L. McLaughlin, *Validation of Discard-at-Failure Maintenance Mathematical Model,* RADC-TR-65-214, RCA, February 1966.
84. O. E. Davis, *Economic Decision Criteria for Repair versus Throwaway Maintenance,* Vitro Laboratories, Silver Spring, Md., Technical Report No. 01816.01-3, September 1964.
85. D. Hoffman, *Throwaway or Repair Implications on Maintenance Cost Program,* Raytheon Company, Report No. BR-4298, September 1967.
86. *Summary Description of COAMP-II: Part I – Model Description,* RCA, Report No. 67-526-MATE-5C, Rev. April 1968.
87. *Optimum Repair-Level Analysis,* AFLCM/AFSCM **375-6,** US Air Force, May **1968.**
88. *Level of Repair for Aeronautical Material, General Program Requirements,* Naval Air Systems Command AR-60, June 1970.
89. **K.** D. Harris, H. R. Gehle, and D. F. Meer, *AMC Replacement Unit Repair Level Analysis Model,* US Army Management Engineering Training Agency, December 1970.
90. R. D. White and D. A. Tyburski, *Generalized Electronics Maintenance Model (GEMM),* US Army Electronics Command, May, **1971.**
91. AMCP 706-115, Engineering Design Handbook, *Environmental Series, Part One, Basic Environmental Concepts.*
92. *A Course of Instruction in Risk Analysis,* ALM-**3315-H,** US Army Logistic Management Center, Ft. Lee, Va.

CHAPTER 6

MAINTAINABILITY TEST AND DEMONSTRATION

SECTION I

INTRODUCTION

6-1 GENERAL REQUIREMENTS

The Maintainability Test and Demonstration task is one of the twelve Maintainability Engineering Program tasks whose primary function is to "verify maintainability" that has been "designed-in" and "built-in" to the equipment (Ref. 1). Up to this point in the equipment development, the tasks of the Maintainability Program have been analytical in nature, providing a confidence that both the quantitative and qualitative maintainability requirements would be met. This confidence has been achieved in terms of:

1. Establishing the design criteria
2. Making allocations and predictions
3. Conducting designer indoctrination and design monitoring
4. Participating in design reviews, design analysis, and design trade-offs
5. Analyzing the specific design configurations in terms of the requirements for logistic life cycle maintenance support.

However, these evaluations do not reflect practical experience with the actual hardware. Therefore, it is essential to close the maintainability loop of analytical confidence evaluation and realistic confidence evaluation by exercising the hardware in an operational environment and performing actual maintainability tests and demonstrations involving the prime equipment and its associated life cycle logistic resources (i.e., support equipment, technicians, technical data).

Due to costs and time schedules, it is not always practical to exercise the hardware through a total life cycle and total population deployment under field environment in order to arrive at an accept/reject decision. Fortunately, there are statistical methods using relatively small sample sizes by which a satisfactory accept/reject evaluation can be made concerning the maintainability objectives.

In planning maintainability tests and demonstrations where statistical acceptheject decision-making methods are employed, there are three essential features that must be recognized by both the developer and the user:

1. The quantitative parameters of time and man-hours required to perform a maintenance action include administrative and logistic delay times beyond the control of the hardware designer. Direct maintenance tasks and sequence flow are controllable by building into designs the qualitative characteristics that define the times to perform specific maintenance functions (i.e., preparation, access, diagnosis, replacement, checkout). Delay times can be minimized further by using standardized or multi-use parts in the design, thereby improving the ability of the logistic system to respond. However, the type of administrative and logistic delay factors indirectly caused by the equipment design characteristics are beyond the designer's control. This is especially true when the direct maintenance steps, maintenance concept, allocation, predictions, and procedures during the design and development phases have been reviewed and accepted by the producer and the consumer.

2. The design parameters for maintenance time are analyzed, summed, and established by considering the contribution of the frequency of maintenance actions (failure rates for corrective maintenance and scheduled maintenance rates for preventive maintenance). Once these rates have been established during the hardware design and development phases, the requirements for the design maintenance time parameters are set. If the failure rates in the hardware change from the originally predicted values, the maintenance action times may not change accordingly. Therefore, in maintainability test design one utilizes random selection of maintenance action tasks by an unbiased random sampling method

that is based upon the predicted maintenance frequency rates.

3. The determination of the producer risk α and the consumer risk β, the test method to be used, and the selection of the sample size for maintainability demonstrations must be negotiated early in the program when the maintainability demonstration plan is established and approved. The producer risk represents the risk that a product conforming to specifications will be *rejected* in the test. The consumer risk represents the risk that a product not conforming to specification will be *accepted* in the test. These factors must be considered during the planning and accomplishment of the maintainability program tasks of allocation and prediction in order to avoid over- or under-design of the maintainability features, and over- or underestimating the integrated life-cycle logistic requirements.

The tests and demonstrations normally are performed in accordance with the requirements and procedures set forth in MIL-STD-471 ***Maintainability Demonstration*** (Ref. 2), and MIL-STD-473 ***Maintainability Verification/Demonstration/Evaluation for Aeronautical Systems*** (Ref. 3). These are guides for the planning, preparation, and submission of the test plan, the actual testing to arrive at an accept/reject decision, and test evaluation. The planning and conduct of the formal accept/reject maintainability demonstration plans are negotiated factors that must be carefully established and agreed upon prior to contract awards. The details of the negotiated plans are updated as the design and development progresses. Normally, for Army contracts, the demonstrations are of three types: Development Tests (DT), User Tests (UT), and Maintenance Evaluation—the extent of each depending upon the contractual requirements. For the "Fly Before Buy" type of procurement, the maintainability demonstrations are part of the program designed to arrive at an accept/reject decision before commitment to production.

There are usually no limits imposed for contractor tests and demonstrations of the maintainability factors at various phases of the contract. These are conducted as the contractor deems necessary to check out the designs by such tests as proof of design, mock-up, breadboard, prototype, reliability model, environmental, quality control, and production.

The formal tests of the accept/reject decision to determine contract compliance should be conducted in an operational support environment or simulated environment, using other operational-type maintenance support resources and production-type systems or models as similar as possible to those specified by contract. In cases where the use of simulation-type conditions and hardware prototype models are specified by contract, one must try to achieve results upon which decisions can be made and interpreted to ascertain real-life effects.

6-1.1 PLANNING AND CONTROL JIREMENTS

The following are the minimum requirements for planning and control in order to obtain the maximum benefit from maintainability tests and demonstrations:

1. ***MIL-STD-471 Requirements.*** The guides listed in MIL-STD-471 (Ref. 2) and MIL-STD-473 (Ref. 3) must be responded to by the contractor. Elements for the planning and control of the demonstrations are:

   a. Test conditions
   b. Test teams
   c. Test support materials
   d. Pre-demonstration phase
   e. Formal demonstration phase
   f. Retest phase
   g. Demonstration procedures
   h. Maintenance task selection and sampling
   i. Selection of test method
   j. Corrective and preventive maintenance task:
   k. Servicing and turn-around tasks
   l. Maintenance task performance
   m. Maintainability data collection
   n. Demonstration reporting, evaluation and analysis procedures
   o. Demonstration administration and control procedures.

These guides are basic to the success of a maintainability demonstration.

2. ***Demonstration*** Model (Refs. 4 and 5). For the formal test and demonstration accept/reject decision, the preferred hardware configuration is the "production model" that will be used in the military deployed environment. Quite often this is not feasible. For instance, in the "Fly Before Buy" concept, the configuration is either as similar as possible to the production model or is the prototype model developed to demonstrate the contractor's fulfillment of the customer's performance requirements. In some cases, there is more than one prime contractor involved, and there may be two entirely different configurations to be demonstrated. In other instances, there may be engineering test configurations and service test configurations those being the same model in some cases and in othe cases variations of the prototype, as the contract schedule may dictate. If the design experiment requires

an early decision before the contract proceeds, the use of breadboard, mock-up and/or prototype models may be required. An example is the case where the breadboard represents an operating model similar to the production model but does not reflect the similarity in size or packaging. Another case is where the mock-up model is not fully functional because it allows only substitution of alternate subassemblies or parts, reflecting such maintainability characteristics as access, size, packaging. A further example is the prototype model, which is far superior to the breadboard and mock-up models since it is more representative of the production model. Prototype models are the most desirable to use for formal maintainability demonstration because of costs, scheduling, and improvements in design which have been incorporated to correct apparent deficiencies and safety hazards. The prime disadvantage in prototype models is that they often have handcrafted parts which neatly fit together, whereas mass-produced items may present difficulties with tolerances and other quality control problems.

One of the most effective ways for demonstrating maintainability features is a contractor-submitted mock-up model. Mock-up models serve two fundamental functions: (a) they provide a basis for demonstrating the proposed maintainability quantitative parameters and qualitative design features of the overall equipment, and (b) they provide a designer's tool for visibility, packaging limitation and interface guides, experimentation, and planning prior to final design or drawing release. The same considerations apply to breadboard and other preliminary design models for environmental and design proof tests used during the design process. Though the latter models cannot be used for formal design accept/reject decision-making, they can be used to evaluate and confirm design decisions.

3. *Environment* (Refs. **2, 3, 4,** and 6). Test environment is important. Downtime varies considerably between laboratory-controlled conditions of climate and temperature, and operational conditions (arctic, desert, sea, swamp, mountains, tropics). For instance, where low temperatures require the use of arctic mittens to remove metal items, two to five times as much technician time is necessary for replacement, and downtime is increased accordingly. The same variation in downtime pertains to diagnostics, with the addition of warm-up and conditioning times. These variables are an integral part of and are directly chargeable to the equipment design functions. However, it is unfortunate in some program planning that because of schedule, costs, and/or lack of availability of the associated equipment, tests are not conducted under the operational environment. Thus, the customer is usually faced with the need to simulate these types of environments. Fortunately, the Army has developed several proving grounds and service test sites where the various operational environments can be simulated with a high degree of confidence for use in the formal demonstration of maintainability during engineering test, service test, and maintenance tear-down evaluation. Therefore, during the planning state the contractor should coordinate and develop plans for the use of these test sites.

Environmental factors must be considered and delineated *so* as to include the test facilities and **support** resource requirements and limitation simulations. **This** early demonstration and test planning must be coordinated with maintainability prediction and allocation tasks, and must be related in terms of the operational environment effects on the consumer and producer risks.

4. *Personnel* (Refs. **2, 3, 4,** and **5**). Personnel used to perform the maintainability demonstration tasks should possess backgrounds and skill levels similar to those imposed on ultimate user maintenance and operating personnel (peace time – war time conditions). This factor is easy to describe but difficult to achieve, especially when one considers the following: (a) personnel available for organizational and intermediate maintenance fluctuate from the 95 percentile to the **5** percentile man, and (b) maintenance personnel seldom conduct the same maintenance action and/or function in a routine iterative fashion. Rather, they perform multiple maintenance actions heterogeneously distributed over any one period of time.

Three guides to the selection of test personnel which will assist in surmounting the difficulties described and will improve the confidence level of the test are:

a. The use of enlisted military personnel is the most desirable. The producer (contractor) furnishes the job description for the required tasks and the background skills. Appropriate military personnel are lent to the test site and participate in a contractor training program which covers the specifics of the equipment being demonstrated. The contractor may reserve the right to screen or interview the selectees for background and skills.

b. The use of engineering technician personnel on a loan basis by either the user or the contractor may be necessary. Although such personnel are not regularly assigned to programs associated with the equipment being tested, they may nonetheless introduce some degree of bias into the demonstration. Formalized training is conducted covering specifics of the equipment being demonstrated.

c. Selection of user and/or contractor personnel from those currently assigned to the same or similar programs involving the equipment being demonstrated (engineering aid or maintenance technician) is least desirable. The formal and on-the-job training achieved by this type of personnel with equipment being demonstrated results in a significant bias. These must be compensated for through the use of known relationships of personnel skill versus maintenance time.

In the selection of personnel, a great deal depends on where the demonstration is to take place, i.e., at the customer's or the contractor's facility. For Army purposes of formal demonstration, Organizational Maintenance and at least Direct Support Intermediate Maintenance are conducted at the user's sites or facilities. Most General Support Intermediate Maintenance and Depot Maintenance are conducted at the contractor's facility. Therefore, the order of preference in personnel selection will vary with the test facility and the previously negotiated contractual agreements.

*5. **Parameter Specification*** (Refs. 1 to 9). The formal maintainability demonstration process is basically a verification of compliance with specified measurable parameters. Such quantitative parameters must be defined in the negotiated demonstration planning documents. Qualitative features may be presented, but they are difficult (if not impossible) to quantify and are usually meaningless in the conduct of a formal maintainability demonstration to assist in arriving at an accept/reject decision. The demonstration parameter specifications should be expressed quantitatively; e.g., the parameters demonstrated shall be the system 30-min mean time to repair, at organizational level, and 1 hr mean time to repair at the intermediate level, with a maximum time at the 90 percentile of 1.5 hr and 3 hr, respectively, a producer risk of 20% and consumer risk of 10%.

There are various measurable time parameters, such as the mean time to repair $MTTR$, mean downtime MDT, mean corrective time $\bar{M}_{ct}$, mean preventive time $\bar{M}_{pt}$, median time $\tilde{M}$, maximum time at given percentiles M_{MAX} at 90%, and man time MMH/OH. These are measurable, specific quantitative parameters for maintainability demonstration.

6. Demonstration Plan (Refs. *2, 3,* **5**, and 6).

a. Test Conditions.

Historically, in the formal demonstration tests of the maintainability parameters, documentation has shown that the planning has been done in an "ivory tower", and too little time and effort in the initial stages of the life cycle have been devoted to the planning of the demonstration design experiment. Since it is necessary to resort to statistical analyses and considerations in the experimental design, the statistician is expected to contribute to the decision method. His contribution must be based on the maintenance policy, the reasons for the experiment, and the measurable parameters. To make his participation more valuable, he must require the maintainability engineer to explain why he is doing the experiment, justify the experimental treatments he proposes to compare, and defend his position that the completed experiment will make a suitable accept/reject decision possible (Ref. **5**).

The maintenance policy is the extent of maintenance accomplished at each level, as established by the maintenance concept plan, along with the time constraints, personnel skill levels required, support equipment needs, facilities, environmental conditions, etc. The maintainability demonstration is a method of verifying the effectiveness of the maintenance policy; therefore, the maintenance policy must be clearly set forth in the demonstration plan (Refs. 2 and 3).

Prerequisite to the accomplishment of the formal demonstration is the selection of appropriate test conditions. Factors to be considered in this selection are:

(1) Maintenance policy (including definition of maintainability requirements)
(2) Demonstration test model
(3) Maintenance test data
(4) Support material and data.

In other words, the specifics of MIL-STD-471 or MIL-STD-473 must be responded to in detail. In this manner, adequate planning, budgeting, costing, and definition will assure a meaningful negotiated contract for the maintainability demonstration test. Most importantly, it will provide confidence that the maintainability objective is, in fact, achievable.

b. Test Planning (Refs. 2 and 3).

Administration and control of the demonstration is the key factor in accomplishing the demonstration on schedule, within the budgetary allowances, and in an effective manner. The essential elements of administration and control are:

(1) Organization approach
(2) Degree and responsibility of the contractor and customer participation
(3) Test team approach
(4) Organizational interfaces
(5) Assignment of responsibilities

(6) Test event scheduling
(7) Test monitoring
(8) Cost control
(9) Test data collection, reporting, and analysis.

The demonstration test planning, as previously mentioned, is conceived, proposed, and negotiated during the first phase of contractor participation in a program, such as the validation phase. Because the accept/reject decision affects the producer and customer risks, and the planning of the demonstrations directly affects the prediction and allocation tasks, they should be coordinated and interfaced early in the program phases for equipment development. As the program progresses through the various phases, the plan is updated as to test schedules, demonstration model designation, identification of logistic support resource needs, and personnel selection. The demonstration plan should cover test conditions; test planning, administration, and control; pre-demonstration phase; formal demonstration phase; retest phase; and test documentation, analysis, and reporting.

c. Test Documentation (Refs. 2 to 6):

Test documentation varies according to the type and complexity of the equipment being tested. Typical documentation requirements are:

(1) Task selection work sheets
(2) Demonstration task data sheets
(3) Demonstration work sheets
(4) Demonstration analysis work sheets
(5) Frequency/distribution work sheets
(6) Demonstration test data histograms and cumulative distribution charts
(7) Failure reports
(8) Interim results of demonstration reports and final reports.

Data may be recorded by such means as tape recorders, chart recorders, stop watch, time-lapse photography and still photography, or simple data observation work sheets. A principal point to keep in mind with regard to documentation is to use existing military maintenance data collection forms and formats, expanding these by supplemental pages as the test needs dictate. By using existing formats one is able to take advantage of established procedures for collecting, storing, and analyzing the data, including existing computerized systems for data collection. The maintainability analysis of the observed data in such an exercise is used to evaluate the test results for an accept/reject decision according to the statistical methods applicable to the specifically planned demonstration. The final report of the demonstration is the final effort in fulfilling the maintainability demonstration requirements. It usually is submitted within 30 days after completion of the demonstrations (Refs. 2 and 3).

d. Selection of the Test Sample (Ref. 2, Appendix A, and Refs. 3, 4, 9, and 10):

As mentioned in par. 6-1, the selection of the test sample is an essential feature of the maintainability demonstrations to assure that the test accept/reject statistical decision is representative and not biased. Therefore, the task selection sample is based on the random selection of tasks that are representative of the frequency contribution of any one task to all possible tasks that comprise the measurement set. The contribution of the frequency of occurrence of a maintenance action is based on the failure rate prediction and not on the frequency of occurrence of the maintenance action during any one type of performance testing.

The size of a sample to be demonstrated is representative of the maintenance tasks, based on the equipment complexity and on the probability of test error. The sample size must be determined independently of the availability of the prime equipment and support resources and of allotted program test time. It is desirable to select the sample size which is large enough to be representative and yet small enough to be compatible with the total program costs and schedule requirements. A larger sample will, of course, provide more definitive results, but time and cost of testing increase as testing continues. However, smaller sample size involves increasingly larger risks and may produce inconclusive test results. Corrective task sample and preventive task sample determinations follow the same procedure except that preventive task rates are definitive schedule rates and corrective tasks are based on failure rates and can occur any time. Fortunately, the demonstrated maintainability factors are based on simulation of the task occurrences and not on the rate of occurrences during the test. This reduces the maintainability demonstration cost and duration of test. Sometimes the observations of maintenance task times are made during other equipment performance tests, whenever corrective maintenance and/or preventive maintenance become necessary. The results may be used in the formal maintainability demonstration test analysis in order to reduce the maintenance simulations necessary to meet the computed sample size required. The sample size is determined by the selected test method and the producer and consumer risks.

The task selection uses the following procedure:

(1) List and categorize all significant assemblies, modules, or parts that make up the item to be tested. The listing is based on the mainte-

nance actions required to fulfill the maintenance task needed at the various levels of maintenance as designated in the maintenance concept plan. For example, if at the organizational level, the corrective maintenance function is to replace failed major assemblies (models, groups, etc.) and the equipment list is categorized at the level of replacement only. If at the depot level, the corrective maintenance function is to repair replaced failed assemblies and equipment lists are categorized to the component part replacement throw-away level.

(2) Indicate the quantity Qof each item listed.

(3) Determine the frequency f of occurrence of maintenance for each item listed.

(4) Determine the total rate for each 'item, i.e., $Q_i f_i$.

(5) Determine the percent contribution PC of each item maintenance task to the total of all item maintenance tasks, i.e., divide the total for each by the total of all items

$$PC = 100 \times Q_i f_i \Bigg/ \sum_{i=1}^{n} Q_i f_i \qquad (6\text{-}1)$$

(6) Where applicable, group together items within the same category for which the contribution is less than 2% (or as otherwise specified by the user) and assign a contribution percent equal to the sum of the individual percent contribution for those items.

(7) Apportion the number of tasks to be demonstrated in proportion to the percent contribution group determined above. The sample size used in the apportionment is that which was originally specified for the demonstration as computed or determined by the statistical method applied. When various possibilities exist, a table of random numbers can facilitate the selection process and reduce bias of inducing possible simulated failures of components comprising an item. Random selection is the key to the sample selection in order to develop an unbiased use of the statistics involved in any statistical accept/reject decision.

The maintenance task selection accept/reject decision method and sample size process are completed by the contractor during the early planning and predemonstration phase. The resulting data are included in the demonstration plan for customer review and approval. Once approved, the selected tasks serve as a basis for the formal accept/reject decision for the maintainability demonstration.

Table I, Appendix A, Ref. 2 and Table 14-1, Ref. 4, p. 249, are examples of the method used to determine task selection. The maintainability engineer must keep in mind the level of maintenance being demonstrated in the development of the tables (see par. (1) above). The key is the level to which a diagnostic determination can be made by the technician in the equipment functional level breakdown. The diagnostic level is the point by which the maintenance technician starts the required maintenance task to be demonstrated.

6-2 TEST APPROACHES

Not all maintainability tests are formal acceptheject demonstration tests. There are several other points in the system life cycle and its associated maintainability program tasks, both before and after the formal maintainability accept/reject decision (Refs. 2 and 3), where test data are required for:

1. Decision-making purposes that affect the maintainability design requirements and the life-cycle maintenance support evaluations.
2. Administrative and logistic control
3. Updating corrective actions and/or modifications.

Fig. 6-1 and Table 6-1 show a typical test flow diagram giving the approximate chronological order in which test data are required for decision making at critical points in the equipment life cycle. In most instances, the required data can be derived from the multipurpose tests conducted by the contractor and/or the user (e.g., function, design experiment, reliability, environmental, subassembly, module parts acceptance, engineering, service tests). These are appropriately mentioned by the maintainability engineer to provide maintainability data. The formal maintainability accept/reject decision tests are conducted by the contractor or the user, such as during ET, ST, or maintenance evaluation. Tests, other than the formal maintainability accept/reject decision demonstration, do not require or imply the detailed planning and approvals set forth in this chapter, MIL-STD-471 (Ref. 2), or MIL-STD-473 (Ref. 3). Only applicable steps are used as needed for integrating the maintainability data collection and observations required to assess the maintainability goals and objectives.

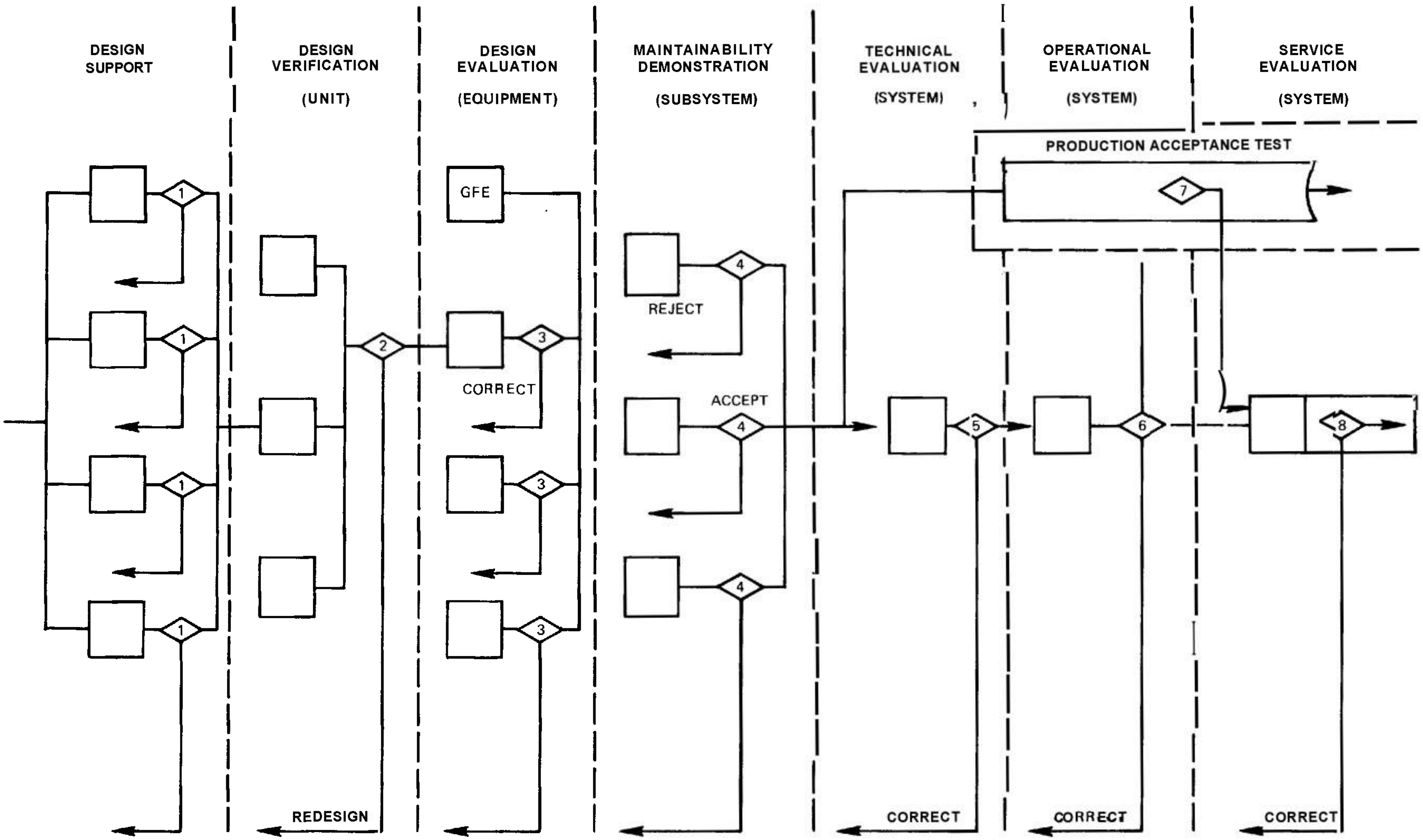


Figure 6-1. Decision Points Requiring Maintainability Test Data

TABLE 6-1.
MAINTAINABILITY TEST DATA REQUIREMENTS AT DESIGNATED DECISION POINTS

Decision Point	Maintainability Data Required	Data Acquisition Method (and Reference)
1 Approval of parts and material for design (Design Support Tests)	Test data, experience data, vendor data, etc.: (1) to prove compatibility of parts and materials with maintainability design requirement and maintenance concept, i.e., durability under maintenance conditions, stability under storage conditions, interchangeability within type designation; and (2) to identify problems and discrepancies for correction.	Integrate maintainability measurement data requirements into standard parts and materials acceptance tests.
2 Approval of design for engineering model development (Design Verification Tests)	Test data supplementing earlier prediction analyses: (1) to define maintainability design problems and critical areas; and (2) to verify problem solution feasibility.	Plan unique maintainability design-verification tests.
3 Approval of engineering model for prototype development (Design Evaluation Tests)	Test data: (1) to confirm earlier failure mode analyses and failure rate predictions; (2) to assess maintainability critical areas; and (3) to identify problems for corrective action.	Integrate maintainability measurement data requirements into functional and environmental tests.
4 Acceptance of prototype design for production (Formal Maintainability Demonstrations)	Demonstration test data: (1) to verify conformance to specified maintainability requirements (e.g., M_{ct} and $M_{max_{ct}}$) allocated to the item under test; and (2) to identify remaining discrepancies for correction.	Design and conduct maintainability demonstration test.

TABLE 6-1.
MAINTAINABILITY TEST DATA REQUIREMENTS AT DESIGNATED DECISION POINTS (Cont)

Decision Point	Maintainability Data Required	Data Acquisition Method (and Reference)
5 Technical approval of installed system (Technical Evaluation Tests—ET-ST)	Test data: (1) to verify compatibility of system installation with maintainability requirements and maintenance concept; and (2) to identify installation interface problems for correction.	Integrate maintainability measurement data into test plans.
6 Approval of installed production system for operational use (Service Tests)	Test data: (1) to confirm suitability of system maintainability; (2) to verify adequacy of maintenance and logistic support plan; and (3) to identify problem areas for correction.	Integrate maintainability measurements data into test plans.
7 Acceptance of individual items for delivery to the Military (Production Acceptance Tests)	Test data: (1) to verify adequacy of controls applied to maintainability critical parameters and processes; and (2) to identify problem areas for correction.	Integrate maintainability critical measurement data requirements into Government acceptance test for the product.
8 Approval of system improvement and modernization change proposals (In Service Deployment Tests)	Test data and operational experience data: (1) to evaluate current fleet maintainability experience; (2) to identify problem areas; (3) to verify need for and adequacy of proposed changes; and (4) to measure maintainability growth.	Integrate maintainability measurement data requirements into data collection and supplementary reporting system.

The test approaches cited in this paragraph are maintainability design features that require consideration in the maintainability allocation and prediction tasks. They are also of concern because of the effects and risks they have on the maintenance action times to be demonstrated during the formal maintainability demonstrations (early in the life-cycle development phases). They affect the diagnostic times (fault isolation and location) which constitute the largest portion of the maintenance downtime. Effectiveness is evaluated during the first three types of test shown in Fig. 6-1 (most desirable if Type 1 test) in order to bring about any necessary design changes that can be made prior to the formal maintainability demonstration (Type **4** of Fig. 6-1).

The primary objective in the use of any maintenance diagnostic technique is to obtain rapid and positive identification of the equipment performance operational readiness condition, or malfunction isolation in cases of failure of the equipment to meet the performance specifications. This in turn contributes to reducing system downtime for maintenance. There are three basic diagnostic techniques available (Ref. 6, Chapter 5):

1. ***Manual.*** Basically a trial-and-error effort requiring technician decision-making and using external meters, oscilloscopes, etc.

2. ***Semi-automatic.*** Representing one or more steps toward automation of fault isolation and involving sufficient internal test units and indications to assist the technician's decision and to direct the next step to be taken. Normally such systems fall short of complete elimination of dependence on the direct participation of technicians in the diagnostic decision.

3. ***Automatic.*** Eliminates the technician's participation in the fault location, and, upon failure, the system is fitted with automatic techniques and switching to a diagnostic mode which, by means of internal circuitry, isolates and identifies the malfunctioning item to a repair-by-replacement level.

Trends during the late 1960's were toward both standardization and modularization of equipment to accelerate employment of automatic and semi-automatic diagnostic techniques. This trend is essential in order not to compromise mission availability while using the available 95 percentile-type technician's skill to materially reduce overall life-cycle support resources and cost. Initial procurement and development costs may rise, but the life-cycle support cost should decrease. Annual life-cycle support costs from 1956 to 1970 rose by at least one order of magnitude over the initial procurement costs, due to higher skill requirements, associated diagnostic support equipment requirements, and higher downtime for maintenance.

6-2.1 TYPES OF TEST APPROACHES

Functional Tests. These are tests that simulate normal operating conditions as closely as possible in order to establish the state of readiness of the equipment to perform its mission. The test equipment simulates all significant inputs, evaluates all final outputs against a set of predetermined standards, deduces the incidence of any probable faults, and indicates to the test equipment operator the system readiness state. Such tests are performed under normal thermal, mechanical, and electrical conditions. This type of test is usually the first and last step in a maintenance action and is called system/equipment check-out. It is used to verify the condition of the equipment before further maintenance actions are required and to verify that a corrective action has been accomplished. Functional tests can be oriented from a system level down to a replaceable subassembly level, depending upon the maintenance concept level being evaluated. Functional tests are applied and required during all evaluation points shown in Fig. 6-1.

Marginal Tests. These tests attempt to isolate potential problem areas by simulation of abnormal operating conditions. They are performed by supplying unrelated stimuli to a system while it is subjected to severe environmental conditions such as vibration, extreme heat, or lowered power supply voltages. Marginal testing has its greatest value when it is used in support of fault prediction, where it identifies many incipient failures resulting from abnormal environmental and operating conditions. It is most useful when the item under test contains a number of identical circuits or parallel paths, with only one or two exposed to severe conditions. However, when items tested follow complex paths, simple marginal testing does not yield useful results both because (1) the complexity of the paths or circuits tends to mask the test results and (2) the variety of circuits and paths requires simulation of many sets of marginal conditions. Therefore, functional testing is most desirable for checking equipment with minimal test equipment, or for either manual or automatic/semi-automatic equipment. Marginal testing is most desirable during the first three evaluation points shown in Fig. 6-1.

Regarding the use of system versus component testing (functional and/or marginal), a system test is more useful and comprehensive. No amount of component testing will ensure definitely that a system will actually perform as specified. System testing is required for this assurance. Component testing usually is employed only to verify that a replaceable component is functioning satisfactorily before replacement or to verify that a replaced item is faulty. Component testing is rarely used as a prime test technique and should only be used as an adjunct to system testing.

Static Tests. In this test, a series of intermittent but sequenced input signals is fed in and the output indications are monitored to measure the operation of an item. This test provides information on the transient behavior of the item although in practice, the transient response is often ignored in favor of a steady static response. A static test is simple and easy to perform. It usually establishes a confidence factor but cannot go beyond that, whereas a dynamic test yields much more information. Static tests usually are applied and evaluated during the first three phases shown in Fig. 6-1.

Dynamic Test. This involves application of a continuous input signal, where the output signals are noted and analyzed to determine whether the item system requirements are met. A test of this type generally simulates a typical mission of the system being tested; thus every item or major subsystem is checked during this test. The dynamic test reveals such additional information as integration rates, phase characteristics, and frequency responses for either typical or boundary inputs. Dynamic tests are preferably used during formal maintainability test (Phase 4, Fig. 6-1, as well as Phases 5 to 8). They also are applied during the first three phases to verify test effectiveness of the design experiment.

Open-loop Tests. Open-loop (or closed-loop) tests are a further refinement of static and/or dynamic tests. The open-loop test provides a zero intelligence feedback to the item being tested. In this type, a system response is evaluated as the item end-product, and no adjustment of the stimuli is made; i.e., a preset stimulus is fed in, and the response is independently evaluated. Open-loop testing usually provides more useful maintenance information than closed-loop testing, because a direct observation of the system transfer function is made, free of the modifying influence of feedback. Thus it provides a ready measure of any degradation of item performance. Also, the open-loop test is simpler and cheaper than closed-loop, because it has no requirements for an analog computer to supply intelligence feedback. Because there is no possibility of test instability, it is the most desirable test for maintenance purposes. Its use is applied during Phase 4 testing (Fig. 6-1), as well as during other phases beyond Phase 4. It is used also in the first three phases in connection with evaluation of the design experiments to verify or gather data for maintainability design purposes.

Closed-loop Tests. In this test, the stimulus is continuously adjusted as a function of the response. Adjustment is computed on the basis of equipment behavior under the test conditions. Closed-loop tests provide engineering information which is useful in evaluating performance, effectiveness of design, adequacy of tolerances, and related characteristics. It is rarely used for field maintenance except for research and development. Closed-loop circuits or paths in equipment designing are the most difficult to maintain and diagnose whenever failures occur in the loop. When a high degree of accuracy is required for a complex performance and when test points radiate noise levels which would degrade the performance, closed-loop tests are very useful. The diagnostic routines usually involve lengthy functional test runs to verify performance, because of inability to measure or stimulate input and receive adequate output tolerance responses. In many instances complex computer programming of stimulus problems have to be recognized in order to verify the performance state of readiness by test (either built-in or external).

In using various test approaches and techniques, all maintainability qualitative characteristics are affected—such as accessibility of test points, optimum placement of parts, test indicators, built-in test features, technician skill requirements, ancillary support resource requirements, modular construction, standardization, and item throw away or disposal upon failure. Therefore, the maintainability engineer must evaluate the test techniques in relation to those qualitative characteristics, and the associated effects on the equipment design and trade-off decisions in order to optimize the maintainability goals and objectives.

For instance, ready accessibility of test points affects the test approaches. The effect depends upon the different needs at the various maintenance concept levels of actions. For example, functional tests at the organizational level require end-to-end input-output accessible test points at the line replaceable subsystem-module level (LRU) to measure the readiness status. This is especially true if use of external test equipment is necessary or built-in test points interface to receive internal stimuli and readouts for automatic test techniques. On the other hand, if the replaceable item repair is performed at the intermediate level, external test points must be accessible to diagnose replaceable items up to the replaceable unit, using static or marginal test techniques. This in turn affects the optimum packaging of replaceable items within the replaced unit.

Closed-loop circuit testing requires many design engineering decisions regarding the circuit and test point accessibility in a modular design, so as to make functional static and dynamic testing possible on replaceable items; open-loop testing requires accessible test points for independent evaluation of the readiness state response.

Regardless of the test type used, the design decision must assure that the comprehensive test is adequate to verify whether a system is operating properly within its design specification and whether the related maintainability design characteristics compromise' the test techniques selected.

Therefore, all maintainability test and demonstrations in all phases shown in Fig. 6-1 must include evaluation of the test techniques by observation of the technician's procedures and proficiency. These observations of the time to perform check-out and diagnostic maintenance tasks and the effects of built-in design qualitative characteristics provide a basis for measuring effectiveness of the maintenance tasks and for corrective action to improve effectiveness. The diagnostic time is a key factor in the achievement of maintainability objectives because, historically, such times are the largest single contributor to maintainability design leading toward dynamic, automatic, and functional test techniques instead of static, manual and/or semi-automatic, and marginal test techniques.

6-2.2 USER-SERVICE TESTS

The user-service test is the most desirable type of formal maintainability demonstration on which to base accept/reject decisions (Refs. 2-6). Such evaluations constitute a statistical approach to the true environment in which equipment is operationally deployed and to conditions where logistic maintenance support resources are used. Although the statistical decision is based on a small sample of the equipment, the decision is made in view of the specified risks, and confidence and accuracy limits of the maintenance task sample. It reflects a close simulation of the system design for maintainability, even though it is not 100% reliable. However, the sampling process of the simulated maintenance tasks does represent randomness of the expected deployed use and provides an adequate simulation within the practical life cycle program development, cost, scheduling, risks, confidence, and accuracy limits.

Current Army procedures require the following phases of user-service testing for maintainability quantitative-qualitative accept/reject decisions; in each phase, the decision-making evaluation is used to establish the design in maintainability goals and objectives:

1. ***Development Tests*** (DT) are made under simulated operational environments, except that technicians are usually more skilled and the support resources are limited due to the nature of the test costs and schedules. It is primarily an engineering design proof test.

2. User Tests (UT) are made under simulated user operational environments using service operators and technicians. They employ the maintenance support resources (personnel, support equipment, spares, technical data, facilities, etc.) which have been developed. In such a situation, actual experience in the operational environment can be observed, recorded, and subsequently analyzed to reflect the true approximate representation of the design in maintainability quantitative and qualitative requirements. Usually the service tests involve only the organization level (operation/organization technician) and the intermediate level (direct support and general support technicians) maintenance tasks. The depot level maintenance normally uses the contractor support facilities, especially in the early deployment phases; demonstration involves maintenance repair tasks using contractor logistic support resources and is performed at contractor facilities.

Technician training must be provided to ensure that the skills necessary to perform the maintenance tasks during the service test demonstrations are adequate for the operationally deployed system.

3. Physical Tear Down is made by the customer to evaluate the capability of equipment tear down to the lowest replaceable item. Normally this test is done under controlled conditions and environment using contractor-customer production or experimental technicians and the associated logistic support resource tools (custom designed or standard). The principal quantitative parameter for maintainability that can be measured is the item replacement task time. Diagnostic task times are not usually involved. This evaluation has the additional advantage of observing such qualitative maintainability design characteristics as accessibility, packaging, or attachments. Also, measurement of the test approaches can be included with little extra cost. The test has the further advantage of evaluating depot level maintenance requirements as well as the other two levels (organizational and intermediate levels), because complete tear down of the system/equipment takes place.

The hardware models used for these three tests usually are engineering prototype models or early production items. In the "make before buy" ("fly before buy") process, the models tested are prototypes that simulate the production model as nearly as possible. The same applies to other logistic maintenance support items, such as technical publications and support equipment.

6-2.3 TASK SELECTION

The basic method of stratification and apportionment of tasks to be demonstrated is presented in detail in the references cited in par. 6-1.1. In relation to the test techniques earlier cited, par. *6-2,* Test Approaches, tests to be used must adhere to the principle that diagnosis and check-out be capable of determining the maintenance functions required at the respective levels of maintenance. For example, if a line replacement unit is replaced at the organizational level, the diagnostic (fault isolation-location) test task must "troubleshoot" the fault with sufficient confidence to assist the technicians in determining that it is essential to replace the unit in order to correct the system malfunction. The check-out task must indicate whether the equipment state of readiness requires an item replacement due to equipment failure. Thus the check-out test must indicate that repair has been accomplished correctly. Simulated faults are not to be inserted if they are potentially hazardous or may damage the equipment. For example, a fault would not be simulated in an aircraft landing gear if it would create a hazard in the landing of an aircraft with potential damage to equipment and injury to personnel.

Careful attention should also be given to choice of samples for the tests. This applies especially to technician skill levels. Random selection techniques should be used in choosing the sample. The sample is chosen so that it correctly reflects the percentage makeup (in terms of skill levels) of the original population. Care must be taken when faults are induced into the system that the maintenance technician is not given information that normally would not be available.

SECTION II

METHODS OF TESTING

6-3 TEST TECHNIQUES

Details of the methods of testing for maintainability goals vary according to the desired objectives during the various phases of the system/equipment life cycle shown in Fig. 6-1, and cited in the test approaches, par. 6-2.

The maintainability design and quantitative goals (time) rely on the reliability quantitative determinations of the need for maintenance (failure rates for corrective maintenance, and scheduled rates for preventive maintenance). Unfortunately, due to costs, schedules and test times, it is not practical to test or operate equipment in a deployed population under combat or peace-time environments or for the designed useful life. In some instances, the failure or wear-out of components/parts for subassemblies/assemblies/modules may never occur.

In some cases, where the selected maintenance tasks for small systems of less than 50 parts involve a limited number of potential tasks of simple diagnosis and replacement, all maintenance tasks can be fault simulated, and the maintenance times observed and weighted by the predicted failure rates to get the mean time to repair the system.

In most cases, the potential maintenance actions are many, and simulation sampling must be used. If, for example, over 6000 different organizational level tasks exist for the C5A aircraft, the sample size for maintainability demonstration would be 228 (Ref. 10), and simulation of faults for the randomly selected tasks would be necessary.

Also contributing to the methods of testing are considerations of the complexity of the various types of equipment and associated technician skills (all mechanical, electro-mechanical, electrical, electronic, optical, or any combinations of types). The techniques are fundamentally the same, varying only to the extent of fault simulation.

In all cases of fault simulation, the safety of personnel and potential damage to system/equipment are considered. Safety fault tree analysis could be the basis for determining simulations. Also the Failure Modes-Effects-Criticality-Analysis is used to evaluate and determine fault simulations.

Fault simulations are determined during the predemonstration phase in accordance with the plan for selection of task and sample. The sample demonstration data sheets must indicate the type of failure to be simulated. A technician (assigned to assist the test director and maintainability test monitor) induces the failure into the equipment during the absence of the technician who will perform the maintenance. The fault of the "bug" is checked to assure that it produces a malfunction in the equipment. The technician who performs the maintenance action is then assigned to the maintenance tasks. He prepares for the tasks by reviewing the reported symptom, using the available technical manual, and looking for approved alternate procedures. He conducts a visual inspection and sets up the equipment, its allied support test equipment and facility power input, and starts the maintenance actions required. He first performs a functional check-out test and verifies the symptom. He then performs the required maintenance tasks until he has completed the action and the check-out functional tasks. This will prove that repair action has been accomplished and that the items are ready for operational use. All actions are timed and recorded on the data worksheets. In most cases the effects of the maintainability features built into the equipment are noted, especially where the times are excessive. In all cases where the predicted times are exceeded, the reason must be recorded in order to establish a basis for corrective action. The following are typical fundamental maintenance actions to be observed, monitored, and recorded, as a minimum:

1. Preparation and visual inspection time
2. Functional check-out time
3. Diagnostic time:
   a. Fault locate
   b. Fault isolate
4. Repair time:
   a. Gain access
   b. Remove and replace
   c. Adjust, align, calibrate

d. Close access

5. Clean, lubricate, service time

6. Functional check-out of the repair action.

The introduction of fault simulation may vary but the following are typical examples:

1. Electrical-Electronic:
   a. Open or short circuits
   b. Insertion of failed components
   c. Misalignment or misadjustment
   d. Insertion of broken connector pins or springs
2. Mechanical-Electro Mechanical:
   a. Insertion of broken springs
   b. Use of worn-out bearings, failed seals, broken relays, or open-short coils
   c. Misalignments or misadjustments
   d. Insertion of failed indicators, broken or worn gears, missing keys
   e. Use of failed or worn components
3. Optical Systems:
   a. Use of dirty reflectors
   b. Misalignment or misadjustment
   c. Insertion of broken parts or components
   d. Use of faulty sensors, faulty indicators.

Scheduled preventive actions are in accordance with established step-by-step procedures of routine visual, mechanical measurement; electrical functional check-out; cleaning; and lubrication. In a few cases they involve scheduled item replacement actions and associated check-out actions. Therefore fault simulation 'is rather limited because in most cases scheduled maintenance tasks are on a calendar, mileage, or rounds basis and 100% are demonstrable. For routine tear down and overhaul task selection, sampling usually is employed.

6-4 SAMPLING AND SAMPLE SIZE SELECTION

Sampling is the process of gathering data about a population in order to infer some of its characteristics. Sampling is performed because it may be impossible, too expensive, or too time consuming to observe the entire population.

In many cases considerable knowledge about a population can be inferred from a properly chosen sample of moderate size. In the typical situation, it is possible to sample only a small part of the population. Thus, it is important to exercise care in the choice of the sample to assure that it is representative of the population. The sample size often will be limited by lack of opportunity, time, or money, and these constraints must also enter the selection process.

To illustrate the latter point, suppose one wishes to infer the mean time to remove and replace a component. In order that the inference be valid over the entire population of repairmen, it is necessary that the sample include repairmen of varying. skill levels and experience, i.e., be representative of the total population. It would be misleading to take all data on only one individual, or solely on individuals of one particular skill level. Information on the composition of the population of repairmen may be available from personnel records. If not, it may be necessary to sample the population of repairmen to gain knowledge of its composition.

How many observations should be made, and how should they be collected, in any given situation? The answers to these questions depend on the amount of knowledge of the population, the desired accuracy, and the population characteristic(s) we wish to infer. For example,'does one know the distributional form of the relevant random variable(s), and if so, the value of any of the parameters of the distribution? How close should the estimates of the population quantities be to these quantities? Which population characteristic is being estimated—the mean, the median, the variance—and (again) how accurately is it to be estimated?

If it were known that all elements of a population were identical, then, clearly it would suffice to make only one observation in order to know all about the population. The less homogeneous a population, the more observations must be made in order to achieve a certain level of information from inferences. The information content of an estimate is measured (inversely) in terms of its variance. Variance, in turn, is usually inversely proportional to the size of the sample. Thus, the information content of an estimate is directly proportional to the size of the sample upon which it is based.

6-4.1 STATISTICAL METHODS IN SAMPLE SELECTION

In this paragraph, methods of determining minimum sample sizes for estimating a proportion of successes and the mean of a normal distribution will be discussed, as will sample size selection for the lognormal distribution. It will be seen that sample size determination requires the specification of a desired accuracy and confidence. The relationship among these three quantities is such that if two of these are given, one can solve for the third. Alternatively, fixing the value of only one of them entails a trade-off between the other two.

1. *Estimating a Proportion of Successex*

First, consider sampling without replacement from an infinite population of items. Each item selected will be judged according to a dichotomous criterion and the two possible choices will be called "success" and "failure", where success or failure consists of the occurrence or nonoccurrence, respectively, of the event of interest.

We cite the following facts from statistics. Given a sample of size n in which s successes have been observed, the maximum likelihood estimate of p, the "true" proportion of successes in the population, is $\hat{p} = s/n$. For large values of n the distribution of $\hat{p}$ is approximately normal with mean p and variance $\sigma_{\hat{p}}^2 = p(1 - p)/n$. Using this approximation, determine first the sample size n needed to assure that $\hat{p}$ is within δ on either side of p, with probability $1 - \alpha$. (Subsequently n will be determined for $\hat{p}$ on one side of, and within δ of p, with probability $1 - \alpha$).

It is required to determine the sample size n so that

$$P(|\hat{p} - p|) \leq \delta) = 1 - \alpha \tag{6-2}$$

This probability statement is equivalent to

$$P\left(\frac{|\hat{p} - p|}{\sigma_{\hat{p}}} \leq \frac{\delta}{\sigma_{\hat{p}}}\right) = 1 - \alpha \tag{6-3}$$

Invoking the approximate normality of $\hat{p}$ with mean p and standard deviation $\sigma_{\hat{p}} = \sqrt{p(1 - p)/n}$, the quantity $(\hat{p} - p) / \sigma_{\hat{p}}$ is approximately a standard normal variate (i.e., mean zero, variance one). The probability statement can be portrayed as shown in Fig. 6-2.

This shows that $\delta/\sigma_{\hat{p}}$ is equal to $z_{1-\alpha/2}$, the $100(1 - \alpha/2)$ percentile of the standard normal distribution. Substituting $\sqrt{p(1-p)/n}$ for $\sigma_{\hat{p}}$ in expressing this equality and solving for n, we obtain

$$n = (z_{1-\alpha/2}^2)p(1 - p)/\delta^2 \tag{6-4}$$

Since the value of p is not known, (else we would not be estimating it!), the following alternatives exist for estimating n:

a. If nothing is known about the value of p, or if it is believed that p is near 1/2, take

$$n = (z_{1-\alpha/2}^2)/(4\,\delta^2) \tag{6-5}$$

b. If it is believed that p is less than 1/2, substitute the largest reasonable guess for p in Eq. 6-4 for n.

c. If it is believed that p is greater than 1/2, substitute the smallest reasonable guess for p in Eq. 6-4 for n.

The statements 1a, 1b, and 1c follow directly from the fact that $p(1 - p)$ is a parabola taking the value 0 at $p = 0$ and $p = 1$, and attaining its maximum of 1/4 at $p = 1/2$. The most conservative rule to follow is 1a; rules 1b and 1c yield smaller values of n than rule 1a.

Example 6-1 Determine the sample size n necessary to assure that $\hat{p}$ is within 0.05 on either side of p with probability 0.90 if

a. Nothing is known about the true value of p.

b. It is believed that $p \leq 0.40$.

c. It is believed that $p \geq 0.80$.

Solution: $\alpha = 0.10$, $1 - \alpha/2 = 0.95$, and $z_{0.95} = 1.645$; $\delta = 0.05$. Thus, rounding up to the nearest integer:

a. $n = (1.645)^2/[4(0.05)^2] = 271$

b. $n = (1.645)^2\,(0.4)\,(0.6)/(0.05)^2 = 260$

c. $n = (1.645)^2\,(0.8)\,(0.2)/(0.05)^2 = 174.$

Now determine the sample size n needed to assure that $\hat{p}$ is within δ on *one side of p* with probability $1 - \alpha$. There are two cases: $\hat{p} > p$ and $\hat{p} < p$. We first derive the result for $\hat{p} > p$. Here we want to be assured that we don't overestimate p "too much", but are not concerned about an underestimate. The derivation is analogous, and the result is identical, for $\hat{p} < p$. In that case we want to be protected against "too large" an underestimate of p.

We want, for $\hat{p} > p$,

$$P = (\hat{p} - p < \delta) = 1 - \alpha \tag{6-6}$$

This is equivalent to

$$P = \left(\frac{\hat{p} - p}{\sigma_{\hat{p}}} < \frac{\delta}{\sigma_{\hat{p}}}\right) = 1 - \alpha \tag{6-7}$$

from which it follows that $\delta/\sigma_{\hat{p}}$ is $z_{1-\alpha}$ the $100(1 - \alpha)$ percentile of the standard normal distribution. Substituting $\sqrt{p(1-p)/n}$ for $\sigma_{\hat{p}}$ into this statement of equality and solving for n yields

$$n = (z_{1-\alpha}^2)[p(1 - p)/\delta^2] \tag{6-8}$$

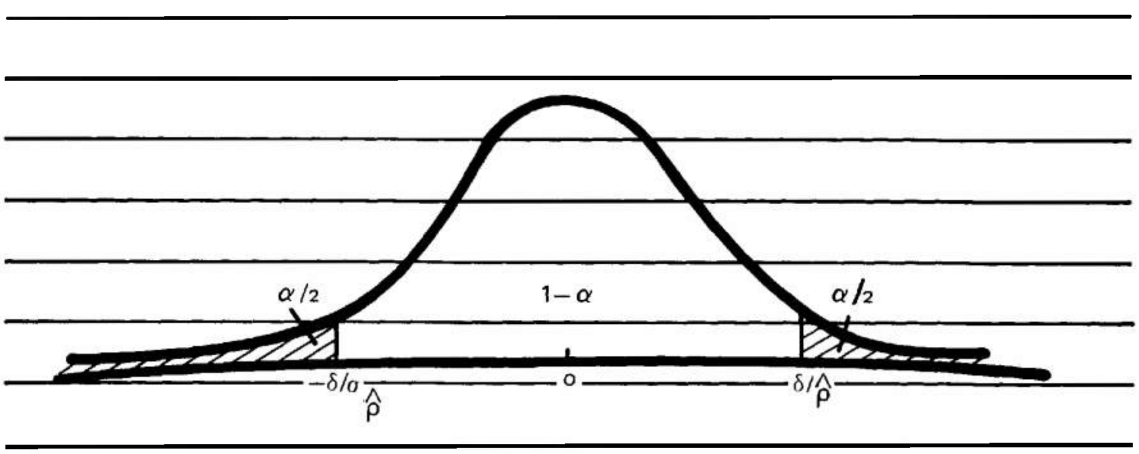


Figure 6-2. Probability Statement

Once again the alternatives a, b, and c previously set down must be followed. The most conservative choice is

$$n = (z^2_{1-\alpha})/(4\delta^2) \tag{6-9}$$

Example 6-2: Determine the sample size n necessary to assure that $\hat{p}$ does not overestimate p by more than 0.05 with probability 0.90 if

a. Nothing is known about the true value of *P*.
b. It is believed that p **i** 0.40.
c. It is believed that p ≥ **0.80.**

Solution: **a = 0.10,** 1 − **a** = 0.90, and $z_{0.90}$ = **1.282;** 6 = **0.05.** Thus, rounding up to the nearest integer:

a. n = **(1.282)^2/[4(0.05)^2] = 165**
b. n = **(1.282)^2 (0.4) (0.6) / (0.50)^2 = 158**
c. n = **(1.282)^2 (0.8) (0.2) /(0.05)^2 = 106.**

The sample sizes indicated in the preceding examples may be quite expensive or time-consuming to obtain in some situations. They are, however, the smallest sample sizes that can be guaranteed to provide the specified accuracy **(i.e.,** 6) and confidence [i.e., **(1 − a)** or **(1 − α/2)]** requirements. If the sample size n given by one of the preceding equations cannot be made available, this is an indication that an unrealistic (or, at least, unrealizable) specification of accuracy (6) and confidence **[(1 − a)** or **(1 − α/2)]** has been made.

A realistic approach in this case is to consider the trade-off between accuracy and confidence achievable for the largest sample size that can be made available. This trade-off is easily seen, for example, in **Eq. 6-5,** dealing with Example **6-la.** Suppose one could make available a sample of size n = **100.** Substituting **this** value into **Eq. 6-5** yields

$$400\,\delta^2 = z^2_{1-\alpha/2} \text{ or } \delta = (z_{1-\alpha/2})/20 \tag{6-10}$$

The relationship can easily be tabulated (Table 6-2) and graphed (Fig. **6-3).**

Fig. **6-3** shows, for example, that with a sample of size 100 in Example **6-**1a, if one wants an accuracy of δ = 0.07, he can achieve it with a confidence of 1 − α = **84%.** Similarly a confidence of 1 − α = 95% will provide an accuracy of 0.0980.

The same kinds of tables and graphs can be prepared for other sample sizes and the other cases.

2. *Estimating the Mean of a Normal Distribution*

Consider sampling from a normal distribution with unknown mean. Recall the following facts from statistics. Given a sample of size n, x_1, x_2, . . . , x_n, from a normal distribution with mean μ and variance σ^2, the maximum likelihood estimate of μ is $\bar{x}$, the sample mean. The distribution of $\bar{x}$ is normal with mean μ and variance σ^2/n. These facts are now used to determine the sample size n needed to assure that $\bar{x}$ is "close" to μ. In case a is known, we can ask for the n which guarantees $\bar{x}$ is within 6 on either side of μ with probability 1 − a, or $\bar{x}$ **on** one side of, and within 6 of, μ with probability 1 − **a.** If a is unknown, the situation is somewhat more complicated. It has been shown by Dantzig (Ref. **11)** that it is not possible with a single sample of fixed size to construct a confidence interval of preassigned length and given confidence coefficient for the mean of a normal distribution when the variance is not known. (However, a two-sample procedure can achieve this.) This means that, for a unknown, a different criterion is necessary for determining n. We can determine n so that $\bar{x}$ is within some multiple ϵ of the population standard deviation a of μ with probability 1 − **a.** Finally, we present a method for the choice of n when the sample mean is to be within a given percent of the population mean.

Given a known; $\bar{x}$ on either side of μ.

In this case we require that

$$P(|\bar{x} - \mu| \leq \delta) = 1 - \alpha \tag{6-11}$$

This probability statement is equivalent to

$$P\left(\frac{|\bar{x} - \mu|}{\sigma/\sqrt{n}} \leq \frac{\delta}{\sigma/\sqrt{n}}\right) = 1 - \alpha \tag{6-12}$$

which can be portrayed graphically as shown in Fig. **6-4.**

This shows that $\sqrt{n}\delta/\sigma$ is equal to $z_{1-\alpha/2}$, the **100(**1 − α/2) percentile of the standard normal distribution. Expressing this equality and solving for n yields

$$n = (z^2_{1-\alpha/2})(\sigma^2/\delta^2) \qquad (6\text{-}13)$$

Example 6-3: Determine the sample size n necessary to assure that $\bar{x}$ is within 1 unit on either side of μ with probability 0.90. It is known that $\sigma = 4$.

Solution: $a = 0.10$, $1 - \alpha/2 = 0.95$ and $z_{0.95} = 1.645$; $6 = 1$, $\sigma = 4$. Thus from Eq. 6-13, $n = (1.645)^2 (4)^2 /(1)^2 = 44$, rounding up to the nearest integer.

Given σ known; $\bar{x} > \mu$, $\bar{x} - \mu < 6$.

Here we want to be assured that we don't overestimate μ "too much", but are not concerned about an underestimate. (The derivation is analogous, and the result is identical, for $\bar{x} < \mu$. In that case we want to be protected against "too large" an underestimate of p.) We require

$$P(|\bar{x} - \mu| < a) = 1 - \alpha \qquad (6\text{-}14)$$

This is equivalent to

$$P\left(\frac{|\bar{x} - \mu|}{\sigma/\sqrt{n}} < \frac{\delta}{\sigma/\sqrt{n}}\right) = 1 - \alpha \qquad (6\text{-}15)$$

from which it follows that $\sqrt{n}\delta/\sigma$ is $z_{1-\alpha}$, the 100 (1 − a) percentile of the standard normal distribution. Expressing this an an equality and solving for n yields

$$n = (z^2_{1-\alpha})(\sigma^2/\delta^2) \qquad (6\text{-}16)$$

Example 6-4: Determine the sample size n necessary to assure that $\bar{x}$ does not overestimate μ by more than 1 unit with probability 0.90. It is known that $\sigma = 4$.

Solution: $a = 0.10$, $1 - a = 0.90$, and $z_{0.90} = 1.282$; $\delta = 1$, $\sigma = 4$.

Thus $n = (1.282)^2 (4)^2/(1)^2 = 27$, rounding up to the nearest integer.

Given σ unknown; $\bar{x}$ on either side of μ.

In this case an appropriate criterion is to require that

$$P(|\bar{x} - \mu| \leqslant \epsilon\sigma) = 1 - \alpha \qquad (6\text{-}17)$$

where E is some multiple of the population standard deviation σ.

This probability statement is equivalent to

$$P\left(\frac{|\bar{x} - \mu|}{\sigma/\sqrt{n}} \leqslant \sqrt{n}\,\epsilon\right) = 1 - \alpha \qquad (6\text{-}18)$$

TABLE 6-2.
TRADE-OFF BETWEEN ACCURACY AND CONFIDENCE FOR A SAMPLE OF SIZE n = 100 FOR EXAMPLE 6-1a

Confidence (1 − a)	a!	$1-\alpha/2$	$z_{1-\alpha/2}$	δ
0.75	0.25	0.875	1.15	0.0575
0.80	0.20	0.90	1.28	0.0640
0.85	0.15	0.925	1.44	0.0720
0.90	0.10	0.95	1.645	0.08225
0.95	0.05	0.975	1.96	0.0980
0.99	0.01	0.995	2.575	0.12875

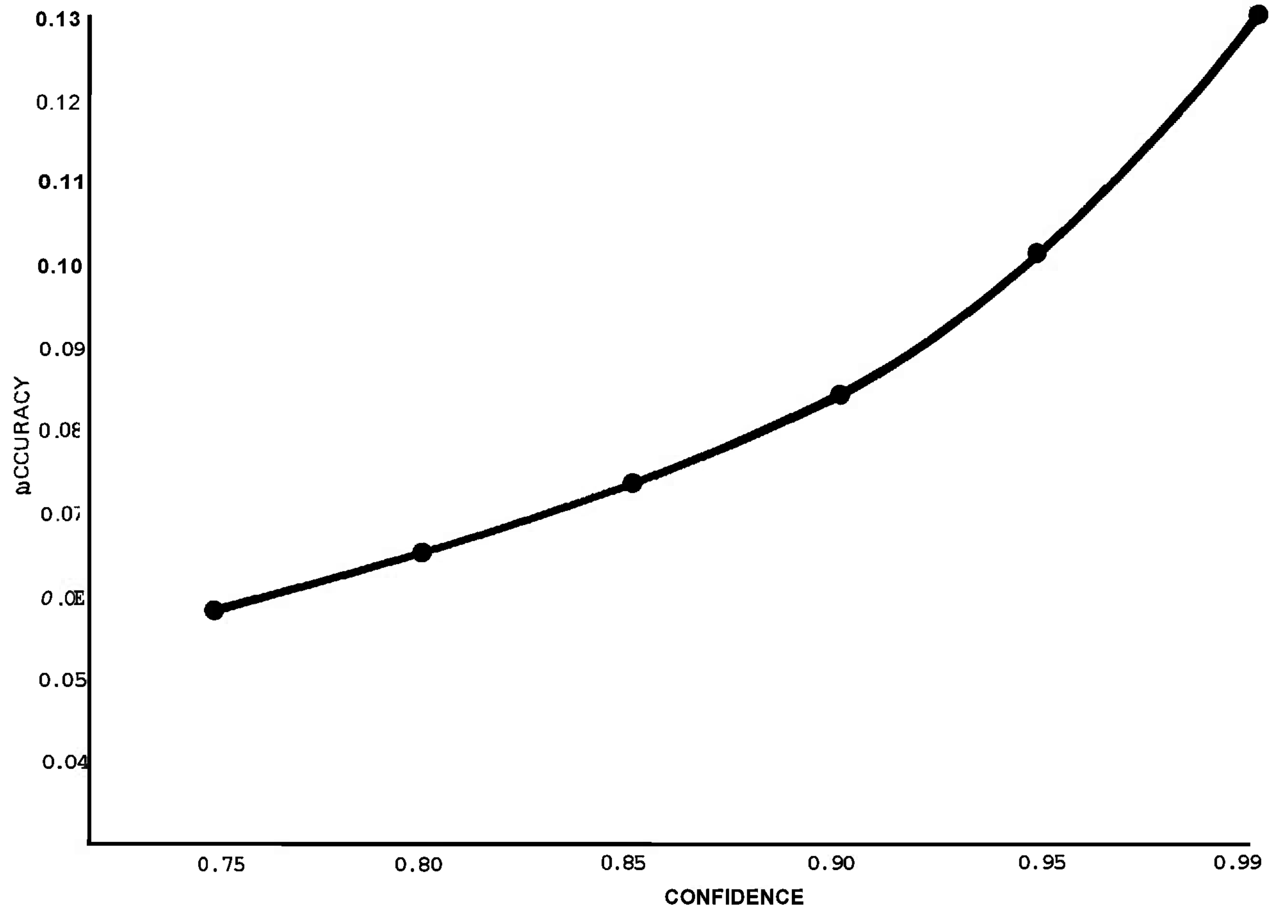


Figure 6-3. Trade-off Between Accuracy and Confidence

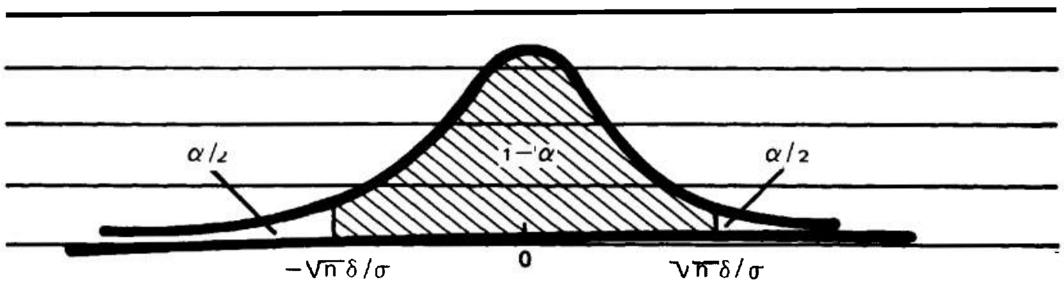


Figure 6-4. Graphic Presentation of Eq. 6-12

so that $\sqrt{n}\epsilon$ is equal to $z_{1-\alpha/2}$, the 100(1 − $\alpha/2$) percentile of the standard normal distribution. Expressing this equality and solving for n yields

$$n = (z^2_{1-\alpha/2})/\epsilon^2 \quad (6\text{-}19)$$

Example 6-5: Determine the sample size n necessary to assure that $\bar{x}$ is within 0.25 σ on either side of μ with probability 0.95.

Solution: $\alpha = 0.05$, $1 - \alpha/2 = 0.975$, and $z_{0.975} = 1.96$; $\epsilon = 0.25$. Thus from Eq. 6-19, $n = (1.96)^2/(0.25)^2 = 62$, rounding up to the nearest integer.

Given σ unknown; $\bar{x} > \mu$, $\bar{x} - \mu < \epsilon\sigma$.

Our requirement is that

$$P(|\bar{x} - \mu| < \epsilon\sigma) = 1 - \alpha \quad (6\text{-}20)$$

This probability statement is equivalent to

$$P\left(\frac{|\bar{x} - \mu|}{\sigma/\sqrt{n}} < \sqrt{n}\epsilon\right) = 1 - \alpha \quad (6\text{-}21)$$

so that $\sqrt{n}\epsilon$ is equal to $z_{1-\alpha}$, the 100 (1 − α) percentile of the standard normal distribution. Expressing this equality and solving for n yields

$$n = (z^2_{1-\alpha})/\epsilon^2 \quad (6\text{-}22)$$

Example 6-6: Determine the sample size n necessary to assure that $\bar{x}$ does not overestimate μ by more than 0.25σ with probability 0.95.

Solution: $\alpha = 0.05$, $1 - \alpha = 0.95$, $z_{0.95} = 1.645$; $\epsilon = 0.25$. Thus from Eq. 6-22, $n = (1.645)^2/(0.25)^2 = 44$, rounding up to the next integer.

Given σ known. Estimating μ to within a given percentage.

This method is described in MIL-HDBK-472 (Ref. 12). We follow our previous notation.

The idea is to determine n so that for given probability $1 - \alpha$, and percentage k (expressed as a decimal) of the population mean μ.

$$P(|\bar{x} - \mu| \leq k\mu) = 1 - \alpha \quad (6\text{-}23)$$

In order that this probability statement be correct, it is necessary that

$$k\mu = (z_{1-\alpha/2})\sigma_{\bar{x}} = (z_{1-\alpha/2})\sigma/\sqrt{n} \quad (6\text{-}24)$$

Solving for n yields

$$n = [(z_{1-\alpha/2})\sigma/(k\mu)]^2 \quad (6\text{-}25)$$

While this method does not require separate knowledge of the population mean μ and population standard deviation σ, it does require knowledge of the ratio σ/μ. This quantity is called the population ***coefficient of variation*** and is usually denoted by V. (MIL-HDBK-472 denotes it C_X) It is claimed that certain classes of equipment have characteristic values of V. To the extent that this is true, the formula for n, which can be rewritten as

$$n = [(z_{1-\alpha/2})V/k]^2 \quad (6\text{-}26)$$

is useful. For example, MIL-HDBK-472 quotes a value of $V = 1.07$ for ground electronic equipment. Of course the formula will be useful in any situation in which V is known.

A nomograph solving Eq. 6-26 for n for a range of values of V and k, and for 90% confidence (i.e., $1 - \alpha = 0.90$) is given is Fig. 3-2, page 3-10 of MIL-HDBK-472. (It should be noted that MIL-HDBK-472 ***incorrectly*** states this nomograph is for 95% confidence). However, the equation for n is easily solved without the nomograph.

Example 6-7: It is known that the coefficient of variation V for a normal distribution is 0.5. Find the sample size necessary so that the absolute difference between the sample mean and the population mean does not exceed 10% of the population mean with probability 0.95.

Solution: We must choose n so that

$$P(|\bar{x} - \mu| \leq 0.10\,\mu) = 0.95 \tag{6-27}$$

The value of n from Eq. 6-26 is

$$n = (z_{0.975}V/k)^2 = [(1.96)(0.5)/0.10]^2 = 96 \tag{6-28}$$

3. Estimation For the *Lognormal* Distribution

A positive random variable X has the lognormal distribution with parameters ν and ω, if the natural logarithm of X, i.e., $\ln X$, has a normal distribution with mean ν and standard deviation ω. The probability density function $f(x)$ of X is

$$f(x) = (\omega x\sqrt{2\pi})^{-1} \exp[-(\ln x - \nu)/(2\omega)] \tag{6-29}$$

where

ν = mean of the natural logarithms of X_i

$\hat{\omega} = \sqrt{\Sigma(\ln X_i - \hat{\nu})^2/(n-1)}$

$\hat{\nu}$ = estimated value of ν

The mean, median, and variance of the lognormal distribution are:

mean: $\mu = e^{\nu + \omega^2/2}$

median: $m = e^{\nu}$

variance: $\sigma^2 = e^{2\nu + \omega^2}(e^{\omega^2} - 1)$

The median is often used as a measure of central tendency for the lognormal distribution because it does not involve ω. The 100pth percentile of the lognormal distribution is given by $e^{\nu + z_p\omega}$ where z_p is the 100pth percentile of the standard normal distribution.

Many maintainability demonstration plans are couched in terms of the 90th or 95th percentile of the lognormal distribution. See, in particular, MIL-STD-471 (Ref. 2). These plans are sequential in nature, not of fixed sample size. (We describe them in par. 6-7.)

Suppose one wanted to estimate the median m of a lognormal random variable X on the basis of a sample $X_1, X_2, \ldots, X_n$. Note that $\nu = \ln m$ is the population mean of the normal random variable $\ln X$, so that an estimate of m is $\hat{m} = e^{\hat{\nu}}$ where $\hat{\nu}$ is the sample mean of

$$\ln X_1, \ln X_2, \ldots, \ln X_n;$$

i.e.,

$$\hat{\nu} = \sum_{i=1}^{n} \ln X_i/n$$

If ω is known, then a sample size n can be determined so that, for given values of δ and α,

$$P(|\nu - \hat{\nu}| < \delta) = 1 - \alpha \tag{6-30}$$

as in Example 6-3. Specifically,

$$n = (z^2_{1-\alpha/2})(\omega^2/\delta^2) \tag{6-31}$$

The preceding probability statement, Eq. 6-30, can be written as

$$P(\hat{\nu} - \delta < \nu < \hat{\nu} + \delta) = 1 - \alpha \tag{6-32}$$

and also, in terms of $m = e^{\nu}$, as

$$P(e^{\hat{\nu}-\delta} < m < e^{\hat{\nu}+\delta}) = 1 - \alpha \tag{6-33}$$

or

$$P(\hat{m}/e^{\delta} < m < \hat{m}e^{\delta}) = 1 - \alpha \tag{6-34}$$

This interval for estimating m does not have a fixed length (indeed, its length is a random variable), but the ratio of the upper to the lower end of the interval is a constant $e^{2\delta}$. Note that it was necessary to assume knowledge of ω in order to determine the sample size. The point of this paragraph is to emphasize that one cannot determine a sample size that will allow construction of a confidence interval of predetermined confidence and length for the median of the lognormal distribution as we were able to do for the binomial and normal distributions. What is true for the median, the 50th percentile, is perforce true for other percentiles which involve ω explicitly; hence, the reason for demonstration plans for the lognormal distribution being sequential rather than of fixed sample size.

4. ***Relation Among Accuracy, Confidence, and Sample Size.***

Under the subpars. 6-4.1(1) and 6-4.1(2) we have discussed the selection of a sample size n to meet joint requirements of accuracy of an estimator and the confidence (probability $1 - \alpha$) with which that accuracy is achieved. It should be clear that in each of the models discussed there is a relation among the quantities of accuracy, confidence, and sample size such that if any two of them are specified, the third is determined.

To illustrate this point it will suffice to consider Example 6-1 in which we sought n so that, for given δ and α,

$$P(|\hat{p} - p| < \delta) = 1 - \alpha \tag{6-35}$$

We found that (Eq. 6-9)

$$n = (z^2_{1-\alpha/2})/(4\delta^2) \tag{6-36}$$

Had n and α been specified, δ would be determined by the equation

$$\delta = (z_{1-\alpha/2})/(2\sqrt{n}) \tag{6-37}$$

Had n and δ been specified, we would solve for α by first obtaining $z_{1-\alpha/2}$ from

$$z_{1-\alpha/2} = 2\delta\sqrt{n} \tag{6-38}$$

and then read $1 - \alpha/2$ from tables of the cumulative normal distribution.

6-5 REDUCTION OF TESTING

As stated previously, the equipment ideally would be exercised and operated within an actual operational environment and population using the established logistic support resources for the duration of the specified useful life. However, this is not practical because of cost and development scheduling. Fault simulation therefore should be used for selected tasks, based on statistical samplings.

Because frequency of maintenance action is a prime consideration in determining the designed-in maintainability quantitative and the qualitative design features, the maintainability engineer must be familiar with the methods and results used by reliability engineers or with maintenance historical data on similar equipment. Also, the frequency of maintenance actions affects not only maintainability but all other system effectiveness and operational design parameters. Therefore, the maintainability engineer must be prepared to participate in all trade-off studies involving reliability. During the early design development phases—once the failure rates have been predicted and the equipment reliability design has been reviewed and approved—the maintainability design parameters are concurrently established, reviewed, and approved. At this point in the equipment cycle, the baseline for the demonstrable maintainability factors is firmly established, and both the appropriate samples and statistical tests can be selected (see pars. 6-4, 6-7, and 6-8). Therefore, any test conducted by the reliability engineers to evaluate reliability predictions—overstressing, increased quantity of equipment tested, environmental testing, intermittent test results, instant failure test results, dormant failure rate determinations, wear-out failure rate, continuous operation versus intermittent operation results, sample size versus time required for adequate testing limitation—must be reviewed by the maintainability engineer. Their effect on the maintainability design parameters must be considered and included in the design for maintainability. This task is part of the maintainability design surveillance, design monitoring, design review, and associated trade study tasks. The description of the various reliability engineering methods, techniques, and statistics of reducing test time for reliability demonstration of the frequency of maintenance actions is contained in various reliability textbooks and symposium literature reports and is not included in this handbook.

6-6 DEMONSTRATION DATA

6-6.1 COLLECTION OF DEMONSTRATION DATA

The maintainability data collection systems of military procurement agencies should be used to the maximum extent possible to record data collected during maintainability demonstrations. Supplementary or unusual data collection systems may be incorporated as necessary, if approved in the demonstration plan. However, if such unusual data collection is necessary, it should be an addendum to existing systems. The primary purpose in using existing systems is to speed up evaluation and analysis of the data. In most existing data collection systems, provisions are not made for recording detailed timing of the individual maintenance tasks, but include only a gross summation of the time to perform the overall function. The data collection of the demonstration results is not the same as gathering data for historical purposes as set forth in Chapter 9.

There are four formats for supplemental data collection:

1. The detailed individual test data collection format must provide for recording, as a minimum, the following individual tasks:

a. Description of the test conducted—such as number, corrective and/or preventive maintenance, name of the recorder/observer, name of the technician performing the task, date, equipment identification, and associated resource equipment identification
b. Fault or scheduled action simulation and/or description of method of simulation—including part identification, failure symptom description, mode of detection, etc.
c. Time data observed for corrective maintenance–such as preparation, inspection, functional checkout, diagnosis (fault location and isolation), disassembly, remove and replace, reassembly, align/adjust, functional checkout, service/clean/lubricate–used to develop the prediction task step times.
d. Time data observed for preventive maintenance—such as preparation, functional checkout, adjust/align/calibrate, overhaul, scheduled replacements, service/clean/lubricate.
e. Effectiveness or discrepancy of observed qualitative features
f. Other comments.

2. The work sheet format for recording the observed results listed in the previous paragraph consists of the elements described in par. la, together with detailed elapsed times observed for elements of pars. 1c and 1d showing task number, task description, number of technicians used, active maintenance start and stop time, and administrative/logistic start and stop times.

3. The data analysis formats are adopted for the recording of the observed data and associated computations, and for charting results of the analysis (such as histograms and cumulative distribution charts). The table formats usually show the demonstration number, observed times, frequency distribution percents, logarithms of the times, class interval frequency distributions, numerical values of customer/producer risks, confidence limits and accuracy, mean time computations, etc. as dictated by the statistical method used to evaluate the observed data.

4. The summary sheet format of the test results consists of demonstration criteria of specified predicted values, achieved values, and the accept/reject decision or evaluations. The format is tailored to the statistical demonstration method data used.

The given formats and requirements apply to the formal maintainability demonstration. They are modified for other maintainability tests (Fig. **6-1** and Table **6-1)** as necessary. In all cases the'delay times not attributable to the intrinsic quantitative design parameters are recorded in order to be used as baseline factors in calculating or determining the overall effects on total down time in the determination of system/equipment "operational availability". Also, notation is included to show ancillary aids used in recording observations—such as time-lapse photography, still photography, tape recorders, chart recorders, and stop watches.

6-6.2 REPORTING, STORAGE, AND RECOVERY OF DEMONSTRATION DATA

The primary purpose of reporting of maintainability demonstration data is to "Report Demonstration Results" and their associated "Report Decision-Making Findings". A secondary purpose is to report maintenance data results for historical purposes in current maintenance data systems. Existing formats are used, properly identified and coded to enable the results and evaluations to be of value for historical purposes and to show trends in new equipment. The details of the methods, techniques, and'the purposes are given in Chapter 9. Where the quantity of data is large, computer facilities and resources should be used to the fullest extent in the evaluations. These resources should be incorporated in the demonstration planning in order to eliminate test costs and reduce associated demonstration evaluation time.

6-7 DEMONSTRATION METHODS

Demonstrating that contractually specified quantitative maintainability requirements have been met is an important aspect in the procurement of maintainable military equipment. For this reason, and to avoid the confusion arising from each service having its own specifications for maintainability demonstration, a tri-service Department of Defense Task Group issued a Military Standard on this subject in February **1966.** This document, MIL-STD-471, ***Maintainability Demonstration*** (Ref. ***2)*** superseded a number of existing requirement documents issued by the various services. Its scope is not limited to any class of equipment but is intended, at least implicitly, to be universally applicable. Section 1.2 states that "This standard is intended for use in demonstration of maintainability at any level (system, subsystem, equipment, etc.) and at any level of maintenance under any defined set of maintenance conditions."

In May, 1971, the Department of Defense issued a new Military Standard, MIL-STD-473, ***Maintainability Verification/Demonstration/Evaluation for Aeronautical Systems*** (Ref. 3). It "is applicable to all Department of Defense procurements for aeronautical systems (aircraft and associated subsystems), and when specified by the procuring activity is also applicable to other procurements." For such procurements only, MIL-STD-473 supersedes MIL-STD-471.

As these two military standards are the applicable documents for maintainability demonstration, they will be described in this paragraph. The methods described in these two military standards will be detailed here according to whether they are sequential plans (par. 6-7.1), plans which are based on an application of the central limit theorem (par. 6-7.2), techniques applicable when requirements are stated in terms of equipment repair time (par. 6-7.3), or techniques to be used when the underlying distribution of the data is unknown (par. 6-7.4). In addition to the MIL-STD's themselves, this paragraph draws on an article by Mazzola (Ref. 13) which gives the background to the preparation of MIL-STD-471, as well as a summary of the six methods therein. A ***Maintainability Engineering Handbook*** (Ref. 14), issued by the Naval Ordnance Systems Command, and Bird's article (Ref. 15) concerning MIL-STD-471 were also consulted.

6-7.1 SEQUENTIAL TEST PLANS

Sequential test plans are plans in which the number of observations comprising the test or the duration of the test is not specified in advance, but is determined on the basis of the progress of the test as it proceeds.

For example, in testing whether the mean corrective downtime for an item meets its specification, the test will be passed quickly if the bulk of corrective downtimes is less than the specified mean, will be failed quickly if the bulk of corrective downtimes exceeds the specified mean, and will require a longer time if the mean corrective downtime is close to the specification.

Method 1 of MIL-STD-471 is a sequential procedure for testing whether specifications of the mean corrective maintenance downtime $\overline{M}_{ct}$ ***and*** of the maximum corrective maintenance downtime $M_{max_{ct}}$ are *jointly* satisfied. This method permits a decision to accept only when an accept decision is made for both $\overline{M}_{ct}$ and $M_{max_{ct}}$.

The method assumes that corrective maintenance task times follow a lognormal distribution. MIL-STD-471 states that the level of risks of the tests is only slightly different if the corrective maintenance task time distribution is normal or exponential. The method is based on maximum level of risk for producer ("α" risk) and consumer ("β" risk) of 16 percent. Constraints on its use are: (a) the specified value of $\overline{M}_{ct}$ may not be less than 10 min nor exceed 100 min (i.e., $10 \leqslant \overline{M}_{ct} \leqslant 100$); and (b) the ratio of specified $M_{max_{ct}}$ to specified $\overline{M}_{ct}$ must not exceed 3:1. MIL-STD-471 asserts that deviation from these conditions (a) and (b) is uncommon. To the extent that this assertion is correct, constraints (a) and (b) will not seriously limit the application of this test method.

To apply Method 1 of MIL-STD-471 one must specify values for both $\overline{M}_{ct}$ and $M_{max_{ct}}$. Then one applies plan A_1 to determine conformance to the $\overline{M}_{ct}$ specification and plan B_1 or plan B_2 to determine conformance to the $M_{max_{ct}}$ specification. Plan B_1 is chosen if $M_{max_{ct}}$ is defined to be the 90th percentile; plan B_2 if $M_{max_{ct}}$ is taken as the 95th percentile.

The accept/reject criteria for plans A_1, B_1, and B_2 are detailed in Tables 6-3, 6-4, and 6-5, respectively. When one plan yields an accept decision, attention to that plan ceases and the remaining plan is continued until a decision is reached for it. The equipment is rejected when a decision to reject is reached on ***either*** plan; in order to accept the equipment, ***both*** plans must yield an accept decision.

If no accept or reject decision has been made after 100 observations, the decision shall be made according to the following rules:

1. Plan A_1: accept only if 29 or fewer observations exceed $\overline{M}_{ct}$.
2. Plan B_1: accept only if **5** or fewer observations exceed $M_{max_{ct}}$.
3. Plan B_2: accept only if 2 or fewer observations exceed $M_{max_{ct}}$.

The accept/reject criteria for plans A_1, B_1, and B_2 as defined by Tables 6-3, 6-4, and 6-5 and the preceding rules are graphically represented in Fig. 6-5.

The risks associated with plan A_1, B_1, and B_2 are summarized in Table 6-6.

6-7.2 TEST PLANS BASED ON THE CENTRAL LIMIT THEOREM

The central limit theorem, upon which all but one of the tests described in this paragraph are based, states that if $X_1, \ldots, X_n$ are identically distributed, independ-

TABLE 6-3. ACCEPT/REJECT CRITERIA FOR PLAN A, OF TEST METHOD 1 (FROM TABLE I, MIL-STD-471)

Number of Tasks Performed *N*	Observation Exceeding		Number of Tasks Performed *N*	Observation Exceeding	
	Accept	Reject		Accept	Reject
5	-	5			
6	-	6	56	13	20
7	-	6	57	13	21
8	-	6	58	13	21
9	-	7	59	14	21
10	-	7	60	14	22
11	-	7	61	14	22
12	0	7	62	14	22
13	0	8	63	15	23
14	0	8	64	15	23
15	1	8	65	15	23
16	1	9	66	16	23
17	1	9	67	16	24
18	1	9	68	16	24
19	2	9	69	17	24
20	2	10	70	17	25
21	2	10	71	17	25
22	3	10	72	17	25
23	3	11	73	18	25
24	3	11	74	18	25
25	4	11	75	18	26
26	4	12	76	19	26
27	4	12	77	19	27
28	4	12	78	19	27
29	5	12	79	20	27
30	5	13	80	20	28
31	5	13	81	20	28
32	6	13	82	20	28
33	6	14	83	21	28
34	6	14	84	21	29
35	7	14	85	21	29
36	7	15	86	22	29
37	7	15	87	22	30
38	7	15	88	22	30
39	8	15	89	22	30
40	8	16	90	23	31
41	8	16	91	23	31
42	9	16	92	23	31
43	9	17	93	24	31
44	9	17	94	24	32
45	9	17	95	24	32
46	10	17	96	25	32
47	10	18	97	25	33
48	10	18	98	25	33
49	11	18	99	25	33
50	11	19	100	26	33
51	11	19			
52	12	19			
53	12	20			
54	12	20			
55	12	20			

TABLE 6-4.
ACCEPT/REJECT CRITERIA FOR PLAN B_1 OF TEST METHOD 1
(FROM TABLE 11, MIL-STD-471)

Number of Tasks Performed *N*	Observations Exceeding $M_{max_{ct}}$ 90th Percentile		Number of Tasks Performed *N*	Observations Exceeding $M_{max_{ct}}$ 90th Percentile	
	Accept	Reject		Accept	Reject
			51		3
2	-	2	52	-	3
3	-	2	53	-	3
4	-	2	54		3
5	-	2	55	-	3
6	-	2	56	-	3
7	-	2	57	0	3
8	-	2	58	0	3
9	-	2	59	0	3
10	-	2	60	0	3
11	-	2	61	0	3
12	-	2	62	0	3
13	-	2	63	0	3
14	-	2	64	0	3
15	-	2	65	0	3
16	-	2	66	0	3
17	-	2	67	0	3
18	-	2	68	0	3
19	-	2	69	0	3
20	-	2	70	0	4
21	-	2	71	0	4
22	-	2	72	0	4
23	-	2	73	0	4
24	-	2	74	0	4
25	-	2	75	0	4
26	-	2	76	0	4
27	-	2	77	0	4
28	-	3	78	0	4
29	-	3	79	0	4
30	-	3	80	0	4
31	-	3	81	0	4
32	-	3	82	0	4
33	-	3	83	0	4
34	-	3	84	0	4
35	-	3	85	0	4
36	-	3	86	0	4
37	-	3	87	0	4
38	-	3	88	0	4
39	-	3	89	0	4
40	-	3	90	0	4
41	-	3	91	0	4
42	-	3	92	0	4
43	-	3	93	0	4
44	-	3	94	0	4
45	-	3	95	0	4
46	-	3	96	0	4
47	-	3	97	0	4
48	-	3	98	0	4
49	-	3	99	1	4
50	-	3	100	1	4

TABLE 6-5.
ACCEPT/REJECT CRITERIA FOR PLAN B_2 OF TEST METHOD 1
(FROM TABLE III, MIL-STD-471)

Number of Tasks Performed N	Observations Exceeding $M_{max_{ct}}$ 95th Percentile		Number of Tasks Performed N	Observations Exceeding $M_{max_{ct}}$ 95th Percentile	
	Accept	Reject		Accept	Reject
			51	1	4
2	-	2	52	1	4
3	-	2	53	1	5
4	-	2	54	1	5
5	-	2	55	1	5
6	-	2	56	1	5
7	-	2	57	1	5
8	-	2	58	1	5
9	-	2	59	1	5
10	-	2	60	1	5
11	-	2	61	1	5
12	-	2	62	1	5
13	-	2	63	1	5
14	-	3	64	1	5
15	-	3	65	2	5
16	-	3	66	2	5
17	-	3	67	2	5
18	-	3	68	2	5
19	-	3	69	2	5
20	-	3	70	2	5
21	-	3	71	2	5
22	-	3	72	2	5
23	-	3	73	2	6
24	-	3	74	2	6
25	-	3	75	2	6
26	0	3	76	2	6
27	0	3	77	2	6
28	0	3	78	2	6
29	0	3	79	2	6
30	0	3	80	2	6
31	0	3	81	2	6
32	0	3	82	2	6
33	0	3	83	2	6
34	0	4	84	2	6
35	0	4	85	3	6
36	0	4	85	3	6
37	0	4	87	3	6
38	0	4	88	3	6
39	0	4	89	3	6
40	0	4	90	3	6
41	0	4	91	3	6
42	0	4	92	3	6
43	0	4	93	3	7
44	0	4	94	3	7
45	0	4	95	3	7
46	1	4	96	3	7
47	1	4	97	3	7
48	1	4	98	3	7
49	1	4	99	3	7
50	1	4	100	3	7

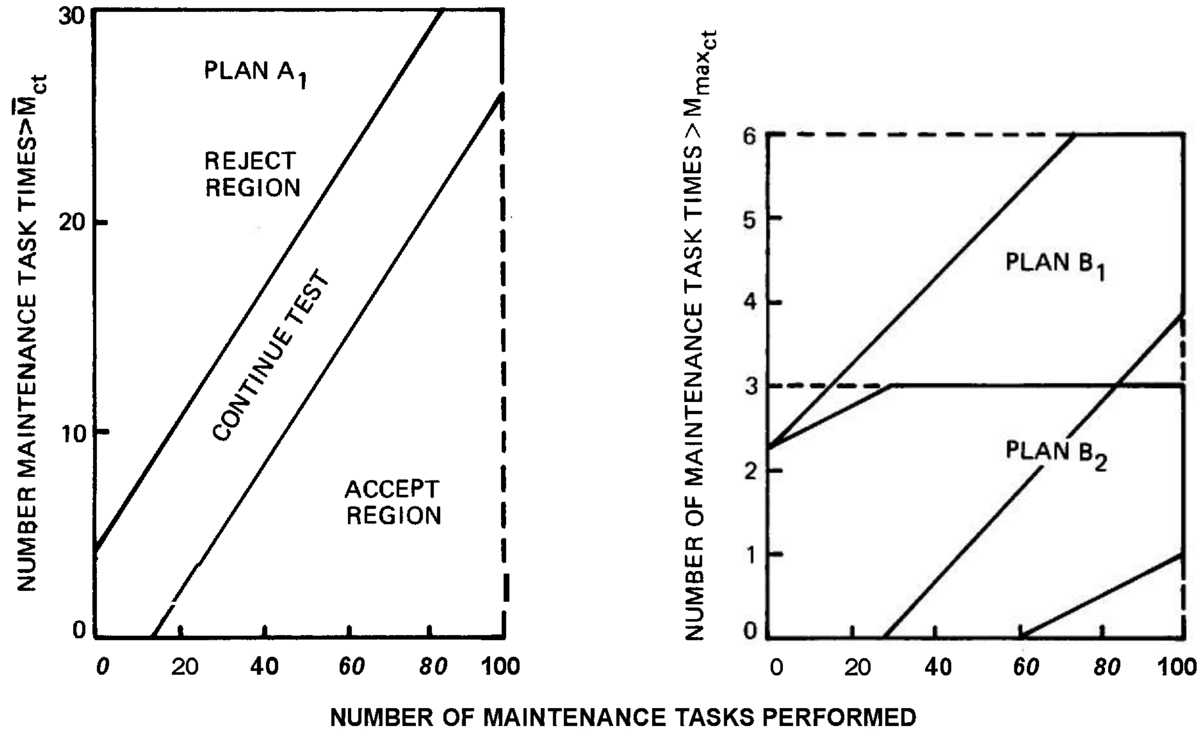


Figure 6-5. Graphic Representation of Plans A, and B_2 of Test Method 1

ent random variables each with mean μ and variance σ^2, and if

$$\bar{X} = \sum_{i=1}^{n} X_i/n \tag{6-39}$$

is their average, then the distribution of

$$z_n = \sqrt{n}[(\bar{X} - \mu)/\sigma] \tag{6-40}$$

approaches the normal distribution with mean 0 and variance 1 as n becomes infinite. The theorem does not say anything about the rate of approach, **i.e.,** for what value of n one may suppose that the limiting distribution has effectively been reached.

6-7.2.1 Tests in MIL-STD-471

Test Method 2 of MIL-STD-471 is designed to demonstrate the following specified values:

1. $\bar{M}_{ct}$: the specified mean corrective maintenance downtime
2. $\bar{M}_{pt}$: the specified mean preventive maintenance downtime
3. $\bar{M}$: the mean maintenance downtime
4. $M_{max_{ct}}$: the maximum corrective maintenance downtime.

TABLE 6-6. RISKS ASSOCIATED WITH THE USE OF PLANS A_1, B_1, AND B_2

1. Plan A_1: Producer's risk, α = 6% at $k \leqslant 0.22$;
Consumer's risk, β = 6%, $k \geqslant 0.39$

where

$k = \dfrac{r(A)}{N_c}$ = proportion of repair times exceeding $\bar{M}_{ct}$

$r(A)$ = number of repair times exceeding $\bar{M}_{ct}$

N_c = number of corrective maintenance tasks performed in the test

2. Plan B_1: Producer's risk, α = 10% at $k \leqslant 0.02$;
Consumer's risk, β = 10% at $k \geqslant 0.10$

3. Plan B_2: Producer's risk, α = 10% at $k \leqslant 0.01$;
Consumer's risk, β = 10% at $k \geqslant 0.05$

where

$k = \dfrac{r(B)}{N_c}$

$r(B)$ = number of repair times exceeding $M_{max_{ct}}$

All the demonstration plans except the one for $M_{max_{ct}}$ make use of the central limit theorem. (The demonstration plan for $M_{max_{ct}}$ assumes a lognormal distribution for corrective maintenance task times.) Hence, with that exception, these plans are not restricted by the form of the underlying distribution of downtime, provided the size of the sample of maintenance tasks is adequate. A minimum sample size of 50 corrective maintenance tasks is required for the demonstration of $\bar{M}_{ct}$ and $M_{max_{ct}}$, and an additional sample of 50 preventive maintenance tasks is required for the demonstration of $\bar{M}_{pt}$.

When Test Method 2 is to be used, the parameter(s) $\bar{M}_{pt}$, $\bar{M}_{ct}$, $\bar{M}$, or $\bar{M}_{max_{ct}}$ and the producer's risk α associated with each proposed test plan must be specified. If $M_{max_{ct}}$ is specified, the percentile defining the "maximum" must be stated. Typically, it will be the 90th or 95th percentile.

One selects 50 maintenance tasks, records the time to perform each task, and computes the sample estimates of each of the specified parameters as follows:

1. Sample mean corrective downtime

$$M_{ct'} = \sum_{i=1}^{N_c} M_{ct_i}/N_c \qquad (6\text{-}41)$$

where

$M_{ct'}$ = observed mean corrective downtime in the sample of maintenance tasks used in the test

M_{ct_i} = individual corrective maintenance task time in the test sample

N_c = sample size, i.e., the number of corrective maintenance tasks performed during the test

2. Sample mean preventive downtime $\bar{M}_{pt'}$

$$\bar{M}_{pt'} = \sum_{i=1}^{N_p} M_{pt_i}/N_p \qquad (6\text{-}42)$$

3. Sample mean total maintenance downtime $\bar{M}'$

$$\bar{M}' = (f_c \bar{M}_{ct'} + f_p \bar{M}_{pt'})/(f_c + f_p) \qquad (6\text{-}43)$$

where

f_c = number of predicted corrective tasks occurring in a designated period of time

f_p = number of expected preventive tasks performed during the same time period

4. Sample maximum corrective downtime $M'_{max_{ct}}$ is computed from:

$$M'_{max_{ct}} = \text{antilog}\,(\overline{\log M_{ct'}} + \Phi S_{\log M_{ct'}}) \qquad (6\text{-}44)$$

where

$$\overline{\log M_{ct'}} = \sum_{i=1}^{N_c} \log M_{ct_i}/N_c$$

$$= \text{mean of the logarithms of } M_{ct_i} \qquad (6\text{-}45)$$

Φ = value from table of normal distribution one-tailed test corresponding to the specified percentage point, e.g., Φ = 1.645 for the 95th percentile

(We follow the notation in **MIL-STD-471** here. Φ is called z elsewhere in the present handbook. Common logs to the base 10 are used.)

$$S_{\log M_{ct'}} = \left\{\left[\sum(\log M_{ct_i})^2 - \sum(\log M_{ct_i})^2/N_c\right]\Big/(N_c - 1)\right\}^{1/2} \qquad (6\text{-}46)$$

$S_{\log M_{ct'}}$ = sample standard deviation of $\log M_{ct}$ (denoted as $\sigma_{\log M_{ct'}}$ in the **MIL-STD**)

Each of these computed sample statistics is then compared to the corresponding acceptheject criterion to determine conformance to the specified requirement as illustrated in the example that follows for determining conformance to the specified mean corrective maintenance **time $\bar{M}_{ct}$ at specified producer's risk α:**

$$\bar{M}_{ct'} + \Phi S_{M_{ct}}/\sqrt{N_c} \le \bar{M}_{ct}\ \text{(specified)} \qquad (6\text{-}47)$$

$$\overline{M}_{ct'} + \Phi S_{M_{ct}}/\sqrt{N_c} > \overline{M}_{ct} \text{ (specified)} \qquad (6\text{-}48)$$

where

$$\overline{M}_{ct'} = \sum M_{ct_i}/N_c \qquad (6\text{-}120)$$

Φ = value, from normal distribution function for a one-tailed test employing a large sample and central limit theorem, corresponding to the specified producer's risk a; e.g., Φ = **1.645** for a = 5%; Φ = 1.282 for a = 10%

$$S_{M_{ct'}} = \left\{\left[\sum(M_{ct_i})^2 - \left(\sum M_{ct_i}\right)^2 \Big/ N_c\right] \Big/ (N_c - 1)\right\}^{1/2} \qquad (6\text{-}49)$$

= standard deviation of sample corrective maintenance times (denoted as $\sigma_{M_{ct'}}$ in the MIL-STD)

Acceptance for maximum corrective time is based on the observed value $M'_{max_{ct}}$ being equal to or less than the specified value $M_{max_{ct}}$, i.e.,

$$\text{Accept if: } M'_{max_{ct}} \leq M_{max_{ct}} \text{ (specified)} \qquad (6\text{-}50)$$

$$\text{Reject if: } M'_{max_{ct}} > M_{max_{ct}} \text{ (specified)} \qquad (6\text{-}51)$$

where

$$M'_{max_{ct}} = \text{antilog}\left[\overline{\log M_{ct}} + \Phi\left\{\left[\sum(\log M_{ct_i})^2 - \left(\sum \log M_{ct_i}\right)^2 \Big/ N_c\right] \Big/ (N_c - 1)\right\}^{1/2}\right] \qquad (6\text{-}52)$$

The acceptheject criteria for determining conformance to the specified mean preventive maintenance time $\overline{M}_{pt}$ at specified producer's risk a is:

$$\text{Accept if: } \overline{M}_{pt'} + \Phi S_{M_{pt'}}/\sqrt{N_p} \leq \overline{M}_{pt} \text{ (specified)} \qquad (6\text{-}53)$$

$$\text{Reject if: } \overline{M}_{pt'} + \Phi S_{M_{pt'}}/\sqrt{N_p} > \overline{M}_{pt} \text{ (specified)} \qquad (6\text{-}54)$$

where

$$\overline{M}_{pt'} = \sum M_{pt_i}/N_p \qquad (6\text{-}55)$$

Φ = as defined for **Eq.** 6-48

$$S_{M_{pt'}} = \left\{\left[\sum(M_{pt_i})^2 - \left(\sum M_{pt_i}\right)^2 \Big/ N_p\right] \Big/ (N_p - 1)\right\}^{1/2} \qquad (6\text{-}56)$$

where

$S_{M_{pt'}}$ = sample standard deviation of preventive maintenance downtime (denoted $\sigma_{M_{pt'}}$ in the MIL-STD)

N_p = sample size, i.e., the number of preventive maintenance tasks performed during the test.

The accept/reject criteria for determining conformance to the specified mean of all maintenance actions $\overline{M}$ at specified producer's risk a is:

Accept if:

Accept if:

$$\bar{M}' = \Phi\{[N_p(f_c S_{M_{ct}'})^2 + N_c(f_p S_{M_{pt}'})^2]/[N_c N_p(f_c + f_p)]\}^{1/2} \leq \bar{M} \text{ (specified)} \quad (6\text{-}57)$$

Reject if:

$$\bar{M}' = \Phi\{[N_p(f_c S_{M_{ct}'})^2 + N_c(f_p S_{M_{pt}'})^2]/[N_c N_p(f_c + f_p)]\}^{1/2} > \bar{M} \text{ (specified)} \quad (6\text{-}58)$$

where

f_c = number of expected corrective maintenance tasks occurring during a representative operating time T

f_p = number of expected preventive maintenance tasks occurring during a representative operating time T

and all other symbols are the same as previously defined.

6-7.2.2 The Test of MIL-STD-473

While the test methods of MIL-STD-471 based on the Central Limit Theorem control only the producer's risk α, the test method based on the Central Limit Theorem that is described in MIL-STD-473 controls the consumer's risk β as well. The methods of MIL-STD-471 specify a sample size of 50 for each procedure, while the method of MIL-STD-473 has a ***minimum*** sample size of 50.

The notation that follows is needed to describe the procedure of MIL-STD-473. The procedure is phrased in terms of chargeable maintenance downtime (CMDT) per flight.

M = true, unknown, mean CMDT per flight

M_o = specified CMDT per flight

σ = true, unknown, standard deviation of CMDT per flight

X_i = the CMDT used after the ith flight

n = number of flights

$\bar{X}$ = mean of sample of n CMDT's $= (\sum_{i=1}^{n} X_i)/n$

S = standard deviation of n CMDT's

$$S = \left[\sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)\right]^{1/2} \quad (6\text{-}59)$$

z_α = standard normal deviate exceeded with probability α. Table 6-7 showing α vs z_α follows.

The test procedure of MIL-STD-473 tests the null hypothesis

H,: $M \leq M_o$

against the alternative hypothesis

H,: $M > M_o$

The sample size n required to implement the test is given by

$$n = \max\left\{50, \frac{(z_\alpha - z_{1-\beta})^2}{[M - M_0/\sigma]^2}\right\} \quad (6\text{-}60)$$

Eq. 6-60 involves:

1. α, the producer's risk, which is the probability that the null hypothesis will be rejected when, in fact, it is true.

2. β, the consumer's risk, which is the probability that the null hypothesis will be accepted when $M - M_o$ has some positive value.

3. σ, the population standard deviation of CMDT.

Generally, σ will not be known, but will have to be determined on the basis of previous data, maintainability mathematical models, or a specification requirement.

The procedure for testing the null hypothesis against the alternative hypothesis is:

$$\text{Accept } H_0 \text{ if: } \bar{X} \leq M_0 + z_\alpha S/\sqrt{n} \tag{6-61}$$

$$\text{Reject } H_0 \text{ in favor of } H_1 \text{ if: } \bar{X} > M_0 + z_\alpha S/\sqrt{n} \tag{6-62}$$

Example: Input: $\alpha = 0.10$, $\beta = 0.10$, $\sigma = 1$, $M - M_o = 0.3$

Then $z_\alpha = 1.28$ and $z_{1-\beta} = -1.28$

$$n = \max\left\{50, \left(\frac{1.28 + 1.28}{.3}\right)^2\right\} = 73$$

(since 73 > 50) (6-63)

6-7.3 EQUIPMENT REPAIR TIME (ERT) TECHNIQUES

Test Method 3 of MIL-STD-471 is to be used for demonstrating maintainability when it is stated in terms of the population median ERT of the distribution of corrective maintenance downtime, assuming that this distribution belongs to the lognormal family. A sample of 20 corrective maintenance tasks is specified by the MIL-STD.

The equipment under test will have met the ERT requirement if

$$\log MTTR_g \leq \log ERT + 0.397S \text{ (see Eq. 6-65)} \tag{6-64}$$

where

$\log ERT$ = logarithm of the median of the population of corrective maintenance downtimes

$MTTR_g$ = sample geometric mean time to repair

$$\log MTTR_g = \sum_{i=1}^{N_c} \log M_{ct_i}/N_c \tag{6-65}$$

where

$$S = \left\{\left[\sum_{i=1}^{N_c} (\log M_{ct_i})^2/N_c - (\log MTTR_g)^2\right]\right\}^{1/2} \tag{6-121}$$

where

N_c = number of corrective maintenance tasks (specified as 20 in the MIL-STD)

M_{ct_i} = duration of ith corrective maintenance task.

This acceptance criterion has a producer's risk α of 0.05.

The rationale for this acceptance criterion and verification of its producer's risk is as follows. If M_{ct} has the lognormal distribution with median ERT, then log M_{ct} is normally distributed with mean log ERT. The mean log $MTTR_g$, of a sample of size N_c of the values log M_{ct_1}, . . ., log M_{ct}/N_c is also normally distributed with mean log ERT. If S denotes the sample standard deviation of the numbers log M_{ct_1}, . . ., log M_{ct}/N_c, then $\sqrt{N_c - 1}$ [log $MTTR_g$ − log ERT]/S has the Student's-t distribution with $N_c - 1$ degrees of freedom.

TABLE 6-7.
STANDARD NORMAL DEVIATE EXCEEDED WITH PROBABILITY α

α	0.01	0.05	0.10	0.15	0.20	0.30
z_α	2.33	1.65	1.28	1.04	0.84	0.52

Note: $z_{1-\alpha} = -z_\alpha$

Thus one rejects the hypothesis that median corrective maintenance downtime does not exceed a specified level ***ERT*** in favor of the alternative hypothesis that median corrective maintenance downtime does exceed ***ERT***, based on a sample of size N_c and with a producer's risk of α if

$$\sqrt{N_c - 1}\ (\log MTTR_g - \log ERT)/S < t_{N_c-1,\alpha} \qquad (6\text{-}66)$$

where $t_{(N_c-1,\alpha)}$ is the value of a random variable distributed as Student's-t with $N_c - 1$ degrees of freedom that is exceeded with probability α. For $N_c = 20$ and $\alpha = 0.05$, we find from tables of percentage points of the Student's-t distribution that $t(19, 0.05) = 1.729$. Eq. 6-66 can be rearranged to read:

Reject if:

$$\log MTTR_g < \log ERT + (t_{N_c-1,\alpha})S/\sqrt{N_c - 1} \qquad (6\text{-}67)$$

Eq. 6-67 becomes Eq. 6-64 if one sets $N_c = 20$ and $\alpha = 0.05$.

Finally, in order to have a consumer's risk β of 0.05 associated with this procedure, the specified value of ***ERT*** is determined by

$$ERT\ (\text{specified}) = 0.37\, ERT_{max} \qquad (6\text{-}68)$$

where ERT_{max} is the largest value of ***ERT*** that should be accepted no more than 10% of the time and where the standard deviation of the logarithm of M_{ct} is taken as ***0.55.***

6-7.4 TECHNIQUES WHEN UNDERLYING DISTRIBUTIONS OF DATA ARE UNKNOWN

The methods described in this paragraph are applicable generally and do not require any assumptions concerning the statistical distribution of the time to perform a maintenance task. They are called ***distribution-free*** or ***nonparametric*** statistical methods. They have the advantage of being valid no matter what the underlying population distribution; their chief disadvantage is that they require larger sample sizes to achieve the same power than do procedures which make specific distributional assumptions. (The power of a statistical test is the probability that it will reject a null hypothesis when, in fact, some alternative hypothesis holds.) Of course, one should generally use a distribution-free procedure when he had no basis for making specific distributional assumptions. This statement is made on the premise that a valid procedure is preferable to an invalid one, even though the valid procedure may involve a larger sample size than the invalid one. One exception must be noted to this dictum. A number of statistical procedures are insensitive to certain departures from assumptions used in their derivation. (This property is called ***robustness.)*** This topic is considerably advanced beyond this handbook, however.

MIL-STD-473 gives one distribution-free procedure and MIL-STD-471 another. The latter also contains two other procedures that are not tests, ***per se.*** All of these will be described in this paragraph.

6-7.4.1 Procedure of MIL-STD-473

Let T be a critical maintenance time, i.e., a maximum value for maintenance time (expressed either in elapsed time or man hours) that is to be exceeded only infrequently, say with probability 0.05 or 0.10. If X denotes the random variable, maintenance time, then the procedure of MIL-STD-473 tests the null hypothesis:

$$H_0:\ P(X > T) = p_0$$

against the alternative hypothesis

$$H_1:\ P(X > T) = p_1 > p_0$$

with p_0, p_1, α (the producer's risk), and β (the consumer's risk) given. Another way of stating the null and alternative hypothesis is $H_0 : T = X_{p_0}$ and $H_1 : T = X_{p_1}$ where X_p is the $100(1 - p)$ percentile of the distribution of maintenance time. This test procedure is distribution free, however. The form of the distribution need not be specified in order to apply the test.

The test procedure is to take a sample of size n of maintenance tasks, count the number of these tasks whose duration exceeds T (denote this number by r), accept H_0 if $r \leq c$ and reject H_0 if $r > c$. Here c is an appropriate acceptance number. The values of c and n are obtained by Eqs. 6-69 through 6-72.

If $0.20 \leq p_0 \leq 0.80$:

$$H_0: P(X > T) = p_0$$

$$H_1: P(X > T) = p_1 > p_0$$

$$n = \{[z_\beta \sqrt{p_1(1-p_1)} + z_\alpha \sqrt{p_0(1-p_0)}\,]/(p_1 - p_0)\}^2$$

(use next higher integer) (6-69)

and

$$c = n[z_\beta p_0\sqrt{p_1(1 - p_1)} + z_\alpha p_1\sqrt{p_0(1 - p_0)}][z_\beta\sqrt{p_1(1 - p_1)} + z_\alpha\sqrt{p_0(1 - p_0)}]^{-1}$$

(use next lower integer) (6-70)

If $p_o < 0.20$, c and n are the smallest integers satisfying

$$\sum_{j=0}^{c} e^{np_0}(np_0)^j/j! \geq 1 - \alpha \qquad (6\text{-}71)$$

and

$$\sum_{j=0}^{c} e^{np_1}(np_1)^j/j! \leq \beta \qquad (6\text{-}72)$$

Table 6-8 solves Eqs. 6-71 and 6-72, and provides the values of n and c for $p_o < 0.20$; $\alpha = 0.05, 0.10, 0.20$; $\beta = 0.05, 0.10, 0.20$; and $k = p_1/p_o = 1.5, 2.0, 2.5, 3.0, 4.0, 5.0, 10.0$. The value of c is read directly from Table 6-8 for the appropriate values of k, α, and β. To obtain n, read off D from Table 6-8 for the appropriate value of k, α, and β; n is the greatest integer less than D/p_o. For $0.20 \leq p_o \leq 0.80$ the values of n and c can be directly computed from Eqs. 6-71 and 6-72. Note that the value of n must be obtained first from Eq. 6-69; the computation of c in Eq. 6-70 involves this value of n.

This test procedure is based on the binomial distribution. However, for computational convenience the determination of n and c is based on the normal approximation to the binomial for $0.20 \leq p_o \leq 0.80$; for $p_o \leq 0.20$ the determination of n and c is based on the Poisson approximation to the binomial.

6-7.4.2 Procedures of MIL-STD-47 1

1. *Method 4:*

Method 4 of MIL-STD-471 is a fixed sample size procedure which controls the producer's risk α, but not the consumer's risk β. In the terminology just used to describe the procedure of MIL-STD-473,' this method tests the hypothesis

$$H_o : P(X > T) = p_o$$

against the alternative

$$H_1 : P(X > T) > p_o$$

at a producer's risk of α.

MIL-STD-471 considers only $p_o = 0.95$ and $p_o = 0.50$, with $\alpha = 0.25$ and $\alpha = 0.10$. A sample size of $n = 50$ is quoted for all procedures. The number r of maintenance times in the sample of 50 which exceed the critical time T is determined. The null hypothesis H_0 is accepted if $r \leq c$; H_0 is rejected if $r >$ c. The value of the acceptance number c appropriate to $p_o = 0.95$ is for a test of a specified value $T = M_{max_{ct}}$ or $T = M_{max_{pt}}$, the value $p_o = 0.50$ is for a test of the medians $T = \tilde{M}_{ct}$ or $T = \tilde{M}_{pt}$.

2. *Method 5:*

Method 5 of MIL-STD-471 describes a distribution-free tolerance interval based on the largest and smallest observation in a random sample. The numbers L and U form lower and upper tolerance limits with coverage at least $100Z$ percent at probability P if the probability is P that at least $100Z$ percent of a population lie in the interval from L to U. The numbers L and U are based on a sample $X_1, \ldots, X_N$ from the population. A tolerance interval is distribution-free if the relation among the coverage Z, the probability P, and the sample size N does not depend on the population distribution.

In this method L is taken as the smallest observation in the sample, and U is the largest observation in the sample. For Z and P taking on values 0.90, 0.95, 0.99, Table 6-10 gives the appropriate value of N. The values of N listed in Table 6-10 are a solution of the equation (Ref. 16, pp. 409-10)

$$H_o : P(X > T) = p_0$$

$$H_1 : P(X > T) > p_0$$

$$1 - P = Z^{N-1}[N(1 - Z) + Z] \qquad (6\text{-}73)$$

TABLE 6-8.
SAMPLING PLANS FOR SPECIFIED p_o, p_1, α, AND β WHEN p_o IS SMALL (e.g., p_o < 0.20)

$k = \frac{p_1}{p_o}$							α = 0.10						α = 0.20					
	P = 0.05				β = 0.20		β = 0.05		β = 0.10		β = 0.20		β = 0.05		β = 0.10		β = 0.20	
	c	D	c	D	c	D	c	D	c	D	c	D	c	D	c	D	c	D
1.5	66	54.1	54	43.4	39	30.2	51		40	33.0	29	23.2	36	31.8	27	23.5	17	14.4
2	22	15.7	18	12.4	14	9.25	17		14	10.3	10	7.02	12	9.91	9	7.29	6	4.73
2.5	13	8.46	10	6.17	3	4.70	10		8	5.43	6	3.90	7	5.58	5	3.84	3	2.30
3	9	5.43	7	3.98	6	3.29	7		5	3.15	4	2.43	4	3.09	3	2.30	2	1.54
4	6	3.29	5	2.61	4	1.97	4		3	1.75	2	1.10	3	2.30	2	1.54	1	0.824
5	4	1.97	3	1.37	3	1.37	3	1.75	2	1.10	2	1.10	2	1.54	1	0.824	1	0.824
10	2	0.818	2	0.818	1	0.353	1	0.532	1	0.532	1	0.532	1	0.824	1	0.824	0	0.227

To find the sample size n, for given p_o, p_1, α, and β, divide the appropriate D value by p_o and use the greatest integer less than the quotient.

Example: $p_o = 0.05$, $p_l = 0.20$, α 0.10, $\beta = 0.05$, **and** $k = \frac{0.20}{0.05} = 4$. **Then** $n = \frac{D}{0.05} = \frac{2.43}{0.05} = 48$. The **acceptance number** is $c = 4$.

TABLE 6-10.
N FOR CERTAIN COMBINATIONS OF Z AND P

P \ Z	0.90	0.95	0.99
	38		
0.95	47		490
0.99	66	135	690

3. *Method 6*

Method 6 of MIL-STD-471 is a procedure for demonstrating if mean and maximum preventive maintenance times satisfy specified values. It is a purely ad hoc procedure with no apparent basis in statistical theory.

The method simply calls for comparing the observed mean preventive maintenance time $\bar{M}_{pt'}$ with the specified mean preventive maintenance time $\bar{M}_{pt}$. An accept decision is made if

$$\bar{M}_{pt'} \leq \bar{M}_{pt} \tag{6-74}$$

a reject decision is made if

$$\bar{M}_{pt'} > \bar{M}_{pt} \tag{6-75}$$

Similarly, the method calls for comparing the observed maximum preventive maintenance time $M'_{max_{pt}}$ with the specified maximum preventive maintenance time $M_{max_{pt}}$. An accept decision is made if

$$M'_{max\,pt} \leq M_{max\,pt} \tag{6-76}$$

a reject decision is made if

$$M'_{max\,pt} > M_{max\,pt} \tag{6-77}$$

The calculation of observed mean preventive maintenance time $\bar{M}_{pt'}$ is to be made by the formula

$$\bar{M}_{pt'} = \sum_{i=1}^{k} f_i(\bar{M}_{pt_i}) \Big/ \sum_{i=1}^{k} f_i \tag{6-78}$$

where

f_i = frequency of occurrence of the ith task
k = number of different preventive maintenance tasks
$\bar{M}_{pt_i}$ = mean preventive maintenance time for the ith task.

MIL-STD-471 requires that all preventive maintenance tasks be performed. The preventive maintenance

TABLE 6-9.
ACCEPTANCE NUMBERS: MAXIMUM NUMBER OF TASKS IN A SAMPLE OF 50 WHICH MAY EXCEED THE CRITICAL TIME T

	p_o = 0.50	p_o = 0.95
a = 0.25	22	1
a = 0.10	20	0

task with the longest deviation is to be determined. This establishes a reference period, and all other preventive maintenance tasks are selected and apportioned over it. No sample size is specified. It should be pointed out that no statements concerning producer's or consumer's risk can be made concerning this method.

In order to calculate $M'_{max_{pt}}$ one must specify which percentile defines the maximum. Say it is the 100γth percentile. Then one orders the n observed preventive maintenance times from shortest to longest and takes $M'_{max_{pt}} = M_{pt[n\gamma]+1}$. Here M_{pt_j} is the jth ordered preventive maintenance time and $[n\gamma]$ is the largest integer contained in $n\gamma$. For example, if $n = 35$ and $\gamma = 0.90$, then $n\gamma = 31.5$ and $[n\gamma] = 31$. If $n\gamma$ is an integer, the $M'_{max_{pt}}$ is not uniquely defined, but may take on any value between $M_{pt[n\gamma]}$ and $M_{pt[n\gamma]+1}$ (Ref: 16, p. 410).

6-7.5 GOODNESS-OF-FIT TECHNIQUES

A number of the techniques discussed in the preceding paragraphs were applicable regardless of the distribution of the underlying random variable. Other techniques did depend on the distribution of the random variable. For example, a demonstration plan for $M_{max_{ct}}$ assumed a lognormal distribution for corrective maintenance task times. Other statistical techniques involve the assumption of a normal distribution or an exponential distribution, to give but two examples.

If a particular distribution needs to be assumed in order to use a prediction model, then serious errors may result if this assumption is incorrect. Thus, it is important to have some reasonable assurance that distributional assumptions are valid. For a treatment of the sorts of errors that can ensue from misspecification of the model, see, for example, Barlow and Proschan (Ref. 17).

How can one be sure that the distributional assumptions necessary to implement a test or other statistical procedure are justified? On occasion, the nature of the process generating the random observations is so well understood that the form of their probability distribution may be taken as known. For example, Goldman and Slattery (Ref. 18, page 46) discuss situations in which the distribution of certain downtimes may be assumed *a priori* to be exponential or lognormal. (See also par. 4-5.1.2.) There are also statistical tests of whether a given sample is associated with a specified family of distributions. A discussion of some of these tests comprises the remainder of this paragraph.

A point that must be made at the outset is that a test of distributional assumptions cannot assure the adequacy of a fit. The only conclusions possible from a test of goodness-of-fit are "the data do not conform to the model" or "the data do not give evidence of nonconformity to the model". It is not possible to conclude the adequacy of a fit, only that the inadequacy of a fit has not been shown. (This is analogous to the distinction between a judicial finding of "innocent" and one of "not guilty").

A distinction between two kinds of test of fit must also be made. Given a set of data, one could test whether these came from a completely specified distribution (i.e., not only the form of the distribution given, but also the values of all its parameters) or merely from a given family of distributions. Thus, for example, one could test that a sample was drawn from a normal distribution with mean 10 and variance 4, or one could test that the sample came from some normal distribution. As might be expected, there is a relationship between the test procedures for these two tests. We will point out where they are similar and also where distinctions must be observed.

Our discussion will feature a general discussion of the chi-square and Kolmogorov-Smirnov procedures, and specific discussions of tests of fit for the normal, lognormal, and exponential distributions.

6-7.5.1 Chi-square Procedures

Suppose a set of data is divided into k nonoverlapping classes or cells $A_1, A_2, \ldots, A_k$ such that, together, they comprise all possible outcomes of the process generating the data; i.e., each conceivable datum will fall into one and only one class. These classes may be defined prior to the data becoming available or after the data are recorded. The individual data measurements need not be known, only the number of observations falling into each class. That is, for purposes of a chi-square test, it is only necessary to know a set of k numbers $n_1, n_2, \ldots, n_k$ (where n_i is the number of observations in class A_i), and not the total sample

$$x_1, x_2, \ldots, x_N$$
$$(\text{where } N = n_1 + n_2 + \ldots + n_k).$$

We first discuss the chi-square test that the sample comes from a completely specified distribution function $f(x,\theta)$. Here the form of the distribution is known and the value of the parameter θ is known as well. θ may be a scalar quantity, as the mean in the exponential distribution; or it may denote a pair of numbers, as the mean and variance of a normal distribution; or it may denote more than two numbers. We use our knowledge of the distribution to calculate $p_i(\theta)$, the probability content of the class A_i, $i = 1, \ldots, k$. Suppose that the upper and lower end points of A_i are U_i and L_i, respectively. Then

$$p_i(\theta) = \int_{L_i}^{U_i} f(x, \theta)\,dx \qquad \textbf{(6-79)}$$

With the knowledge of the probabilities p_i (θ) and the observed frequencies n_i for the k classes, one calculates the test statistic X_2 (θ) given by

$$\chi^2(\theta) = \sum_{i=1}^{k} \{[n_i - N_{p_i}(\theta)]^2 / N_{p_i}(\theta)\} \qquad \textbf{(6-80)}$$

This quantity must be compared with an appropriate critical point of the chi-square distribution with (k — 1) degrees of freedom. Thus, if the test is to be made at the 100α percent confidence level, the hypothesis that the data come from the distribution $f(x, 8)$ is rejected if

$$\chi^2(\theta) > X^2_{k-1;1-\alpha}$$

where

$$\chi^2_{k-1;1-a}$$

is the 100(1 — a) percentile of the chi-square distribution with (k — 1) degrees of freedom.

When one has a choice of the number of classes (i.e., when the data do not come already grouped) he should divide the data into as many classes as possible, subject to the constraint that the expected frequency, $N_{p_i}(\theta)$, in each class should be at least 5. After forming classes and calculating the $N_{p_i}(\theta)$'s it may be necessary to combine adjacent classes to achieve this condition.

Next, we discuss a modification to the chi-square test of a fully specified hypothesis needed to make it a test of whether data come from a particular distributional family. Thus we test, for example, whether a sample can reasonably be supposed to come from some exponential distribution, rather than testing that that sample came from the exponential distribution with given mean 8.

Now if the parameter of the distribution is not specified, we cannot calculate the $p_i(\theta)$ and thus not $\chi^2(\theta)$. The correct way to proceed is to choose 8 so as to minimize $\chi^2(\theta)$. The resulting minimum value has the chi-square distribution with (k — p — 1) degrees of freedom, where p is the number of independent parameters involved. The value of θ, say θ^*, which yields the minimum value of $\chi^2(\theta)$, $\chi^2(\theta^*)$, is called the chi-square minimum estimate of 8.

The foregoing often is summarized by saying that the only difference between the chi-square test of a fully specified hypothesis and that of membership in family of distributions is that in the latter, one first estimates the parameters and then subtracts one degree of freedom for each independent parameter estimated. What is lost sight of in this characterization is that the method of estimation must be chi-square minimization which involves only the number of observations in the classes. If one uses, say, the maximum likelihood estimate $\hat{\theta}$ of the parameter(s), based on the individual observations $x_1, x_2, \ldots, x_N$, the quantity $\chi^2(\hat{\theta})$ cannot properly be compared with percentage points of the chi-square distribution with (k — p — 1) degrees of freedom. (For further details on the correct distribution of $\chi^2(\hat{\theta})$ see Ref. 19.)

However, for practical purposes, it is difficult to calculate the minimum value of $\chi^2(\theta)$ without a computer, and it becomes a matter of convenience to substitute the maximum likelihood estimate into $\chi^2(\theta)$ and compare the resulting value with a percentage point of the chi-square distribution with (k — p — 1) degrees of freedom. It is claimed that the result so obtained often does not differ greatly from the chi-square minimum method.

Before leaving the chi-square method we point out some of its advantages and disadvantages. Its chief advantage is that it can be applied to test the fit of a sample to any distribution, fully specified or not. Its chief disadvantage is its low power (i.e., poor ability to reject the hypothesis of fit when there is actually a significant departure from the hypothesis) and its dependence on the number of classes and the position of the boundaries between them.

With these generalities out of the way, we now give the formulas for the probability content of the classes for the exponential, normal, and lognormal distributions, and the Kolmogorov-Smirnov test of goodness-of-fit.

6-7.5.2 Exponential Distribution

A random variable distributed according to the exponential distribution with mean θ takes on all nonnegative values; thus $L_1 = 0$ and $U_k = \infty$.

$$p_i(\theta) = \int_{L_i}^{U_i} [\exp(-x/\theta)/\theta]\,dx$$
$$= \exp(-L_i/\theta) - \exp(-U_i/\theta)$$
$$\text{for } i = 2, \ldots, k-1 \qquad (6\text{-}81)$$

More specifically for the first ($i = 1$) and last ($i = k$) class we have

$$p_1(\theta) = 1 - \exp(-U_1/\theta) \qquad (6\text{-}82)$$

and

$$p_k(\theta) = \exp(-L_k/\theta) \qquad (6\text{-}83)$$

respectively.

6-7.5.3 Normal Distribution

A normally distributed random variable having mean μ and variance σ^2 takes on all values. Thus $L_1 = -\infty$ and $U_k = \infty$.

$$p_i(\mu,\sigma) = (\sqrt{2\pi}\sigma)^{-1}\int_{L_i}^{U_i} \exp[-(x-\mu)^2/(2\sigma^2)]\,dx$$
$$= \Phi[(U_i-\mu)/\sigma] - \Phi[(L_i-\mu)/\sigma] \qquad (6\text{-}84)$$

where

$$\Phi(z) = (\sqrt{2\pi})^{-1}\int_{-\infty}^{z}\exp(-t^2/2)\,dt \qquad (6\text{-}85)$$

is the widely tabulated cumulative distribution function of the standard normal distribution.

6-7.5.4 Lognormal Distribution

A lognormal distributed random variable with parameters ν and ω takes on all positive values. Here ν is the mean, and ω is the standard deviation of $\ln X$. Thus $L_1 = 0$ and $U_k = \infty$.

$$p_i(\nu,\omega) = \int_{L_i}^{U_i} (\sqrt{2\pi}\,\omega x)^{-1}\exp[-(\ln x - \nu)^2/(2\omega^2)]\,dx$$
$$= \Phi[(\ln U_i - \nu)/\omega] - \Phi[(\ln L_i - \nu)/\omega] \qquad (6\text{-}86)$$

where, again, $\Phi(z)$ is the cumulative distribution function of the standard normal distribution function.

We next discuss the Kolmogorov-Smirnov test of goodness-of-fit.

6-7.5.5 Kolmogorov-Smirnov Tests

Suppose that the random variable X has cumulative distribution function $F(x)$, that an ordered sample $x_1 < x_2 < \ldots < x_n$ is available, and that one wants to test the hypothesis that this sample was drawn from a population having cumulative distribution function $F(x)$. One can perform such a test by calculating the empirical cumulative distribution function corresponding to this sample

$$F_n(x) = \begin{cases} 0, & x < x_1 \\ i/n, & x_i \leq x < x_{i+1},\ i = 1, \ldots, n-1 \\ 1, & x \geq x_n \end{cases} \qquad (6\text{-}87)$$

calculating D_n, the maximum absolute discrepancy between these functions, i.e.,

$$D_n = \max_x |F_n(x) - F(x)| \qquad (6\text{-}88)$$

and comparing D_n with appropriate percentage points of the Kolmogorov-Smirnov statistic. (See Table 6-11 for rejection criteria.)

The Kolmogorov-Smirnov test for a completely specified hypothesis has the desirable property that the distribution of D_n does not depend on $F(x)$. This makes it possible to prepare one table of percentage points for this test which is valid for all underlying distributions $F(x)$. It also tends to be more powerful than the chi-square test, which has a similar property.

TABLE 6-11.
CRITICAL VALUES FOR THE KOLMOGOROV-SMIRNOV TEST OF GOODNESS OF FIT

	Significance level				
Sample size n	0.20	0.15	**0.10**	***0.05***	**0.01**
1	0.900	0.925	0.950	0.975	0.995
2	.684	.726	.776	.842	.929
3	.565	.597	.642	.708	.829
4	.494	.525	.564	.624	.734
5	.446	.474	.510	.563	.669
6	.410	.436	.470	.521	.618
7	.381	.405	.438	.486	.577
8	.358	.381	.411	.457	.543
9	.339	.360	.388	.432	.514
10	.322	.342	.368	.409	.486
11	.307	.326	.352	.391	.468
12	.295	.313	.338	.375	.450
13	.284	.320	.325	.361	.433
14	.274	.292	.314	.349	.418
15	.266	.283	.304	.338	.404
16	.258	.274	.295	.328	.391
17	.250	.266	.286	.318	.380
18	.244	.259	.278	.309	.370
19	.237	.252	.272	.301	.361
20	.231	.246	.264	.294	.352
25	.21	.22	.24	.264	.32
30	.19	.20	.22	.242	.29
35	.18	.19	.21	.23	.27
40				.21	.25
50				.19	.23
60				.17	.21
70				.16	.19
80				.15	.18
90				.14	
100				.14	
Asymptotic Formula:	$\frac{1.07}{\sqrt{n}}$	$\frac{1.14}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Reject the hypothetical distribution $F(x)$ if $D_n = \max_x |F_n(x) - F(x)|$ exceeds the tabulated value,

(For α = 0.01 and 0.05, asymptotic formulas give values which are too high—by 1.5 percent for n = 80.)

This table is taken from Massey, F.J., Jr., "The Kolmogorov-Smirnov test for goodness of fit," J. Amer. Stat. Assn., 46, 68-78 (1951), except that certain corrections and additional entries are from Birnbaum, Z.W., "Numerical tabulation of the distribution of Kolmogorov's statistic for finite sample size," J. Amer. Stat. Assn. 47, 425-441 (1952).

A disadvantage of the Kolmogorov-Smirnov test, relative to the chi-square test, is that it is not as easy to modify it to test for fit to a family of distribution functions. However, this defect has been remedied for the exponential and normal (and hence also lognormal) distributions. Lilliefors (Refs. **20** and **21**) has tabulated percentage points of a Kolmogorov-Smirnov test for exponentiality and normality where one substitutes maximum likelihood estimates for parameters. (To test for lognormality of a sample $x_1, \ldots, x_n$, one tests for normality the values $\ln x_1, \ldots, \ln x_n$.) We give Lilliefors' tables of critical points in his Kolmogorov-Smirnov tests for exponentiality and normality in Tables **6-12** and **6-13,** respectively. Lilliefors writes N for sample size and $S_N(x)$ for the empirical cumulative distribution instead of the n and $F_n(x)$ used earlier in this paragraph. His $F^*(x)$ is the cumulative distribution function of the family under examination with maximum likelihood estimates substituted for parameters.

We close this discussion by giving some references to additional tests of fit as well as to additional discussion concerning the chi-square and Kolmogorov-Smirnov tests. They are Chapter 8 of Hahn and Shapiro (Ref. 22), and paper by Epstein (Ref. 23), Chapter **4** of Aitchison and Brown (Ref. 24), and the references contained therein.

6-8 CONFIDENCE INTERVALS FOR DEMONSTRATION METHODS

Parameters describing a population characteristic, such as mean corrective maintenance downtime $\overline{M}_{ct}$, are seldom known and are estimated by a corresponding quantity in a sample drawn from the population (M_{ct}, in the case of $\overline{M}_{ct}$). This yields a single number, called a ***point estimate,*** as an estimate of the population parameter.

To cite only a single number as an estimate of a population parameter gives no indication of the precision of that estimate. Accordingly, statisticians have devised the concept of a ***confidence interval*** to give a ***range*** of values, not merely a single value, to estimate a population parameter. Associated with each method of determining a confidence interval is a number called the ***confidence coefficient.*** The confidence coefficient is the probability that the method yields an interval which covers the parameter value.

Confidence intervals are of two kinds: two-sided, and one-sided. A two-sided confidence interval is defined by a lower confidence limit L and an upper confidence limit U. Both L and U depend on the sample values drawn, the desired confidence coefficient, and the distribution being sampled. (The latter dependence is not present for distribution-free confidence intervals. These will be discussed in this paragraph.) Thus if θ is the parameter, L the lower confidence limit, U the upper confidence limit, and $1 - \alpha$ the confidence coefficient, one can write

$$P(L \leqslant \theta \leqslant U) = 1 - \alpha \tag{6-89}$$

A one-sided confidence interval can be of two varieties, upper or lower. An upper, one-sided confidence interval with confidence coefficient $1 - \alpha$ for the parameter is determined by the upper confidence limit U. These quantities are interrelated by the equation

$$P(\theta \leqslant U) = 1 - \alpha \tag{6-90}$$

Similarly a lower, one-sided confidence interval with confidence coefficient $1 - \alpha$ for the parameter θ is determined by the lower confidence limit L, with

$$P(\theta \geqslant L) = 1 - \alpha \tag{6-91}$$

By the $100q$th percentile of a random variable X, we mean a number ξ_q such that $P(X \leq \xi_q) = q$, i.e., $100q$ percent of the population lies below ξ_q.

With these definitions set down, we now exhibit upper and lower confidence intervals for parameters of a number of distribution functions. We consider the mean and percentiles of the normal, lognormal, and exponential distributions, as well as the parameter of the binomial distribution. Additionally, we introduce the concept of a tolerance interval and relate it to that of a confidence interval for a percentile. Finally, we discuss distribution-free confidence and tolerance intervals.

6-8.1 NORMAL DISTRIBUTION

Two cases must be distinguished here; i.e., the population variance σ^2 known, and σ^2 unknown.

6-8.1.1 σ^2 Known

A two-sided confidence interval with confidence coefficient $1 - \alpha$ for the mean μ of a normal distribution with known variance σ^2 is given by

$$\overline{X} - K_{\alpha/2}\sigma/\sqrt{n} < \mu < \overline{X} + K_{\alpha/2}\sigma/\sqrt{n} \tag{6-92}$$

Here $\overline{X}$ is the sample mean, n the sample size, and $K_{\alpha/2}$ is the $100(1 - \alpha/2)$ percentile of the standard normal distribution. (For example: if $\alpha = 0.05$, $K_{0.025} = 1.96$; if $\alpha = 0.10$ $K_{0.05} = 1.645$, etc.)

A one-sided upper confidence interval with confidence coefficient $1 - \alpha$ for the mean μ of a normal distribution with known variance σ^2 is given by

$$\mu \leq \overline{X} + K_\alpha \sigma/\sqrt{n} \tag{6-93}$$

A one-sided lower confidence interval with confidence coefficient $1 - \alpha$ for the mean μ of a normal distribution with known variance σ^2 is given by

$$\mu \geq \overline{X} - K_\alpha \sigma/\sqrt{n} \tag{6-94}$$

The $100q$th percentile ξ_q of a normal random variable with mean μ and variance σ^2 is given by $\xi_q = \mu + K_q\sigma$.

TABLE 6-12. TABLE OF CRITICAL VALUES OF D (TESTING FOR EXPONENTIALITY)

The values of D given in the table are critical values associated with selected values of N. Any value of D which is greater than or equal to the tabulated value is significant at the indicated level of significance. These values were obtained as a result of Monte Carlo calculations, using 5,000 samples for N = 3 (2), 19, 20, 25, 30, interpolation for N = 4 (2) 18, and extrapolating (see text) for N over 30.

Sample Size N	Level of Significance for $D = \max_X \lvert F^*(X) - S_N(X) \rvert$				
	0.20	0.15	0.10	0.05	0.01
3	0.451	0.479	0.511	0.551	0.600
4	.396	.422	.449	.487	.548
5	.359	.382	.406	.442	.504
6	.331	.351	.375	.408	.470
7	.309	.327	.350	.382	.442
8	.291	.308	.329	.360	.419
9	.277	.201	.311	.341	.399
10	.263	.277	.295	.325	.380
11	.251	.264	.283	.311	.365
12	.241	.254	.271	.298	.351
13	.232	.245	.261	.287	.338
14	.224	.237	.252	.277	.326
15	.217	.229	.244	.269	.315
16	.211	.222	.236	.261	.306
17	.204	.215	.229	.253	.297
18	.199	.210	.223	.246	.289
19	.193	.204	.218	.239	.283
20	.188	.199	.212	.234	.278
25	.170	.180	.191	.210	.247
30	.155	.164	.174	.192	.226
Over 30	$\frac{0.86}{\sqrt{N}}$	$\frac{0.91}{\sqrt{N}}$	$\frac{0.96}{\sqrt{N}}$	$\frac{1.06}{\sqrt{N}}$	$\frac{1.25}{\sqrt{N}}$

TABLE 6-13.
TABLE OF CRITICAL VALUES OF D (TESTING FOR NORMALITY)

The values of D given in the table are critical values associated with selected values of N. Any value of D which is greater than or equal to the tabulated value is significant at the indicated level of significance. These values were obtained as a result of Monte Carlo calculations, using 1,000 or more samples for each value of N.

Sample Size N	Level of Significance for $D = \max_X \lvert F^*(X) - S_N(X) \rvert$				
	0.20	0.15	0.10	0.05	0.01
4	0.300	0.319	0.352	0.381	0.417
5	.285	.299	.315	.337	.405
6	.265	.277	.294	.319	.364
7	.247	.258	.276	.300	.348
8	.233	.244	.261	.285	.331
9	.223	.233	.249	.271	.311
10	.215	.224	.239	.258	.294
11	.206	.217	.230	.249	.284
12	.199	.212	.223	.242	.275
13	.190	.202	.214	.234	.268
14	.183	.194	.207	.227	.261
15	.177	.187	.201	.220	.257
16	.173	.182	.195	.213	.250
17	.169	.177	.189	.206	.245
18	.166	.173	.184	.200	.239
19	.163	.169	.179	.195	.235
20	.160	.166	.174	.190	.231
25	.142	.147	.158	.173	.200
30	.131	.136	.144	.161	.187
Over 30	$\frac{0.736}{\sqrt{N}}$	$\frac{0.768}{\sqrt{N}}$	$\frac{0.805}{\sqrt{N}}$	$\frac{0.886}{\sqrt{N}}$	$\frac{1.031}{\sqrt{N}}$

A two-sided confidence interval with confidence coefficient 1 — a for the 100qth percentile t, of a normal distribution with mean μ unknown but with variance σ^2 known is given by

$$\bar{X} - K_{\alpha/2}\sigma/\sqrt{n} + K_q\sigma < \xi_q < \bar{X} + K_{\alpha/2}\sigma/\sqrt{n} + K_q\sigma \qquad (6\text{-}95)$$

A one-sided upper confidence interval with confidence coefficient 1 — a for the 100qth percentile ξ_q of a normal distribution with mean μ unknown but with variance σ^2 known is given by

$$\xi_q = \bar{X} + K_\alpha\sigma/\sqrt{n} + K_q\sigma \qquad (6\text{-}96)$$

A one-sided lower confidence interval with confidence coefficient 1 — a for the 100qth percentile of a normal distribution with mean μ unknown but variance σ^2 known is given by

$$\xi_q \geqslant \bar{X} - K_\alpha\sigma/\sqrt{n} + K_q\sigma \qquad (6\text{-}97)$$

Next we consider similar confidence intervals when the variance is unknown.

6-8.1.2 σ^2 Unknown

A two-sided confidence interval with confidence coefficient 1 $-\alpha$ for the mean μ of a normal distribution with variance unknown is given by

$$\bar{X} - (t_{\alpha/2;n-1})S/\sqrt{n} < \mu < \bar{X} + (t_{\alpha/2;n-1})S/\sqrt{n} \quad (6-98)$$

where S is the sample standard deviation

$$S = \left[\sum_{i=1}^{n} (X_i - \bar{X})^2/(n-1) \right]^{1/2} \quad (6-99)$$

and $t_{(\alpha/2,\ n-1)}$ is the 100 $(1-\alpha/2)$ percentile of the Student-t distribution with $(n-1)$ degrees of freedom.

A one-sided upper confidence interval with confidence coefficient 1 $-\alpha$ for the mean μ of a normal distribution with variance unknown is given by

$$\mu \leqslant \bar{X} + (t_{\alpha;n-1})S/\sqrt{n} \quad (6-100)$$

A one-sided lower confidence interval with confidence coefficient 1 $-\alpha$ for the mean μ of a normal distribution with variance unknown is given by

$$\mu \geqslant \bar{X} - (t_{\alpha;n-1})S/\sqrt{n} \quad (6-101)$$

Confidence intervals for the 100qth percentile ξ_q of a normal distribution with both mean and variance unknown involve the use of percentage points of the noncentral t-distribution. These have been tabulated by Resnikoff and Lieberman (Ref. 25), and by Scheuer and Spurgeon (Ref. 26), (see also the references in Scheuer and Spurgeon).

(The noncentral t-distribution differs from the ordinary, or central, t-distribution in that it involves two parameters: f the number of degrees of freedom, and δ the noncentrality parameter. The ordinary t-distribution involves only the one parameter, degrees of freedom. It may also be considered as a special case of the noncentral t-distribution with noncentrality parameter $\delta = 0$.)

A one-sided upper confidence interval with confidence coefficient 1 $-\alpha$ for ξ_q, the 100qth percentile of a normal distribution with both mean and variance unknown is given by

$$\xi_q \leqslant \bar{X} + t_\alpha S[(n-1)/n]^{1/2} \quad (6-102)$$

where $\bar{X}$ is the mean, and S the standard-deviation of a sample of size n; and t_α is the upper 100q percentage point of the noncentral t-distribution with $(n-1)$ degrees of freedom and noncentrality parameter $\sqrt{n}\,K_{1-q}$.

A one-sided lower confidence interval with confidence coefficient 1 $-\alpha$ for ξ_q is given by

$$\xi_q \geqslant \bar{X} + t_{1-\alpha}S[(n-1)/n]^{1/2} \quad (6-103)$$

where $t_{1-\alpha}$ is the lower 100 α percentage point of the noncentral t-distribution with $(n-1)$ degrees of freedom and noncentrality parameter $\sqrt{n}\,K_{1-q}$.

A two-sided confidence interval with confidence coefficient 1 $-\alpha$ for ξ_q is given by

$$\bar{X} + (t_{1-\alpha/2})S[(n-1)/n]^{1/2} < \xi_q < \bar{X} + t_{\alpha/2}S[(n-1)n]^{1/2} \quad (6-104)$$

where $t_{\alpha/2}$ and $t_{1-\alpha/2}$ are, respectively, the upper and lower 100 $\alpha/2$ percentage points of the noncentral t-distribution with noncentrality parameter $\sqrt{n}\,K_{1-q}$.

6-8.2 LOGNORMAL DISTRIBUTION

Recall that X has a lognormal distribution with parameters μ and σ^2 if $Y = \ln X$ has a normal distribution with mean μ and variance ω^2. The mean M and variance V of X are:

$$M = \exp[\mu + \sigma^2/2]$$
$$V = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1] \quad (6-105)$$

Thus, if one has a sample $x_1, x_2, \ldots, x_n$ from a lognormal distribution, then $y_1 = \ln x_1, y_2 = \ln x_2, \ldots, y_n = \ln x_n$ are a normally distributed sample, and one can use the methods detailed in par. 6-8.1 to obtain a confidence interval for μ.

However, theory provides no means of obtaining exact confidence intervals for M. For large samples, one can invoke the Central Limit Theorem to get an approximate confidence interval for M. Defining

$$a = \exp(\bar{y} + S_y^2/2)$$
$$b^2 = \exp(2\bar{y} + S_y^2)[\exp(S_y^2) - 1] \quad (6-106)$$

with $\bar{y}$ the mean, the S_y^2 the variance, of the transformed observations $y_1, y_2, \ldots, y_n$ (i.e., $y_i = \ln x_i$), an asymptotic confidence interval for M with confidence coefficient $1 - a$ is

$$a - K_{\alpha/2} b/\sqrt{n} \leq M \leq a + K_{\alpha/2} b/\sqrt{n} \qquad (6\text{-}107)$$

Additional discussion may be found in the book, The Lognormal *Distribution* by Aitchison and Brown (Ref. **24).**

6-8.3 EXPONENTIAL DISTRIBUTION

For the exponential distribution it is convenient to give results for a more general sampling situation than one in which the entire sample is available, i.e., censored sampling. In censored sampling, n items comprise the sample but only the first r sample values are known. (For example, n items are put on life test. Because the anticipated downtime needed to experience the failure of all of them may be quite long, the experimenter decides in advance to terminate the test after a predetermined number of r of them have failed.) Specifically, if the ordered sample values are $x_1 < x_2 < \ldots < x_n$ and the sample is censored at the r the observation, then only the values $x_1, x_2, \ldots, x_r$ are known. The integer $r (1 \leq r \leq n)$ is determined independently of the sampling process. If $r = n$ the entire sample is observed; if $r < n$, only a part of the sample is observed. Censored sampling often arises in reliability testing where a sample of n items is placed on life test and the test is terminated when a predetermined number r of the test items have failed.

The maximum likelihood estimate of $\hat{\theta}$, the mean of an exponential distribution, based on a sample censored at r out of n observations is given by

$$\hat{\theta}_{r,n} = \left[\sum_{i=1}^{r} x_i + (n - r)x_r\right]/r \qquad (6\text{-}108)$$

An alternate expression is (with $x_0 = 0$)

$$\hat{\theta}_{r,n} = \left[\sum_{i=1}^{r} (n - i + 1)(x_i - x_{i-1})\right]/r \qquad (6\text{-}109)$$

With this definition on the record, we now give expressions for confidence intervals for the mean and for percentiles of the exponential distribution.

A one-sided upper confidence interval with confidence coefficient $1 - \alpha$ for θ, the mean of an exponential distribution, based on a sample censored at r out of n observations, is given by

$$\theta \leq 2r\hat{\theta}_{r,n}/\chi^2_{\alpha;2r} \qquad (6\text{-}110)$$

where $\chi^2_{\alpha;2r}$ denotes the 100αth percentile of the chi-square distribution with $2r$ degrees of freedom.

A one-sided lower confidence interval with confidence coefficient $1 - a$ for 8, the mean of an exponential distribution, based on a sample censored at r out of n observations, is given by

$$\theta \geq 2r\hat{\theta}_{r,n}/\chi^2_{1-\alpha;2r} \qquad (6\text{-}111)$$

A two-sided confidence interval with confidence coefficient $1 - \alpha$ for θ, the mean of an exponential distribution, based on a sample censored at r out of n observations, is given by

$$2r\hat{\theta}_{r,n}/\chi^2_{1-\alpha/2;2r} < \theta < 2r\hat{\theta}_{r,n}/\chi^2_{\alpha/2;2r} \qquad (6\text{-}112)$$

A one-sided upper confidence interval with confidence coefficient $1 - a$ for ξ_q the $100q$th percentile of an exponential distribution with unknown mean θ, based on a sample censored at r out of n observations, is given by

$$\xi_q = (-2r\hat{\theta}_{r,n})\ln(1 - q)/\chi^2_{\alpha;2r} \qquad (6\text{-}113)$$

A one-sided lower confidence interval with confidence coefficient $1 - a$ for ξ_q the $100q$th percentile of an exponential distribution with unknown mean θ, based on a sample censored at r out of n observations, is given by

$$\xi_q \geq (-2r\hat{\theta}_{r,n})\ln(1 - q)/\chi^2_{1-\alpha;2r} \qquad (6\text{-}114)$$

A two-sided confidence interval with confidence coefficient $1 - a$ for ξ_q the $100q$th percentile of an exponential distribution with unknown mean 8 based on a sample censored at r out of n observations, is given by

$$\frac{(-2r\hat{\theta}_{r,n})\ln(1 - q)}{\chi^2_{1-\alpha/2;2r}} \leq \xi_q \leq \frac{(-2r\hat{\theta}_{r,n})\ln(1 - q)}{\chi^2_{\alpha/2;2r}} \qquad (6\text{-}115)$$

6-8.4 BINOMIAL DISTRIBUTION

A binomial distribution is appropriate in a situation in which there is a fixed number *n* of independent trials, each trial resulting in either "success" or "failure", and the probability of success p is the same at each trial. By "success" is meant any event of interest.

Based on a sample in which *r* successes occurred in *n* trials, one would like to construct upper, lower, and two-sided confidence intervals for *p*, the binomial parameter. If one has a table of the cumulative binomial distribution, one can obtain such confidence intervals (see, e.g., Ref. 27, page 369 and Ref. 32). However, tables and graphs yielding such intervals have been constructed, obviating the need for any calculation. See, for example, Ref. 28, and NAVWEPS Report 8090 (Ref. 29).

Tolerance Intervals. Suppose $X_1, \ldots, X_n$ is a sample of a random variable X having a continuous cumulative distribution function $F(X)$. If $T_1(X_1, \ldots, X_n) < T_2(X_1, \ldots, X_n)$ are two functions depending on the sample such that for $0 < \beta < 1$

$$P\{[F(T_2) - F(T_1)] \geqslant \beta\} = \gamma \tag{6-116}$$

then (T_1, T_2) is called a $100\beta\%$ ***tolerance interval*** with probability y for the population. This means that, with probability y, at least $100\beta\%$ of the population lies between T_1 and T_2. (β is called the *coverage* of the tolerance interval).

The quantities T_1 and T_2 are two-sided tolerance limits. One-sided upper and lower tolerance limits can also be defined. We say that $X \geq T_1$ is a one-sided lower tolerance interval with coverage β and probability γ if

$$P\{[1 - F(T_1)] \geqslant \beta\} = \gamma \tag{6-117}$$

Similarly X ■T_2 is a one-sided upper tolerance interval with coverage β and probability γ if

$$P\{[F(T_2) \geqslant 1 - \beta]\} = \gamma \tag{6-118}$$

It can be seen that a lower tolerance interval with coverage β and probability γ is a lower confidence interval with confidence coefficient for the 100(1 − β)th percentile, $\xi_{1-\beta}$; also that an upper tolerance interval with coverage β and probability y is an upper confidence interval with confidence coefficient y for the 100(1 − β)th percentile, $\xi_{1-\beta}$. No such relationship exists between two-sided tolerance intervals and two-sided confidence intervals for a quantile.

Two-sided tolerance intervals for a normal distribution of the form $X \pm kS$ are available. Values of k for different values of sample size n, coverage β, and probability y have been tabulated by Eisenhart, Hastay, and Wallis (Ref. 30, Chapter 2).

Tolerance intervals valid for any continuous distribution (i.e., so-called distribution-free or nonparametric tolerance intervals) are available. If $X_1 < X_2 < \ldots < X_n$ are sample values arranged in increasing order, and if a and b are integers with $1 \leq a < b$ ■n, ther a tolerance interval (X_a, X_b) with coverage β has probability $1 - I_\beta[b - a, n - (b - a) + 1]$ where $I_\beta(u, v)$ is the incomplete beta function defined by

$$I_\beta(u, v) = \int_0^\beta \Gamma(u, v)[\Gamma(u)\Gamma(v)]^{-1} X^{u-1}(1 - x)^{v-1} dx \tag{6-119}$$

This has been tabulated by Pearson (Ref. 31).

REFERENCES

1. MIL-STD-470, ***Maintainability Program Requirements,*** 21 March 1966.
2. MIL-STD-471, ***Maintainability Demonstration,*** 15 Feb. 1966.
3. MIL-ST D-473, ***Maintainability Verification/-Demonstration/Evaluation for Aeronautical Systems,*** 3 May 1971.
4. B. S. Blanchard and E. E. Lowery, ***Maintainability Principles and Practices,*** McGraw-Hill Book Co., N.Y., 1969, Chapter 14.
5. AMCP 705-1, ***Maintainability Engineering,*** June 1966.
6. F. D. Mazzola, "Maintainability Demonstrations", IEEE Transaction of Reliability, R-16, No. 1, 37 (May 1967).
7. G. H. Almy, "Allocations, Assessment, and Demonstration of System MTTR", Annals of Reliability and Maintainability, 1971 – 10th R&M Conference, American Society of Mechanical Engineers, 172(June 1971).
8. John V. Lavery, Jr., "Maintainability Demonstration Program", Annals of Reliability and Maintainability, 1971 – 10th R&M Conference, American Society of Mechanical Engineers, 168(June 1971).

9. Wayne Nelson, "Producer Risk Tables for Method 4 Maintainability Demonstration Plan of MIL-STD-471", Annals of Reliability and Maintainability, 1971 – 10th R&M Conference, American Society of Mechanical Engineers, 188(June 1971).
10. Hunter M. Sohn, Lockheed-Georgia Company, "Proven Techniques for Maintainability Predictions and Demonstrations", presented at 10th Reliability and Maintenance Conference, Anaheim, California, June 1971 (unpublished).
11. G. B. Dantzig, "On the Non-Existence of Tests of "Student's" Hypothesis Having Power Functions Independent of σ", Annals of Mathematical Statistics, 11, 186-192 (1940).
12. MIL-HDBK-472, ***Maintainability Prediction,*** 24 May 1966.
13. F. D. Mazzola, "Maintainability Demonstration", IEEE ***Transactions*** on ***Reliability,*** R-16, 37-42 (1967).
14. NAVORD OD 39223, ***Maintainability Engineering Handbook,*** Naval Ordnance Systems Command, 1 Feb. 1970.
15. G. T. Bird, "MIL-STD-471, Maintainability Demonstration", ***Journal of Quality Technology, I,*** 134-148 (1969).
16. B. W. Lindgren, ***Statistical Theory,*** 2nd Ed., MacMillan, New York, 1968.
17. R. E. Barlow and F. Proschan, "Exponential Life Test Procedures When the Distribution Has Monotone Failure Rate", Journal of the American Statistical Association, 62, 548-560 (1967).
18. A. S. Goldman and T. B. Slattery, ***Maintainability:*** A ***Major Element of System Effectiveness,*** John Wiley and Sons, Inc., N.Y., 1967.
19. H. Chernoff and E. L. Lehmann, "The Use of Maximum Likelihood Estimates in Chi-Square Tests for Goodness of Fit", Annals of Mathematical Statistics, 25, 579-586 (1954).
20. H. W. Lilliefors, "On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown", Journal of the American Statistical Association, 64, 387-89 (1969).
21. H. W. Lilliefors, "On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown", Journal of the American Statistical Association, 62, 399-402 (1967).
22. G. J. Hahn and S. S. Shapiro, ***Statistical Models in Engineering,*** John Wiley and Sons, Inc., N.Y., 1967.
23. B. Epstein, "Tests for the Validity of the Assumption that the Underlying Distribution of Life is Exponential", Technometrics, 2, 83-101 and 2, 167-183 (1960).
24. J. Aitchison and J. A. C. Brown, ***The Lognormal Distribution,*** Cambridge University Press, 1957.
25. G. J. Resnikoff and G. J. Lieberman, ***Tables of the Noncentral t-Distribution,*** Stanford University Press, Stanford, 1957.
26. E. M. Scheuer and R. A. Spurgeon, "Some Percentage Points of the Noncentral t-Distribution", Journal of the American Statistical Association, 58, 176-182.
27. S. S. Wilks, ***Mathematical Statistics,*** John Wiley and Sons, Inc., N.Y. 1962.
28. C. J. Clopper and E. S. Pearson, "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial", Biometrika, 26, 404-413.
29. NAVWEPS Report 8090, ***Binomial Reliability Table (Lower Confidence Limits for the Binomial Distribution),*** US Naval Ordnance Test Station, China Lake, California, January 1964.
30. C. Eisenhart, M. W. Hastay, W. A. Wallis, ***Techniques of Statistical Analysis,*** McGraw-Hill Book Co., Inc. N.Y. 1947.
31. K. Pearson, ***The Tables of the Incomplete Beta Function,*** Biometrika Press, London, 1934.
32. AMCP 706-109 Engineering Design Handbook, ***Tables of Cumulative Binomial Probabilities.***

CHAPTER 7

ECONOMETRICS
SECTION I

COST FACTORS AND ANALYSES

7-1 INTRODUCTION

The purpose of this chapter is to describe the use of econometric techniques in maintainability (*M*) engineering. After this introduction, par. **7-2** discusses developing cost factors to categorize existing system costs or predict new system costs, as they vary with M-related variables. These factors are used to derive M-related costs, and predict how these costs will change with M-related design and performance variables. An important contribution to success in maintainability cost analysis is the development of a logical, ordered, well-structured breakdown of M-related cost elements—a tree-like structure for planning, analysis, and design. Par. **7-2** concludes with a discussion of breakdown techniques. Par. **7-3** discusses the concepts and use of cost analysis. After developing a cost element structure and list of elements for a particular system, the analysis of system costs requires prediction of the variation of individual element costs. Historical data are desirable for many systems; many difficulties arise in collection and conditioning of the data for use. Just as technical and performance variables affect M-related costs, *M*-related elements affect system variables like manpower. A typical subanalysis illustrates this interaction. Another typical maintainability cost subanalysis illustrates the decision between repair and throw-away modules. Section I concludes with isolated, single-subject subanalysis of particular M-related issues. The reality is much more complex; changing one M-related variable impacts others, often in a complicated way. To treat these interactions, more complete models of the maintainability process are necessary.

Section **II** of this chapter treats mathematical models in cost analysis. Par. **7-4** discusses the development of such more complete models for predicting annual costs of repair and support. It discusses the structure, formulation, and combination of cost equations into such a model.

Pars. **7-5** through **7-7** discuss the use of cost models for system analysis, design, and decision. Par. **7-5** begins the discussion with a treatment of the use of cost analysis for system development and decision at the level of entire systems. This preliminary phase of system analysis concerns itself with choosing among alternate systems for accomplishing the same purposes. Par. **7-6** discusses the economics of trade-off decisions. After a decision to begin preliminary design of one or more systems, the internal configuration of each system must be specified. Economic trade-offs are an essential part of this process of selecting from among subsystem and component alternatives. Par. **7-7** concludes this chapter with a discussion of the management and organizational issues in performing cost analysis, and the final "bill-of-costs" for acquisition and operation of a system which is prepared at the conclusion of system design. This life-cycle cost analysis will reflect the results of all cost trade-offs performed during system design. It can be used to compare competitive designs for a system in order to select the least-cost alternative to buy and operate.

7-1.1 MAINTAINABILITY COSTS

One key economic issue in equipment design is the total cost to the Government of a particular equipment acquisition. Procurement of equipment implies a commitment to the operating and support costs of the equipment as well. Thus the relevant *costs* are those for development, acquisition, operation, and support. It is the total of these cost elements which counts—not the individual pieces. This total cost is affected by many factors, including maintainability and reliability. Many operation and support costs might be reduced if we could increase maintainability. More maintainable equipment reduces the technician's maintenance time and hence its operating cost. Fewer technician hours imply fewer technicians, reducing training costs. More rapid maintenance means faster return to service, reducing the size of the maintenance float required. Smaller backlogs of equipment awaiting maintenance can reduce the size of maintenance or overhaul facilities and the required amount of test equipment. The key comparison is the acquisition cost for such improved maintainability or reliability, and the resultant savings in operating costs. Improved maintainability can be obtained in a number of specific ways; they include (Ref. 1):

a. Discard-at-failure maintenance and module sizing
b. Easier access for maintenance
c. Design of built-in test points
d. Increased self-checking features
e. Greater use of automatic test equipment
f. Use of reduced-maintenance components, e.g., self-lubricating bearings
g. More detailed troubleshooting manuals.

Each approach can increase acquisition cost and decrease operating costs. A specific analysis, comparing these costs, is needed in each case.

7-1.2 OWNERSHIP COSTS

Costs of ownership fall into three major categories: research and development (R&D), acquisition, and operation and support. Equipment acquisition costs can account for half of defense costs for particular equipment. Operation and support costs, accounting for the other half, include manpower as their largest element. In one case the *M*-related operation and support costs have exceeded acquisition costs by a factor of 1500 (Ref. 2). As a result, maintainability strategies which reduce manpower should have high leverage.

7-2 COST FACTORS OF EQUIPMENT

When analyzing the total costs of ownership, it is necessary to break cost categories into elements and subelements which can then be estimated, analyzed, budgeted, reported, and controlled. The basic element breakdown is subdivided by considering a number of basic principles: (1) Each subelement must be a part of the higher level element costs; (2) subelements at any level must add to their summary elements; (3) subelements must be uniquely defined at every level to avoid overlap and double counting; (4) the lowest level for subelements (the most detailed breakdown) must permit direct estimation of costs. Thus different cost elements may be broken down to different levels, as appropriate.

A general caution is usually applied to costing a system, as stated in a Navy publication: ". . . to include all costs incurred in establishing, operating, and maintaining the . . . System, and to exclude all costs which would occur whether the . . . System existed or not" (Ref. 3).

The purpose of the cost analysis must be carefully considered. If we wish to account for the costs of a particular equipment, a complete categorization is necessary. If we are comparing alternatives for the purpose of selection, costs which do not vary among alternatives are often excluded. It is also necessary to consider the point in equipment life when the analysis is being performed. Once a decision has been made to procure a particular type of equipment, earlier costs leading up to that decision may become "sunk", and should be ignored, except for accounting purposes. Sunk costs are those that have already been incurred. "Relevant costs lie in the future, not in the past" (Ref. 4). In contemplating a new system in advance, however, the research and development costs are a legitimate part of the decision to proceed or not, and must be considered. Similarly, existing equipment—e.g., test equipment—is "free" except for its salvage value or alternative use. In deciding between new or existing test equipment, the original cost of the existing equipment is irrelevant to the analysis, providing enough sets of test equipment exist and are available for the purpose considered. New test equipment, not yet procured, must be costed at full acquisition price for comparison. In developing the cost elements of an equipment integrated cost structure, great care must be taken to include or exclude individual elements as appropriate.

7-2.1 INDIVIDUAL COST ELEMENTS

Integrated costs of ownership include many *M*-related costs in each phase of equipment life. Research and development costs include those for feasibility studies and tests of equipment and components. Many studies are "paper analyses" comparing different maintenance doctrines and performing many different trade-offs to arrive at the preferred system. The cost of these analyses and tests are a part of the development. Current procurement policy includes "fly before you buy" in many cases; a prototype equipment is produced and carefully tested and evaluated before the design is frozen. Many *M*-related design elements may be tested as a part of this process; again the costs must be considered. Feasibility studies may extend to support equipment, including automatic test equipment and special tools; these costs must also be included. New concepts in *M*-related design may be developed as components or parts in isolation during a research stage. When these costs are associated with a particular system or equipment they must be included.

7-2.1.1 Investment

M-related elements of investment cost include those for prime equipment, support equipment, system test and evaluation, system engineering management, training, data, new operational facilities, and repair parts.

7-2.1.1.1 Prime Equipment

It is often difficult to estimate the *M*-related investment costs for prime equipment directly. When comparison of alternatives (trade-offs) are involved, the marginal cost (cost difference) for equipment with particular features compared to identical equipment without these *M*-related features often is used. More conveniently, all costs are summed for each alternative and the totals (with operation and support costs properly discounted) compared. This permits all cost-related differences to be considered and balanced at one time.

7-2.1.1.2 Support Equipment

Special tools and test equipment are included here. Often an item of support equipment is used for many items of prime equipment. In such cases an allocation of costs to each item of prime equipment is sometimes used when costs per piece of prime equipment are being calculated. More commonly, total system costs are calculated, based on an estimate of total field population of prime equipment and all relevant support equipment. Special and common (already in DoD inventory) support equipment are frequently found subcategories.

7-2.1.1.3 System Test and Evaluation

When acceptance testing and incentive/penalty measurements take place on equipment delivery, *M*-related test and evaluation measurements are often included. These tests can cover preventive maintenance, simulated repair, simulated overhaul, and related matters. The cost of executing these tests is included in this cost elements.

7-2.1.1.4 System Engineering/Management

Before and during production, the Government and the contractor will perform system engineering efforts including operation analysis, life-cycle costing, value engineering, human engineering, reliability, and maintainability. Planning for system test and evaluation, and the development test of system, subsystem or component models or mock-ups are also included.

7-2.1.1.5 Training

This cost element covers contractor-furnished training services, devices, accessories, aids, equipment, and parts; it includes the cost of instructors, training plans, and course materials. It also covers Government costs for special training not part of normal skill-qualification training. Personnel, Government facilities, instructors, maintenance trainers, Government-prepared training plans, course materials, training aids, and similar resources are included.

7-2.1.1.6 Data

This cost element includes maintenance data. Technical Manuals, drawings, plans, circuit diagrams, and maintenance manuals are some of the elements for consideration.

7-2.1.1.7 Operational Facilities

For major systems, this cost element includes new facilities. For all systems, incremental expansion of existing facilities is also costed here. When contractors are to perform maintenance and overhaul, any charges to the Government for contractor operational facilities would be included.

7-2.1.1.8 Replacement Components and Repair Parts

This cost element reflects the initial provisioning of replacements and repair parts for pipeline, depot, and field stocks. Replacements are components or assemblies used for maintenance replacement purposes in end items of equipment. Repair parts are those "bits and pieces", e.g., individual parts or nonrepairable assemblies, required for the repair of replacements or end items. The cost element does not include costs for consumption of replacements or repair parts during the life of the system; such costs are charged to operation and support costs.

7-2.1.2 Operations and Support Costs

These cost elements reflect ongoing costs of ownership after a system is delivered and placed in service. They include operational and maintenance manning support; such direct operational support as fuel and electricity; maintenance, repair, and alterations support; material support, and related costs.

7-2.1.2.1 Organizational Maintenance Manpower

This cost is sometimes difficult to separate from operational manpower without careful analyses. Particularly when crew members perform many functions, as in the case of first-line maintenance performed by operating personnel, appeal to maintenance plans, schedules, and mean-time-to-repair data can be misleading. **A** system cannot be manned by a fractional technician. Sometimes a system which requires somewhat higher maintenance man-hours can still be manned with a fixed personnel complement. This cost element is another reminder of the difference between analytic theory and military reality, a distinction which must always be kept in mind when designing or analyzing systems. It is customary to distinguish between corrective and preventive maintenance when considering subelements of maintenance manpower. Included in the cost of maintenance manpower is the cost of replacement training during the life of a system, as personnel rotate.

7-2.1.2.2 Maintenance, Repair, and Modification Support

This cost element is the one most commonly considered when designing for reliability and maintainability. Because of its importance, it is further subdivided here into its components: modifications, maintenance and repair (other than overhaul), and overhaul.

7-2.1.2.2.1 Modifications

This cost reflects improvements or retrofits, including installation of next-generation equipment in a system after it has been in the field. It can include improved automated test equipment built into larger systems.

7-2.1.2.2.2 Maintenance and Repair (Other Than Overhaul)

This element covers labor, material, and overhead for maintenance and repair, other than regularly scheduled overhauls, conducted by personnel 'of a maintenance facility. It does not include organizational maintenance and repair, which is covered under Materiel Support (for parts) and Maintenance Manpower (for labor).

7-2.1.2.2.3 Overhaul

This element is used to accumulate depot overhaul costs for all labor, material, and overhead required during regularly scheduled overhauls for open-and-inspect procedures; maintenance, repair and refurbishment; and revision of technical data to reflect overhaul actions.

7-2.1.2.3 Organizational Material Support

This element covers replenishment repair parts and replacement items, and repair of replacement components, necessary for organizational maintenance. It is usually subdivided into replacement components, and repair parts. Replacement components are usually subdivided into replacement items and repair of replacement items, since maintenance doctrine often calls for subsequent repair of the replaced "defective" item.

7-2.2 SYSTEMS AND SUBSYSTEMS

The basic concept of improved maintainability for reduced total cost of ownership and the fragmentation of costs of ownership into separate, individually analyzed elements is directly applicable to individual pieces of hardware. Par. **7-3** discusses the actual estimation of individual element costs. Let us suppose that we have these estimates. Then it becomes a matter of mathematical calculation to expand the concept to subassemblies, subsystems, and complete systems. When the building-block elements (pieces of hardware) have independent costs (as, for example, for sections of maintenance manuals dealing with each element), they are additive. Many costs involve joint economies and must be calculated more carefully. **A** maintenance manual on a subsystem requires only one binding, while individual manuals on subassemblies are more costly. Preventive maintenance tests on systems can use checking procedures which exercise major portions of a system, or the entire system, while separate checking procedures for subsystems or components will usually require more time and effort. When a system is designed with high redundancy, corrective maintenance-related manpower and material consumption will often be reduced since the system can continue to operate with partially degraded components (reducing the out-of-service time, the number of system spares, and increasing the number of malfunction corrections which can be made at one time, with the resulting efficiency of maintenance and repair).

7-2.3 COST ELEMENT BREAKDOWN TECHNIQUES AND CODING SYSTEMS FOR COST STRUCTURES

Par. 7-2.1 discusses individual cost elements important in maintainability, and introduces some principles for identifying individual cost elements, with illustrations of principal *M*-related elements. The basic breakdown structure implied by these elements incorporates several further principles, similar to those in work breakdown structures used for program monitoring and control. The structure can be considered an inverted tree-like arrangement, beginning with the complete system costs as the topmost element, branching out to more and more levels of detail. Each increase in detail should reflect a consistent subdivision. If a hardware category is being subdivided, subdivisions should be conceptually comparable. Subdivisions of a tank might include the weapon system, the drive system, the guidance system, and the armor system. It would be inappropriate to subdivide a tank into weapons, engine, wheels, transmission, guidance, and armor, since the transition from tank to engine, wheels, and transmission omits a level of aggregation otherwise included. Each lower-level subdivision reflects functional subdivision in hardware systems. MIL-STD-881 discusses Work Breakdown Structures.

In subdividing cost elements it is important to use a consistent approach to subdivision, wherever possible. Conflicts will often arise, and must be thought through carefully. For example, in a hardware subdivision for a system which includes both contractor- and Government-furnished equipment, it is often necessary to estimate ownership costs separately for contractor-provided and Government-provided objects and services (including contractor-provided operation and support, such as at a manufacturer-operated overhaul facility). Should there be two identical structures, one for contractor costs and the other for Government costs, coming together at the top? If the Government buys services at different points in the program, the structure should reflect this—perhaps through subdivision at the appropriate phasing point. Thus investment costs would be subdivided into contractor and Government elements; operating costs would contain a few elements for contractor-furnished services at the proper point of management integration in the system. A hardware-related subdivision of operating cost elements such as repair parts, which shows the prime equipment end use of each category of repair parts presents certain problems. Such a requirement can be met in a linear (tree-like) structure only if the dimensionality changes abruptly. Thus operation costs can be subdivided into categories including material usage; material usage can be subdivided to include replacements and repair parts; repair parts can be subdivided into classes by major hardware use. Once the dimensionality changes,.it is difficult to return to the earlier principle of subdivision. A better way of handling the situation is to use a separate structure for each major dimension when that dimension is expected to apply to many elements in the system. The different structures would be coded to the most detailed elements, which could then be aggregated upward as desired. Thus a particular cost can be for repair parts, for parts to be applied to a tank transmission, for Government-manufactured parts, for an arctic version of the tank. This single cost can then be accumulated upward through parts, to materials, to operation and support costs—still for the tank transmission for Government-manufactured parts for the arctic version. The cost can instead be accumulated upward to Government- and contractor-manufactured repair parts for the arctic tank transmission; accumulated to Government-manufactured tank transmission repair parts for all climates, and many other summary categories.

Associated with particular breakdown techniques are numbering or coding systems which imply the breakdown technique.

Table 7-1 shows a simple breakdown for *M*-related cost elements. Each level of the structure is represented by a digit position in the code. If more than 9 subdivisions exist at any level, the digits may be supplemented by letters of the alphabet.

Table 7-2 shows a linear breakdown structure with an abrupt change in dimensionality. Finer functional subdivisions of operation and support costs are broken off to change to a hardware-associated code for some materials. Note the complexity of the code, once we introduce a hardware breakdown into the operation and support structure. Note also that without an implied intermediate structure, we cannot present the cost of replacement components for the Mark Z Gun. Above 331.1 I1 in Table 7-2, the gun costs are aggregated into weapon costs, replacement parts costs, and then spares costs. To get gun spares, we would have to add 331.111 to other gun-related spares costs. To handle such problems easily, we introduce the concept of multiple coding structures, or "threaded lists". Each characteristic of a cost by which we may wish to report it is given a separate coding structure. Some of these structures are simple divisions, such as into contractor costs and Government costs. Others may be more complex, like hardware breakdowns, functional breakdowns, or mission-related breakdowns. We then select the appropriate code from each structure to categorize a particular cost. Highly flexible aggregation and reporting then becomes possible with a data processing system. Taking the example of par. 7-2.3, we might have a hardware coding system preceded by the letter "H", a source code (C for contractor; G for Government), a basic functional cost breakdown with no prefix, and a climate breakdown (where 3 = arctic) preceded by the letter "K". The repair parts cost for the

TABLE 7-1.
SIMPLE *M*-RELATED COST BREAKDOWN

000.	Tank System
100	Research and Development
200.	Investment
210.	Tank
211.	Weapons
211.1	Mark Z Gun
211.11	Gun Bearing Assembly
212.	Drive System
212.1	Transmission
220.	Support Equipment
230.	System Test and Evaluation
240.	System Engineering/Management
241.	Maintainability Planning
242.	Test Planning
249	Inventory Introduction
250.	Training
251.	Equipment
252.	Services
253.	Facilities
260.	Data
261.	Manuals
261.1	Maintenance Manuals
270.	Operational Facilities
280.	Spares and Repair Parts
300.	Operation and Support
310.	Maintenance Manpower
320.	Maintenance, Repair, and Alteration Support
321.	Alterations
322.	Maintenance and Repair (Other Than Overhaul)
323.	Overhaul
330.	Material Support
331.	Spares
331.1	Replacement Parts
331.2	Repair of Replacement Items
332.	Repair Parts
333.	Transportation
340.	System Engineering/Management
348	Supply cost
349	Inventory Maintenance

TABLE 7-2.
COST BREAKDOWN DIMENSIONALITY CHANGE

300.	Operation and Support
310.	Organizational Maintenance Manpower
320.	Maintenance, Repair, and Alteration Support
321.	Alterations
322.	Maintenance and Repair (Other Than Overhaul)
322.1	Weapons
322.11	Mark Z Gun
322.111	Gun Bearing Assembly
322.2	Drive System
323	Overhaul
323.1	Weapons
323.11	Mark Z Gun
323.111	Gun Bearing Assembly
323.2	Drive System
330.	Material Support
331.	Spares
331.1	Replacement Parts
331.11	Weapons
331.111	Mark Z Gun
331.111.1	Gun Bearing Assembly
331.12	Drive System
331.2	Repair of Replacement Items
331.21	Weapons
331.211	Mark Z Gun
331.211.1	Gun Bearing Assembly
332.	Repair Parts

tank transmission for Government-manufactured, arctic version parts might then be coded (see Table **7-2**) as **332H2.1GK3,** or, if uniform separators are used to avoid prefixes, **332/2.1/G/3,** instead. Such a code can be aggregated upward by any combination of dimensions through simple techniques, to solve many of the dimensionality and aggregation problems imposed by more rigid linear coding systems.

7-3 COST ANALYSIS

We have seen in par. **7-2** how costs are categorized and subdivided into elements. Par. **7-3** now discusses the analysis of the costs for these individual elements. Par. **7-3.1** discusses the use of related historical data for estimation of system costs in the daily concept stages, before experience with a proposed or new system has been accumulated. Par. **7-3.2** treats specific data to be collected, and methods for obtaining that data. Par. **7-3.3** uses some historical data to treat the influence on manpower consumption at different states. Par. **7-3.4** concludes Section **I** with a discussion of the cost elements, data, and methods used for throw-away versus repair analyses.

7-3.1 COST ESTIMATION IN EARLY CONCEPT STATES

When we plan a new system, it is important to be able to estimate costs in early concept stages for budgeting and planning purposes, and to compare alternative maintenance policies, *M*-related design alternatives, and other elements of system design under our control. Detailed discussion of such analyses and tradeoffs will be presented in pars. **7-3.3, 7-3.4, 7-5,** and **7-6.** Since we have no hard data on the costs of the planned system or the many hundreds of alternate candidates which may be considered and rejected to arrive at the selected system, we must estimate costs through indirect methods. Using historical data and analogous systems, we can make such estimates of individual element costs and total them, or we can make direct estimates of total system costs.

7-3.1.1 Analogous Systems

An analogous system is one sufficiently like a candidate system in one or more respects that it can provide a basis for estimating element costs or total costs. We will refer to the analogous system as the "source system" and the system being designed as the "weapon system". The process of cost estimation involves the following steps:

a. Select one or more source systems, for estimating particular elements of weapon system costs.

b. Identify similarities and differences between source and weapon systems.

c. Develop a method for adjusting source system costs to be comparable with the weapon system costs.

d. Collect required source system cost data.

e. Adjust for weapon system.

f. Develop weapon system costs.

Consider a new tank (call it the Mark Z), with a heavier frame and more powerful engine, fuel cell power, use of new guidance electronics, and a new type of radar gun director. To begin, we would like a first-cut estimate of total costs of this weapon system. One way to obtain such an estimate is to seek several source systems, each having common features to the weapon system such that in total all features of the weapon system are represented. By analyzing historical data on each source system related to the weapon system features, it may be possible to separate out the cost data for each weapon feature and use them. A second method would be adjustment of source system costs in order to scale them to weapon system complexity or magnitude. An example will make both of these approaches clear.

7-3.1.2 Example: Source Systems With Weapon System Features

Table 7-3 shows several source systems and the weapon system design elements of interest to this example. Table 7-3 shows (see * items) that the Mark 2 Tank has a frame and engine of similar type to the weapon system, that the Mark 3 Armored Personnel Carrier has a similar power source and guidance system, and that the Mark 1 Field Gun has a similar gun director. The Mark 1 Tank apparently has nothing in common with the Mark Z Tank; its cost data would not be used. It remains to determine what the degree of similarity is in each source system, and which cost elements are of concern.

Suppose the Mark 2 Tank had an almost identical frame weight, propulsion system (including transmission, gearing, and tracking), and engine design. We might try that analog to the weapon system without further modification. Suppose that the fuel cell system for the armored personnel carrier had only 75 percent of the capacity of that of the weapon system, but was of identical design. Then through statistical or engineering analysis we would develop direct conversion factors for translating the costs of the source fuel cell system to those of the weapon system. If we had the data or could get it, we might perform regression analysis on fuel cell systems of several different sizes, to relate number of cell-units or capacity to maintenance costs. If we discovered a stable relationship for a series of systems, we might interpolate or extrapolate to the weapon system size in order to predict costs. Alternatively, we might develop an engineering relationship based on number of cells, manner of connection (series and parallel), cell-unit MTBF (and perhaps *MTTR)*, and other physical parameters, to predict weapon system costs. Of course, realism must apply to such cost extrapolations by size. A hole in a small tank could cost just as much to fix as a hole in a big tank. Par. 7-4 discusses the development of statistical and engineering relationships as a part of model development.

Suppose that the Mark 3 Armored Personnel Carrier had an identical computer-assisted guidance system. We would likely still need to develop conversion relationships because of the different deployment and mission factors (including operating profiles, environmental conditions, mission length, and stress) to convert the armored personnel carrier guidance costs into reasonable estimates for the weapon (tank) system.

Suppose the field gun radar were identical to the weapon system gun direction radar. We would still have to develop conversion relationships to the weapon system because of the vibration differences between a fixed and mobile system, and because of the different interface environments (including the differential effect of the external vs fuel cell power supplies).

Thus we see that many conversion factors need to be allowed for in using analogs for cost estimation. They include design factors such as size and weight; environmental factors such as load, stress, combat conditions, and mission profile; and interface conditions. The identification of conversion variables and the development of conversion factors is not an exact science, but requires appeal to reason and judgment, even after statistical analysis of experience data.

TABLE 7-3.
COST DATA SOURCES

	Mark 1 Tank	Mark 2 Tank	Mark 1 Field Gun	Mark 3 Armored Personnel Carrier	Mark Z Tank (Proposed)
Frame/ Engine	Light	Heavy"	N/A	Medium	Heavy
Power	Storage Battery	Storage Battery	External	Fuel" Cell	Fuel Cell
Guidance	Visual	Visual	N/A	Computer* Assisted	Computer Assisted
Fire Control	Visual	Visual	Radar"	N/A	Radar

7-3.1.3 Prior System Data

When all or part of a weapon system is in existence, the cost analysis problem is simplified. Care is still required to be sure that the source system (in this case, all or part of the weapon system) is similar. Often a weapon system is available in prototype or field test form, or in a version designed for some initial mission. Conversion of its cost data before use will still be necessary if mission profiles or other factors differ between the system from which data are collected and the system being planned.

7-3.2 PREDICTION AND COST ANALYSIS DATA

Costs are predicted through use of the results of cost analysis. Cost analysis is based on resource data since costs are incurred through consumption of resources over time. Money is a convenient, common, additive measure of resource consumption; use of men, repair parts, and fuel may all be measured in dollars, and added. Data are usually collected in resource units, and then converted to dollars. Obtaining and converting data often present difficulties of definition, identification, capture, and use.

Identification of data presents another set of problems. What data have been or can be collected? Can they be used to measure relevant costs? *M*-related investment costs are based on the initial acquisition of prime mission equipment, with its *M*-related features, on the acquisition of related support and test equipment, data, training, on the construction, expansion or organization of maintenance facilities, and on the initial pipeline spans and repair parts (which are, by convention, assumed to be an element of investment costs, although they are "consumed" as is the prime mission equipment, over the life of a program). *M*-related operating cost elements include men used over time for maintenance at organizational, direct support, general support, and depot level; material consumption of nonrepairable and repairable parts at these same levels; and related integrated logistic support, data, and transportation costs.

Data for identifying these costs are obtained from contractual and "returned" (experienced) costs, manpower descriptions, field data collection systems, planning factors, policy guidance prescribing resource levels, engineering variables derived from equipment designs, technical reports and papers, and budgets.

7-3.2.1 Manpower Data

Manpower data in early concept design may be obtained from similar systems with similar maintenance doctrines, through examination of organizational and maintenance documents. Maintenance data systems which collect manpower consumption data can also be used. The most effective method for manpower data development is through direct examination of the conceptual structure of the system being designed, by direct calculation of likely manpower needs given a maintenance policy, and (where maintenance policy is a design variable) by calculation of manpower implications of several alternative maintenance doctrines which may be considered in a search for the best alternative. Such an analysis must consider skill levels required for maintenance as well.

7-3.2.2 Material Data

Material consumption data may be calculated as a detailed part or component level from *MTBF* data, mission and operating profiles, and part population data. In early concept design, more aggregated estimation procedures are required because of the lack of full system definition, the limited time available for analysis, and the number of alternative hardware configurations usually considered at this stage. More aggregated cost estimating relationships must be used, based on cost models (see par. 7-4), derived from historical data on similar systems and subsystems, suitably adjusted for reliability, mission and operating profile, and other relevant differences.

Obtaining historical data for relevant source components and subsystems is usually difficult; compromise hardware and factoring of costs to translate to the weapon system are often necessary. Mission and operating profile data on the source system may have to be inferred; such data for the weapon system may have to be standardized through the assumption of an "analytic scenario" or several scenarios. When a weapon system of increased reliability or maintainability is considered, it is important to appeal to similar existing subsystems, or to components with similar parameters; simple assertion of higher reliability or maintainability is not enough. Poorly validated claims otherwise can lead to sharp cost escalation once the equipment is in the field.

7-3.2.3 Time Data

Time for maintenance actions is a critical generator of cost variables. Given a target level of operational availability A_o the maintenance and downtimes for equipment meeting A_o are given by (see Eq. 1-26):

$$A_0 = (MTBM + RT)/(MTBM + MDT + RT) \tag{7-1}$$

where A_o is the long term steady-state availability of a system or equipment operating in its use environment and performing its required missions and functions. In this equation *MTBM* is mean operating time between maintenance actions, both corrective and preventive, that make a system unavailable; *MDT* is mean downtime on account of failure, preventive maintenance, or logistic delay, and ***RT*** is the system average ready time in a cycle of ***MTBM*** + RT + ***MDT***.

Further,

$$MTBM = (MTBPM)(MTBCM)/(MTBPM + MTBCM) \tag{7-2}$$

where *MTBPM* is mean operating time between preventive maintenance and ***MTBCM*** is mean operating time between corrective maintenance.

The mean downtime is

$$MDT = MFDT + MMT + MMDT \tag{7-3}$$

which is composed of *MFDT*, mean fault detection time, a function of fault monitoring methods and frequency; the mean maintenance time, *MMT*, for active (equipment unavailable for use) preventive and corrective maintenance, i.e.:

$$MMT = [(MTBCM)(MPMT) + (MTBPM)(MCMT)]/(MTBCM + MTBPM) \tag{7-4}$$

where ***MTBCM*** and ***MTBPM*** are as before; ***MPMT*** is mean preventive maintenance time; ***MCMT*** is mean corrective maintenance time; and the remaining term *MMDT*, which is the mean maintenance delay time, is the average delay in preparing for maintenance action-composed of average personnel assignment time, average preparation time, and average replacement items requisition time (Ref. 5).

If a system is designed so that no preventive maintenance actions that cause system downtime are required, and if personnel staffing and replacement item provisioning make maintenance personnel and replacement items always available, A_o approaches the familiar "designed-in" inherent availability

$$A_i = MTBF/(MTBF + MTTR) \tag{7-5}$$

Since the consumption of manpower and replacement items is intimately related to the times described, we see the importance of time in economic analysis and design for maintainability.

7-3.3 MAINTAINABILITY AND MANPOWER CONSUMPTION

In the process of resource consumption, maintainability determines the amount of manpower needed to effect repairs at each stage. The physical and policy design for maintainability will influence the ability to perform maintenance at each level and the kinds of maintenance possible. The maintainability parameters determined by these physical and policy design decisions will (through preventive maintenance time and *MTTR*) influence directly manpower consumption at each stage. If built-in test equipment, for example, is a rational part of the design for maintainability of a piece of equipment, we would expect the time for fault detection and isolation to be reduced, compared with that for the same equipment without such features. Higher maintainability induces reduced time for maintenance and thus reduced manpower consumption (assuming men are a scarce resource and the man-hours saved in maintenance time may be productively used elsewhere). Several alternative levels of built-in test equipment might be considered early in concept design. Similar effects can be expected from easier access for faster open-and-inspect procedures. Several altematives might be considered at higher maintenance organizational stages; design alternatives can include direct repair or repair by replacement to speed up repair. At still higher levels, a centralized, specialized module repair facility can reduce manpower through economies of scale. In each case where required availability can be obtained without these procedures, an economic trade-off is necessary to compare the savings from reduced manpower (and other sources) with the costs of design for higher maintainability.

Once cost data and cost analysis have produced prediction methods for each element of cost, such costs may be calculated for each alternative M-related policy and hardware design to be considered. The resultant cost analysis can be used to select from among these alternatives. Such comparisons (trade-offs) are discussed in par. 7-6. To illustrate the cost elements to be considered, a brief discussion is presented in par. 7-3.4.

7-3.4 METHODS FOR THROW-AWAY VERSUS REPAIR DESIGN DECISIONS

The basic principle of an economic trade-off is to calculate all costs for each altemative being considered (i.e., meeting mission and performance requirements including availability) and compare the results. Since each alternative admitted to the economic trade-off meets noncost requirements, the least-cost alternative (after allowance for estimation error and uncertainty) should be selected. Suppose we wish to establish when items are economically thrown away and when they should be repaired. We wish to pick the lower-cost alternative between throw-away and repair. What are some factors to be considered?

The following discussion is adapted from Ref. 1. First, the question of decision point—if we are in a design state, we must address overall system economies; if we are in a provisioning stage, the hardware is fixed and many costs are sunk—a more limited set of costs and benefits associated with sparing levels is appropriate; if we are at the "moment of truth" when a module has failed, we can consider the most limited set of costs related to current mission conditions, resupply time and stock levels as the system is operating (as distinct from the theoretical assumptions in policy design). We must also consider the level at which the decision is to be made, bearing in mind multiple options at the decision point: repair failed module, discard failed module, replace and/repair failed module at organizational or one of several higher levels.

The cost elements to be considered in the most general case (in specific cases, particular elements may be ignored) include:

1. Cost of hardware, including alternative levels of module complexity (measured, for example, in transistors or circuits per module)
2. Cost of test equipment and tools
3. Cost of manpower by skill levels, and cost of training
4. Cost of repair parts
5. Cost of supply, administration, and cataloging
6. Cost of replacement parts
7. Cost of repair facility
8. Cost of packaging and shipping.

In making the decision at the design level, a cost estimate is prepared for each alternative, showing the cost of each appropriate element. The costs for each altemative (with operating and support costs over the life of the equipment suitably discounted) are added and the totals compared in order to arrive at the design decision.

SECTION II

COST ANALYSIS

This section discusses the development and use of mathematical models in cost analysis. It begins in par. 7-4 with a treatment of life-cycle cost model development, from the basic parameters and equations through their combination into complete models of annual (average) costs and life-cycle costs. The exercise of such models to produce the expected costs of a particular system, subsystem, or design alternative is discussed next, in par. 7-5, which also discusses the use of models to consider several alternatives, a range of output parameters (availability, performance), and to perform sensitivity analyses of cost versus manpower or skill utilization. Next (par. 7-6) is a discussion of formal trade-offs in which economics plays a key role in M-related considerations. The concluding part of this chapter, par. 7-7, discusses life-cycle costing as a decision-, procurement-, and performance-monitoring tool.

7-4 MODEL DEVELOPMENT

Cost models are made up of a collection of equations called cost estimating relationships (CERs) which estimate individual element costs. Each CER contains variables describing resource consumption, and parameters reflecting prices, conversion factors, or empirical relationships combined into an equation. The equation, in general form, is called the "structural form". Cost estimating relationships include those for direct calculation, those using physical and engineering relationships, and those derived through empirical statistical methods from historical cost data. This paragraph describes the parts of cost models, discusses specific M-related cost element parameters and variables, and presents a number of CER examples.

7-4.1 COST ESTIMATING RELATIONSHIPS AND BASIC BUILDING BLOCKS

7-4.1.1 Parameters and Planning Factors

A parameter in a CER reflects a conversion factor from one system of units to another. It may be a price, an empirically derived ratio, or a policy parameter. A price like cost per man-hour, for example, converts man-hours into dollars. An example of an empirical ratio is the number of man-hours of corrective maintenance per failure of a given component, which may be obtained as a statistical average. An example of a policy parameter is the number of parts per module. Such parameters enter into cost estimating relationships and often are compiled and published as planning factors.

7-4.1.2 Variables

A variable in a CER characterizes resource consumption over time. It may be a physical or performance measure. Variables generate costs. Examples of variables include failure rate, preventive maintenance man-hours per unit of equipment, and hardware design characteristics.

7-4.1.3 Equations

An "engineering" equation, or "engineering" cost estimating relationship, reflects our belief in the underlying mechanism or relationship which generates costs. Often when a detailed theoretical relationship cannot be developed, particularly for hardware costs, a statistically derived empirical relationship is used. The statistical CER is a simplification or "short cut", and is not necessarily a representation of a physical situation. An example will make the difference clear. Suppose we wished to estimate the annual shipment cost of a type of failed module to a fixed-site depot. If we knew the weight per year of such module (say, W) and a cost per pound of packing and shipping for this module-type (say, CP), then an engineering relationship (reflecting the physical prices of shipment) for annual costs CA might be

$$CA = W(CP) \tag{7-6}$$

Suppose, instead, we had no way of obtaining a direct variable such as weight to measure shipment cost. We might infer that the cost varied with the number of units shipped, which in turn might reflect a fixed number of failures per year, plus a variable number dependent on mission hours per year. We could collect historical data on annual shipment costs and mission hours for a number of years and attempt to fit an equation to those data through statistical methods. A reasonable equation might be

$$CA = a + b(MH) \quad (7\text{-}6a)$$

where a represents the fixed costs per year, b the shipping cost per mission hour per year, and MH is the annual mission hours. The parameters a and b would be estimated by the method of least squares (see Ref. **4** for an excellent treatment of statistical cost model construction and estimation). If we thought, instead, that the rate of increase in shipping costs decreased as mission hours increased, due to "burn-in" effects, we might add a parameter c to reflect this scale effect, and estimate the parameters of the equation

$$CA = a + b(MH)^c \quad (7\text{-}6b)$$

through least squares. We might compare the "fit" of these equations to the historical data through the use of measures of merit of each equation, such as the standard error of estimate, coefficient of variation, or the coefficient of determination R^2.

7-4.1.4 Engineering Cost Estimate

Sometimes particular cost elements can be directly estimated, particularly during middle and late design stages, by examining a system component-by-component. While this estimating method is the most commonly used one, based on detailed "pricing" wherever possible, experience has shown it to be extremely inaccurate and unreliable, despite the appearance of detailed analysis. Particularly at early stages, for costing large components, subsystems, and systems, engineering cost estimates tend to escalate radically over time. It may be thought of as a "legislated" cost, as in the case of a fixed price contract, but recent experience in military procurement shows that even these costs will escalate.

7-4.2 STATISTICAL CER DEVELOPMENT

The first step in developing a statistical **CER** is to define the dependent variables to be sure that data collected are consistent.

The second step is to identify a list of possible independent (predictor) variables. Each candidate independent variable must be related logically to the dependent variable; the **CER** finally obtained must be defensible on rational, as well as statistical, grounds. With the existence of rapid, economical, sophisticated computer programs for estimating the parameters of **CER's,** it is often all too easy to produce statistical relationships with attractive measures of fit which are nonsense.

The third step is to identify structural forms for the **CER's** which also make sense. Simple linear relationships of the form:

$$C = a \quad (7\text{-}7)$$

a constant, or

$$C = bx \quad (7\text{-}8)$$

a linear homogeneous function, are often appropriate for fixed costs or costs which increase linearly with the independent variable, respectively. Eq. 7-7 might reflect the construction costs of a fixed-size depot; Eq. 7-8 is often used for pay and allowances as a function of number of personnel, construction costs as a function of square feet, or facilities maintenance cost as a function of facilities initial investment cost (Ref. **4**). Combining the two we have

$$C = a + bx \quad (7\text{-}9)$$

a typical "fixed-plus-variable" cost estimating relationship.

If we have reason to believe that economies or diseconomies of scale exist, we might consider

$$C = a + bx^d \quad (7\text{-}10)$$

where a, b, and dare constants to be determined. In this relationship, a is an estimated parameter representing fixed cost elements, b represents increasing or decreasing costs per unit over the range of interest, and d represents increasing or decreasing returns to scale. An economy of scale occurs when the cost per unit decreases as the number of units increases. **A** large, specialized repair facility, fully loaded, usually has a lower cost per repair than several smaller facilities with the same total capacity. The reasons for this include the ability to spread fixed costs (major equipment, land, buildings) over a larger number of units (in this case, repairs). Diseconomies occur at still larger scales. People get in each other's way, overtime is necessary after capacity has been reached, facilities are too large to operate efficiently. Par. 7-6 discusses this topic further.

The relationship is estimated by a statistical fit to the data, but underlying physical interpretation of the estimated parameters adds confidence to the validity of the CER which would be absent in the case of a purely statistically derived relationship.

Multivariate statistical CER's are often used, of the form

$$C = a + bx + dy + \cdots \tag{7-11}$$

or

$$C = ax^b y^d \ldots \tag{7-12}$$

among others.

Often, a "polynomial" form is used (Ref. 6), subsuming many of the previously given forms as special cases:

$$C = a_1 x_1^{b_1} + a_2 x_2^{b_2} + a_3 x_3^{b_3} + \cdots = \sum_{i=1}^{n} a_i x_i^{b_i} \tag{7-13}$$

In the polynomial form, letting b, = 0 sets the first term to a constant a,; letting b_1 = 1 produces the product a, x_1. Many physical situations produce cost relationships of polynomial form; a powerful optimization technique, geometric programming, has been developed for rapid solution of constrained design problems having polynomial cost functions (Ref. 6). The fourth step is to fit different structural relationships to the data, with various combinations of relevant-seeming variables. Many curve-fitting methods exist for fitting statistical CER's to data. Christ (Ref. 7) treats the subject in great detail from an economic point of view. A more accessible treatment is contained in Kane (Ref. 8). Numerous computer programs exist for fitting. The most commonly used (Ref. 9) is BMDO2R, a stepwise linear regression program. For a detailed treatment of the proper development of CER's and the use of this method, with examples derived directly from commonly used computer programs, Draper and Smith (Ref. 10) cover the subject completely, albeit with a physical science orientation.

In attempting to identify appropriate sets of variables, it is often helpful to examine the correlation matrix of the dependent and independent variables, in order to select meaningful combinations, and identify dangers. One common pitfall is the mistaken use of a variable to explain costs which is not the generator of costs, but correlated with some cost-generating variable. For example, subsistence costs will usually be well correlated with pay. But "explaining" pay using subsistence costs would be misleading. The correct explanatory variables are manning and skill levels. As subsistence costs changed over time, the first relationship would produce increasingly unreliable results.

In selecting from among alternative sets of independent variables, care must be taken to avoid use of variables which are themselves mutually correlated. This "multicollinearity" may be reduced or avoided by performing a principal components analysis on all variables to select an uncorrelated subset for model building. The BMDO2R program (Ref. 9) may then be used to "build" a CER step-by-step, considering the variables in order of explanatory power. Ref. 10 discusses this process in detail, pointing out the many pitfalls and dangers along the way.

The fifth step is to compare alternative CER's and select one for use. There are several statistical measures which can be used to evaluate each CER in absolute and comparative terms (Refs. 6, 8, 9, 10). They include:

a. The standard error of estimate, which indicates the magnitude of error in the CER's fit to each data point used to construct it. Assuming a normal distribution, two-thirds of the fitted points lie within one standard error of the actuals, and 95% of the fitted points lie within two standard errors of their actual points.

b. The coefficient of variation indicates the relative standard errors (standard error of estimate divided by sample means of dependent variables).

c. The standard error of the regression coefficients, which indicates the range around each estimated regression coefficient where the true regression coefficient is likely to be. There is a chance of 0.67 that the true coefficient is within plus or minus one standard error of the regression coefficient and a 0.95 chance that it is within two standard errors. A standard error as large or larger than the coefficient being estimated is poor. If a fitted CER for $C = a + bX$ had an estimated form $C = 5 + 0.7X$ and the standard error of b was 0.7 this means chances are 0.67 to 1 that b is "really" somewhere between 0 and 1.4. Suppose, on the other hand, that the standard error of b was 0.1. Then we are 67 percent sure that b is between 0.6 and 0.8, and 95 percent sure that b is between 0.5 and 0.9. A t-test (Refs. 7, 8, or 10) would confirm the significance of the estimate of b. If there were high probability that b = 0, X is not a significant explanatory variable for cost C.

d. The coefficient of multiple determination R^2 and multiple correlation R; R^2 is the proportion of total variance we have explained by the cost estimating relationship.

e. The Theil U-statistics, which measure the proportions of estimating error due to misestimation of the mean (estimates consistently high or low), of the variance (estimates consistently worse the larger, or smaller, the value we are predicting), and the covariance (Ref. 11).

f. The Durbin-Watson statistic, which measures "runs" of error in our fit, and indicates a misestimated or incompletely specified structural form, or a violation of the basic least-squares assumptions behind the statistical fit (Refs. 7 and 8).

A further check on different CERs is to use them for data points withheld from the estimating data base, and see how well they predict. The CER estimates also can be plotted against the values of independent variables, to see if they give nonsense curves or turning points.

The sixth step is to report one's work. Fisher (Ref. 4) suggests presenting a report on CER development as follows:

1. A summary of background research, including information about trips to field installations and initial impressions about the hypothesis to be tested

2. A complete presentation and description of the raw data base and the adjustments made to it

3. A description of the hypothesis that seemed worthy of serious examination

4. A description of the testing process itself, indicating the tests used and the reasoning leading to the acceptance of a particular hypothesis (CER) and the rejection of the alternatives

5. A presentation of the complete set of statistical measures pertaining to the accepted regression equation (all adjusted for degrees of freedom). For example:

   a. Standard error of estimate
   b. Relative standard error of estimate (coefficient of variation)
   c. Standard errors of the regression coefficients
   d. The equation for the standard error of forecast (Ref. 12, pp. 568-571, 602, 629-630)
   e. The coefficient of multiple determination and the coefficient of multiple correlation

6. A listing of special warning restrictions that should be observed by users of the regression equation.

A good treatment of the estimating process, with detailed references and examples, is contained in Ref. 10, Chapter 6.

7-4.3 COST ESTIMATING RELATIONSHIPS

This paragraph discusses examples of cost estimating relationships at a highly aggregated, simplified level for incremental full cost estimating (par. 7-4.3.1), to set the stage, describes typical input CER parameters (par. 7-4.3.2), and presents typical CER's using such parameters (par. 7-4.3.3).

7-4.3.1 Generalized Estimation for Budgetary Purposes

One set of generalized cost estimating relationships (Ref. 13) for Army operations and maintenance costs is based on force levels, materiel to levels, personnel levels, and other factors. It is a highly aggregated set of relationships for costing the implications of an incremental force. Materiel costs used for estimation are at "standard cost" (current Army catalog cost) of initial issue materiel items (excluding replacement/consumption, maintenance float, and wartime stockage). The force's materiel must be broken down into weapons, combat materiel items, tactical materiel items, support materiel items, electronic and communication items, missile ground support items, and aircraft (fixed-wing, rotary-wing). Many of the relationships use average costs derived from budgets as margin (incremental) costs for the force increment. Table 7-4 shows some key CER's. These may be useful for generalized budgetary estimation, but must be avoided for M-engineering analyses, since they estimate maintenance costs as a function of acquisition costs. Such an assumption is unwarranted for system and component design and analysis—the implication that increased acquisition costs for automatic test equipment, for example, also increases maintenance costs is not borne out by experience or logic.

7-4.3.2 Detailed Cost Estimating Relationships—Parameters

In development of more detailed M-related cost estimating relationships, one must consider design related factors including MTBF and *MTTR*, which will vary for different equipment within a class, and between equipment of the same type. This paragraph describes some typical cost model input parameters.

7-4.3.2.1 Annual Operating Hours for Equipment (O)

M-related costs are driven by the failure of equipment, itself a function of operating hours. A key parameter—estimated from mission profiles, scenarios, and other data developed by mission analysts—is the operating schedule for the mission equipment. For an-

TABLE 7-4.
INCREMENTAL FORCE DEPENDENT MAINTENANCE RELATED ANNUAL OPERATING COSTS

Cost Element	CER
Maintenance	
Support Maintenance Annual Repair-parts Cost	0.04" Item Cost
All Materiel Items (Except Aircraft)	At Standard Cost
Aircraft	
Fixed-wing	Not Given
Rotary-wing	(5.2 + 0.0042" Helicopter Empty Weight in Pounds)* No. of Flying Hours.
Support Maintenance Civilian Labor	
All Items	0.214" Support Maintenance Annual Repair-Parts Cost (Above)
Depot Maintenance	
Combat Vehicles	0.0046* Item Cost
Tactical and Support Vehicles	0.01056" Item Cost
Electronic and Communication Equipment	0.01027" Item Cost
Missile Systems	0.01155 Item Cost
Aircraft	
Fixed-wing	0.00382" Item Cost
Rotary-wing	0.01686" Item Cost
Residual Depot Maintenance	0.274" Direct Depot Maintenance (Above)

Note: * Indicates Multiplication

nual costing, these data may be summarized as operating hours per equipment in each year. If average annual costs are being estimated, without regard to equipment age and discounting of varying cost flows, average operating hours per year may be used. If the cost details of overhaul and force rotation cycles are to be captured, particularly for derivation of optimal overhaul doctrine, a time profile over the equipment life becomes necessary-calling out operating and overhaul hours in each year. For the purposes of this chapter, relationships will be developed initially for a "typical" year, assuming operating hour data as input from performance and mission analysis. Let O represent operating hours per year in the material to follow.

7-4.3.2.2 Structural Cost of End Item/Component (CAQ)

This cost, which we shall call **CAQ,** represents those elements of acquisition cost of the item or component to be supported represented by the hardware. It ***does not include*** Technical Manuals, training, initial repair parts, or other acquisition cost elements separate from the cost of the physical structure of the end item or component.

7-4.3.2.3 Manpower Per Repair Action (MR)

At each level *I* of maintenance, a particular number of man-hours is required to perform maintenance for each repair action. This parameter may be developed through maintenance engineering analysis, prior experience, or it may be prescribed by policy.

7-4.3.2.4 Manpower Utilization Rate (UI)

Maintenance men are often assigned other tasks or military duties. They are thus not fully utilized on maintenance. Since equipment maintenance requirements are usually expressed in man-hours, a manpower utilization rate is needed to convert from number of maintenance men to maintenance man-hours available at each level *I*.

7-4.3.2.5 Incremental Cost of Manpower (CP)

One simplifying assumption sometimes made in maintenance cost analysis **is** that any "commodity" reaching repair level I has the same cost per man-hour of repair. This assumption is based on the similarity of corrective maintenance "style" at a particular level. The differences in repair cost are captured in equipment ***MTTR*** factors, which vary with repair complexity. We denote repair cost per man-hour at level *I* as CRHI. This cost does not include repair parts nor their logistic support.

7-4.3.2.6 Reorder Cost of Field (Repair) Parts (CHRI, CHTI)

This cost reflects ordering and other logistic cost elements. It is a factor in repair versus discard and stock level analyses. Its value is usually minor compared with material and manpower costs, and, except for special cases, may be excluded from analyses. The factor may be important for repair-discard analyses. Let CHRI and **CHTI** be, respectively, repairable and throw-away costs at level *I*.

7-4.3.2.7 Size and Weight of . Item/Component (EIS, EIW)

These physical data, perhaps summed over all failures per year, are important in calculating transportation and storage costs. In constrained inventory problems where space is limited they often play the role of a "side condition" on a least cost solution.

7-4.3.2.8 Packaging and Transportation Costs (CTRI)

These one-way costs reflect the shipment of items to or from level *I*. The costs are usually expressed as costs per pound or per cubic foot, and are partially dependent on distance. They can have a significant impact on system design analyses, where number and location of repair level facilities are being planned. They are also a contributor to the result of discard versus repair analyses, and of modularization studies.

7-4.3.2.9 Storage and Shop Space Costs (CSI, CSA)

Cost factors are important to inventory size analyses and repair level studies. In many such analyses, it is important to estimate the fixed and variable component of these factor costs, and to capture any size effects that reduce cost per unit (per cubic foot in this case). When facility design studies are involved, it is important to make detailed space analyses rather than gross estimates. These apparently straightforward costs can lead to quite complex analyses in the case of constrained maintenance inventories—where space **is** limited, the "cost" reflects both the space taken up by parts and modules, and the "availability value" of different parts or modules having the same size or weight. CSI is the initial cost per cubic foot of space; CSA is the annual upkeep cost per cubic foot.

7-4.3.2.10 Training Costs (CTM, MHT)

Training costs include those for facilities, instructors, materials, training aids, and trainees. Annual training includes that for replacements and for retraining or upgrading existing personnel. In manning analyses and comparison of alternative maintenance plans, training alternatives may be assumed fixed through derivation of a flat cost per man per year CTM and on estimate of man-hours per trainee per year MHT be used in costing trainee pay and allowances. Initial training must be treated separately as part of the acquisition cost of a system. Basic and recruit training should be excluded. Analyses comparing alternative training approaches cannot use such gross planning factors as CTM and MHT, but must cost a detailed training plan, course-by-course and item-by-item.

7-4.3.2.11 Annual Administrative Cost Per Item (CIA)

These costs are often included for completeness, but usually do not affect the outcome of *M*-related analyses and trade-offs, except in special cases. For example, in a trade-off between a discard-at-failure module or large-scale integrated (LSI) circuit, and repair of individual piece parts (sometimes numbering in the hundreds) that could make up an equivalent set of electronic circuits, the piece parts would be heavily penalized by being charged several hundred times the administrative cost of the single module or LSI circuit.

7-4.3.2.12 Cost of Entry of New Line Items in the Federal Supply System (CII)

This logistic cost for cataloging is similar to CIA in its effect. Except as a penalty against many parts compared with a module, its effect is small.

7-4.3.3 Cost Estimating Relationship Examples

This paragraph describes a number of *M*-related CERs using parameters discussed in par. 7-4.3.2. CERs must be developed for the specific physical and organizational structures being analyzed or planned; they must reflect costs incurred at different levels of maintenance, including the flow of equipment to and from these facilities. CER's must reflect ILS doctrine being planned or analyzed, including provisioning, inventory, and resupply considerations. They must allow for repair policies (discard-at-failure, or repair, or both), and differences caused by components, modules, and special design features (e.g., built-in and automatic test equipment). As a result, CER's and cost models are often "tailored" to the particular case at hand. Par. 7-6 discusses this issue further in connection with trade-offs.

Since no simple set of CER's can be used in all circumstances, the remainder of this section will use an illustrative set largely derived from the TRIM (Throw-away or Repair Implications on Maintenance Cost Programs; Raytheon Co. for the Army) model, but reflecting enough detail for somewhat broader applicability (Ref. 14). The TRIM model reflects a weapon support system composed of batteries, direct and general support units, overseas and CONUS supply depots, and a factory. TRIM terminology is here used to allow the reader to refer back to the original model. The maintenance flow includes throw-away and repairable elements; items of equipment are sent to particular support units because of the nature of the repair facility. The CERs assume "shrinkage" loss of replacement and throw-away units and repair parts, and deal with "peculiar" support items (not yet in the Federal Supply System) as well as "general" support items.

The TRIM CER's, however, are organized by engineering concept rather than cost analysis concept. Many CER's contain several cost elements lumped together. A preferred procedure is the development of CER's by cost element; the TRIM CER's are modified here for that purpose, and a structure based on AR 37-18 (Ref. 15) is used. Estimating methods (or factors from the TRIM model, which are illustrative only) follow each parameter or variable name.

7-4.3.3.1 Related Cost (0.00)

$$C(0.00) = C(2.04) + C(2.051) + C(2.052) + C(2.07) + C(3.04) + C(4.012) + C(4.021) + C(4.022) + C(4.031) + C(4.032) + C(4.05) + C(4.06) \quad (7\text{-}14)$$

where

$C(2.04)$ = publication costs
$C(2.051)$ = building costs
$C(2.052)$ = maintenance equipment costs
$C(2.07)$ = logistic line-item cataloging costs
$C(3.04)$ = acquisition cost of item
$C(4.012)$ = repair cost—maintenance
$C(4.021)$ = repair costs—nonrepairable parts
$C(4.022)$ = repair costs—repairable parts
$C(4.031)$ = logistic processing costs
$C(4.032)$ = logistic cataloging annual costs
$C(4.05)$ = transportation costs
$C(4.06)$ = building maintenance costs

7-4.3.3.2 Publication Costs (2.04)

$$C(2.04) = K(LT)(CAQT) + N(LR)(CAQR) + QR \quad (7\text{-}15)$$

where

K = publication factor—throw-away (0.0056)
LT = Initial inventory of throw-away units (direct estimate)
$CAQT$ = throw-away unit cost (direct estimate)
N = publication factor—repairable (0.0240)

LR = initial inventory of repairable units (direct estimate)
$CAQR$ = repairable unit cost (direct estimate)
Q = number of items of maintenance equipment (direct estimate)
R = maintenance equipment unit cost (direct estimate)

1. K and N given were derived empirically from DoD-wide budgetary ratios and publications ratio differences between throw-away and repairable units.

2. Direct estimation of number of pages of Technical Manuals and cost per page may be preferred.

3. LT and LR, the initial inventories, may be derived through policy and standard inventory calculations, including pipeline considerations.

7-4.3.3.3 Building Costs (2.05 1)

$$C(2.051) = (SS)(CSI) \qquad (7\text{-}16)$$

where

SS = cubic feet required for maintenance buildings at all maintenance levels (direct estimate)
CSI = initial cost per cubic foot ($1 per cubic foot for new cement block)

7-4.3.3.4 Maintenance Equipment Costs (2.052)

$$c(2.052) = C + QR \qquad (7\text{-}17)$$

where

C = R & D cost of maintenance equipment (direct estimate)
R = unit acquisition cost of maintenance equipment (direct estimate)
Q = number of maintenance equipments (direct estimate)

7-4.3.3.5 Logistic Cataloging Costs (2.07)

$$C(2.07) = G(CII) \qquad (7\text{-}18)$$

where

G = number of new line items introduced into Federal Supply System (FSS)
CII = cost of introducing a line item into the FSS' ($25 14—Huntsville)

1. G is the sum of line items in new throw-away modules (1 each), repairable modules (number of new items in all modules), and test equipment. If a particular new item occurs more than once, it counts as one item.

7-4.3.3.6 Acquisition Cost of Item (3.04)

$$C(3.04) = (CAQR)(PR) + (CAQT)(PT) \qquad (7\text{-}19)$$

where

$CAQR$ = acquisition cost, of each repairable item in the system, adjusted for production learning (direct estimate)
$CAQT$ = acquisition cost for throw-away items (direct estimate)
PR = population of repairable items (direct estimate)
PT = population of throw-away items (direct estimate)

7-4.3.3.7 Repair Cost; Maintenance (4.012)

$$C(4.012) = (FR)(1 - S1I)(WM) \qquad (7\text{-}20)$$

where

FR = number of repairable units failing over the system life
$S1I$ = repairable "shrinkage" factor due to loss, damage, etc.
WM = unit repair cost—manpower

These three factors are obtained as follows:

1. The FR factor is given by

$$(FR) = \lambda(PR)OY \qquad (7\text{-}21)$$

where

λ = item failure rate (direct estimate)
PR = population of repairable items (direct estimate)
O = operating hours per year (direct estimate)

Y = system life (10 yr)

2. The $S1I$ factor is tabulated in Table 7-5;

3. The WM factor is

$$(WM) = (UI)(CP)(MR) \quad (7\text{-}22)$$

where

UI = manpower use factor at level I (A use factor of 2 means that 8/2 or 4 hr per day are spent on repair duties.)

1. The UI factor is tabulated in Table 7-6;

CP = manpower cost per hour, including overhead (for military personnel \$8.10/hr is estimated during FY72; for civilians, \$9.60) Figs. 7-1 and 7-2 present estimated costs derived for the TRIM models.

MR = average man-hours per repair action (historical or direct estimate).

These maintenance manpower estimates in the TRIM model exclude those for fault isolation to the defective module (repairable or throw-away), which must be separately estimated, through manning studies or otherwise.

TABLE 7-5.
FACTOR *S1I*

Level	Factor S1I
Operational to Operational	0.0
Operational to 3rd	0.025
Operational to 4th	0.037
Operational to 5th (except CONUS)	0.09
Operational to 5th (CONUS Depot)	0.125
Operational to 5th (factory)	0.1375

TABLE *I-6*.
FACTOR *UI*

Level	Factor UI
2	3
3	1.88
4 and above	1.04

7-4.3.3.8 Repair Costs; Nonrepairable Parts (4.021)

$$C(4.021) = (CAQT)(FTR) + (CAQR)(FRR) \quad (7\text{-}23)$$

where

$CAQT$ = average unit cost of a nonrepairable unit (direct estimate)

$CAQR$ = average cost of repairable unit

FTR = number of nonrepairable units replaced over the system life

FRR = number of repairable units which must be replaced

1. The factors FTR and FRR are given by:

$$(FTR) = \lambda(PT)OY[1 + (ST)] \quad (7\text{-}24)$$

where

λ = item failure rate (direct estimate)

PT = population of nonrepairable units (direct estimate)

O = operating hours per year (direct estimate)

ST = throw-away shrinking factor (0.125)

Y = system life (10 yr)

and

$$FRR = (FR)(S3I) \quad (7\text{-}25)$$

where

FR = number of repairable units failing over the system life (see par. 7-4.3.3.7)

S31 = replacements required for repairable unit shrinkage and supply ineffectiveness at level I

2. The factor S3Z is tabulated in Table 7-7.

7-4.3.3.9 Repair Costs; Repairable Parts (4.022)

$$C(4.022) = (FR)(1 - S1I)(WP) \quad (7\text{-}26)$$

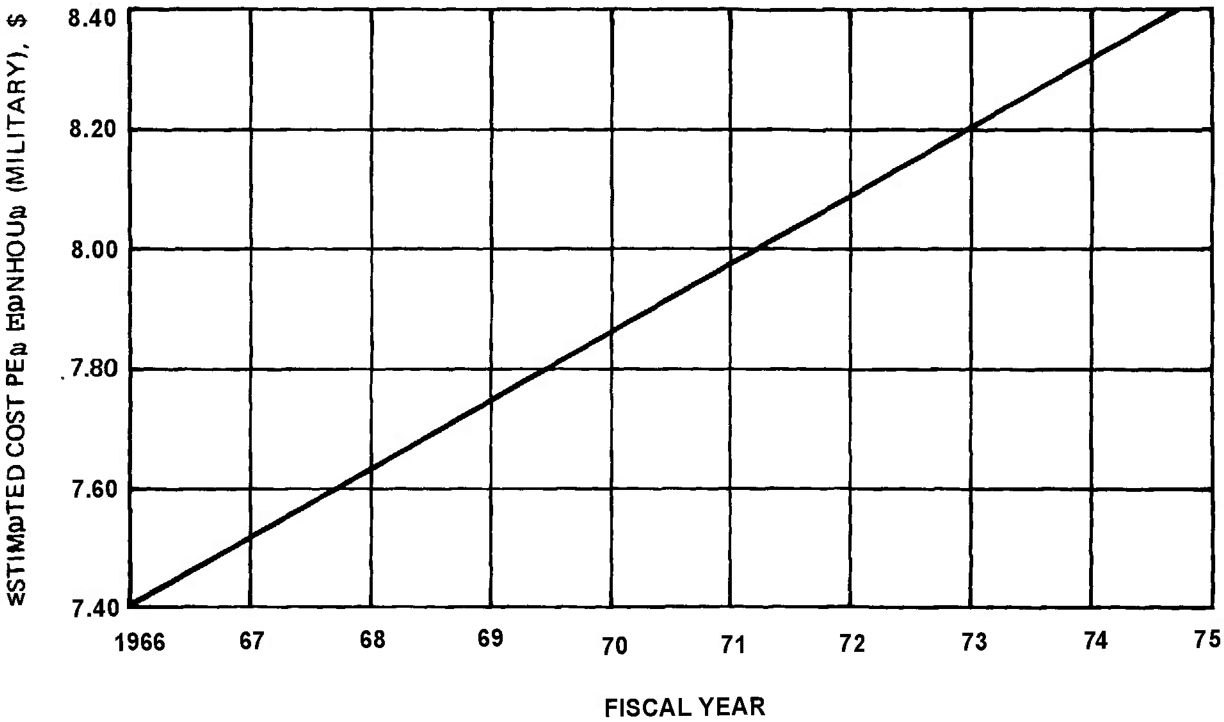


Figure 7-1. Estimated Cost Per Military Manhour vs Time

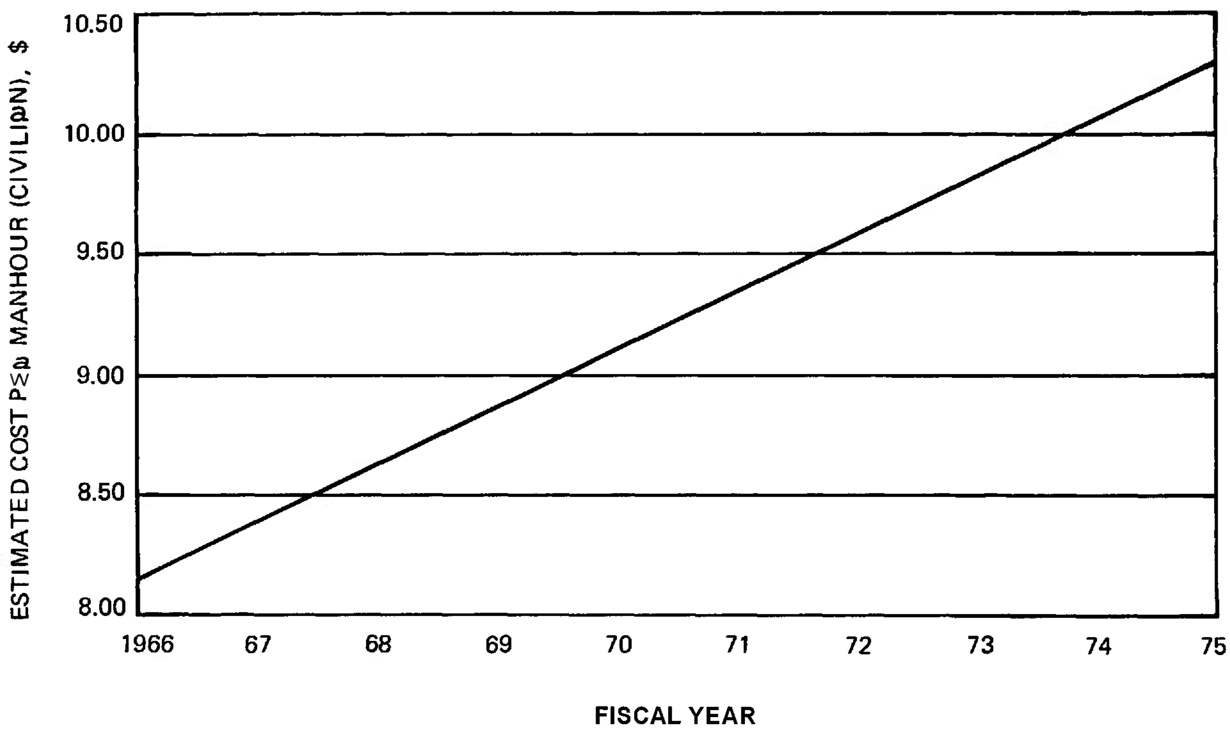


Figure 7-2. Estimated Cost Per Civilian Manhour vs Time

TABLE 7-7.
FACTOR *S3I*

Level	Factor S31
2nd	0.0
3rd	0.0674
4th	0.0827
5th (except CONUS)	0.1958
CONUS Depot	0.2679
Factory	0.2947

TABLE 7-8.
FACTOR *S2I*

Level Transfer	Factor S2I
Supply to 2nd	0.875
Supply to 3rd	0.900
Supply to 4th	0.914
Supply to 5th (except CONUS)	0.965
Supply to CONUS or Factory	1.0

where

FR = number of repairable units failing over system life (see par. 7-4.3.3.*1*)

$S1I$ = repairable "shrinkage" factor (see par. 7-4.3.3.7)

WP = unit repair cost—parts

and WP is given by:

$$(WP) = [X/(S2I)]\,Z \qquad (7\text{-}27)$$

where

X = number of parts used for average repair

$S2I$ = repair parts shrinkage factor

1. The factor $S2I$ is tabulated in Table 7-8.

and

Z = failure weighted parts cost. This average part cost per assembly failure is obtained as the product of parts failure rates and parts costs, divided by the sum of parts failure rates, i.e.,

$$Z = \sum_{i=1}^{n} N_i\lambda_i C_i \Big/ \sum_{i=1}^{n} N_i\lambda_i \qquad (7\text{-}28)$$

where

N_i = population

λ_i = failure rate

C_i = cost of any item-type i

Alternately, when detailed design information is unavailable, a weighted parts cost of $30 to $70 is recommended (for the TRIM model).

7-4.3.3.10 Logistic Processing Costs (4.031)

$$C(4.031) = (CHRI)(FR) + (CHTI)(FTR) \qquad (7\text{-}29)$$

where

CHRI = cost of supply processing action at level I for repairable items

CHTI = cost of supply processing action at level I for throw-away items

FR = number of repairable unit failures (par. 7-4.3.3.7)

FTR = number of nonrepairable units replaced

CHRI includes cost to initiate and fill an order, and replace the shelf item, as in Table 7-9.

CHTI reflects flow of nonrepairable units through the supply chain. It is 50% of 2nd maintenance level CHRI, 50% of 3rd level *CHRI*, plus depot *CHRI*, or $19.50 per item plus the weight factor.

7-4.3.3.11 Logistic Cataloging Annual Costs (4.032)

$$C(4.032) = G(CIA)Y \qquad (7\text{-}30)$$

where

G = number of new line items introduced into the Federal Supply System (par. 7-4.3.3.5)

CIA = cost/line item in the FSS (**$1400—Huntsville**)

Y = system life (10 yr)

7-4.3.3.12 Transportation Costs (4.05)

$$C(4.05) = (CTT)(FTR) + 2(CTRI)(FR) \qquad (7\text{-}31)$$

TABLE 7-9.
COST OF SUPPLY PROCESSING ACTION

Level	CHRI	Remarks
2	\$ 1.00	
3	\$ 10.00	1.14 man-hours at \$7.91/mh, including overhead = \$9, plus \$1 from 2nd level
4	\$ 14.50	2 level plus 50% of third level + 4 level (4 level taken = 3 level = \$9.00)
5	\$ 14.00/item + weight factor	(see Fig. 7-3 for weight factor)

where

CTT = one way transportation cost for throw-away units (direct estimate)

$CTRI$ = one way transportation cost for repairable units to level I (direct estimate)

FTR = number of nonrepairable unit failures (see par. 7-4.3.3.8)

FR = number of repairable unit failures (see par. 7-4.3.3.7)

7-4.3.3.13 Building Maintenance Costs (4.06)

$$C(4.06) = (SS)(CSA)Y \qquad (7\text{-}32)$$

where

SS = cubic feet required for maintenance buildings at all levels (direct estimate) (see par. 7-4.3.3.3)

CSA = annual building upkeep cost per cubic foot (10% of construction cost; \$0.10/cubic foot for cement block)

Y = system life (10 yr)

7-4.3.4 costs Per Unit of Use

The preceding CERs are total costs for any element. Where maintenance level related CERs have been given, they must be calculated for each level and the results added. Then, an analysis of lifetime, number of rounds fired, distance traveled, or other units of use is necessary—followed by a division of total costs by total units of use to obtain costs per unit. Otherwise, complex adjustments must be made to each fixed cost CER (e.g., cost of buildings) to convert them to costs per unit, and every other CER must be converted to a variable cost CER in terms of the particular units of interest.

7-4.4 COST STATES OF A SYSTEM; TIME PROFILES

Par. 7-4.3 has presented cost estimating relationships based on total operating hours, through use of average annual hours multiplied by system life. In many analyses it is necessary to consider a more detailed treatment of time as a variable. A major system or equipment may have an operating profile spanning several years—including overhaul, training, and deployment cycles. In comparing such a system with another intended for similar use, but having a different time profile, averaging operating schedules may conceal important cost effects relevant to design or selection.

In other cases we may be interested in system cost as a function of its maintenance "state" i.e., ready, in periodic maintenance, or failed. Par. 7-4.4.1 discusses time phasing within and between years for cost analysis; par. 7-4.4.2 discusses discounting; par. 7-4.4.3 discusses escalation; and par. 7-4.4.4 discusses combined analytical treatment.

7-4.4.1 Time Phasing

To develop time-phased costs, estimating relationships or discrete costs are needed for each state and operating profile element. Two common methods for handling this are statistical aggregation and piecewise costing. In statistical aggregation, typified by the CERs already presented, the expected time in each state is calculated from operating profile, $MTBF$ and $MTTR$ data. Par. 7-4.3.3.6, for example, used a CER for repair cost based on the item failure rate. In effect,

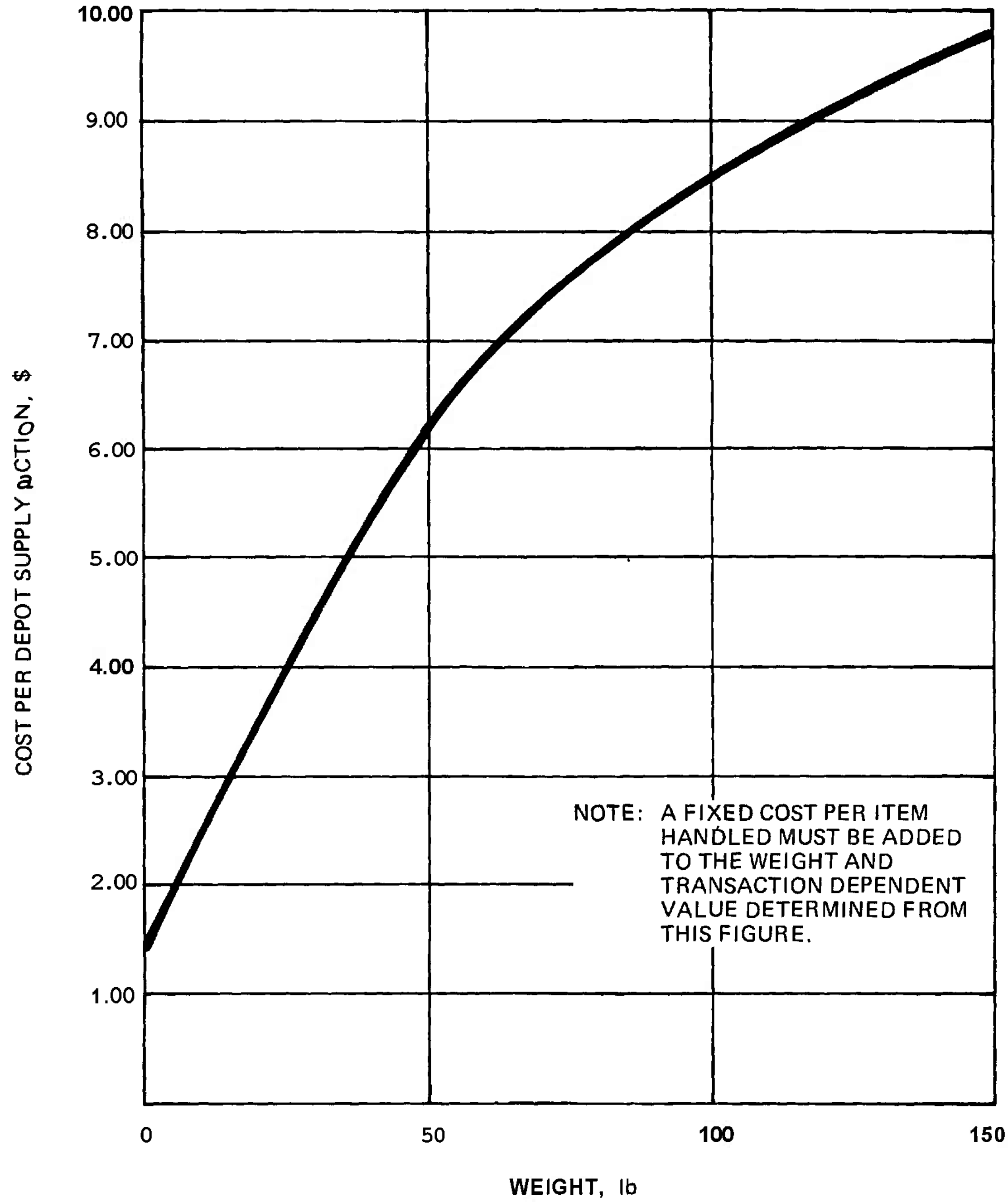


Figure 7-3. Depot Supply Action Cost vs Weight

TABLE 7-10.
PIECEWISE COSTING

Time	State	Number of Failures	Cost of Failures @ $100 Each
1	Training	50	$ 5,000
2	Deployment	20	2,000
3	Deployment	20	2,000
4	Deployment	20	2,000
5	Deployment	20	2,000
6	Training	50	5,000
7	Deployment	20	2,000
8	Deployment	20	2,000
9	Deployment	20	2,000
10	Deployment	20	2,000
11	Rotation	5	500
12	Training	50	5,000
Total		315	$ 31,500

the expected number of failures (number of times in the failed state) was used. Similarly, operating hours are calculated by summing the operating hours in each mode—e.g., training, deployment, and rotation.

Piecewise costing would develop separate **CER's** for each mode, and apply the appropriate ones to each unit of time (quarter or month) over the life of the system.

Table 7-10 presents simiplified data for a system which is intensely stressed during training, moderately stressed during deployment, and unstressed during rotation. The table itself shows a discrete, explicit piecewise costing. Statistical costing would assume 3 training periods per year, 8 deployment periods per year, one rotation per year, and apply an average formula, such as:

Failure cost per year = [(No. of Training Periods)(Training Failure/Period) + (No. of Deployment Periods)(Deployment Failure/Period) + (No. of Rotations)(Rotation Failure/Period)] Cost/Failure or $(3 \times 50 + 8 \times 20 + 1 \times 5)(100) = \$31{,}500$.

But the time pattern would be lost for such a method; pars. 7-4.4.2 and 7-4.4.3 will discuss the importance of presenting this time pattern.

In calculating costs, system life is important. The time period chosen should reflect probable system life. In comparing systems with dissimilar lives, salvage value for each must be put in at the end of the time frame being used and replacement costs and salvage value must be included for elements whose life is shorter than system life.

7-4.4.2 Discounting

Relations among alternatives and calculation of economic impact involve flows of funds over time. When we commit to a particular system, maintenance doctrine, level of modularization, or test equipment, we commit operating costs over time. The flow of these costs over time often varies between systems. A dollar today is worth more than a dollar in ten years. The commitment of operating funds means giving up alternative uses of these funds. Discounting is used to add costs in different years. The mechanism is simple. **A** dollar in the bank today at **6%** interest is worth **$1.79** at the end of ten years (with interest). Alternatively, the present value of **$1.79** received ten years from now, at a discount rate of **6%,** is **$1.** In general, the present value PV of $\$K$ received N years from now, at a discount rate Eis:

$$(PV) = K/(1 + E)^N \tag{7-33}$$

The ***present value*** of a future **$1,** received ten years from now, at a **6%** discount rate is **(1/1.79)** or **0.558.** This is called the present value factor. Tables of present value factors are readily available (Ref. **7**) for different discount rates and time periods. In calculating discounted present value of a flow of funds, we multiply the cost in each year by the present value (present worth) factor for that year at the discount rate being used and sum the total discounted operating costs. To that total we add the acquisition costs—which are not discounted in the first year—to obtain the discounted present value of all costs. Repeating this process for alternatives being considered, we may then compare the different present values to arrive at a decision. Note that this process requires accurate time phasing of costs in each year rather than lifetime costs averaged annually. Ref. **16** is a good introductory reference to discounted cash flow analysis and many other engineering-economic, decision-making techniques.

What should the discount rate be? Cost analysts do not agree on this question; but **6%** to **10%** rates are commonly used for military trade-offs. In many cases it makes no difference to the outcome. In other cases, the decision among alternatives may change with a particular rate. **A** sensitivity analysis, assuming alternate rates, will quickly reveal when this is so, and the discount rate that equalizes the present value of alternatives may be found by trial and error. The decision maker can then decide whether the appropriate discount rate is higher or lower than this "break-even" rate. Since the rate is a measure of the opportunities foregone, one persuasive argument is that it should be the average long-term rate of return of the civilian sector. In general, a rate of **6%** is not unreasonable. Table **7-11** shows a typical calculation of discounted present value. The system shown has a discounted present value (at **6%)** of **$1060,** compared with an undiscounted (discounted at 0%) present value of $1290.

7-4.4.3 Escalation

Just as discounting captures the time-effects of expenditures, escalation captures the change in price levels over time. To estimate costs in each future year based on current price is the equivalent of assuming no escalation. This "constant dollar" assumption is unsatisfactory for budgetary analysis (when we need to know what funds must be requested in the future) and for many kinds of trade-offs. If two alternative systems are being compared, one with automated test equipment and the other with manual diagnosis, we must make assumptions about the rise in the price of men (salaries and fringe benefits) over the life of the system in order to fully credit the automated alternative with all its savings, including those obtained by purchasing the automated equipment at today's prices. The method of calculation is to separate costs that will escalate from costs that will not escalate in each year, apply the correct escalation factors, total the escalated costs in each year, and then apply the discount factors. Table **7-12** shows a pro-forma cost calculation form which includes all effects. The formula for the escalation factor EF in year N at escalation rate E is:

$$(EF) = (1 + E)^N \tag{7-34}$$

7-4.4.4 CER's with Escalating and Discounting

It is often convenient to include escalation and discounting in CER's particularly when they are to be used in computerized cost models. If costs are constant in each year, a CER such as Eq. **7-35** may be modified to that of Eq. **7-36**

$$C = f(A, B) \tag{7-35}$$

$$C = f(A, B)[(1 + E)/(1 + D)]^Y \tag{7-36}$$

TABLE 7-11.
DISCOUNTED PRESENT VALUE CALCULATION

Year	Undiscounted cost ($000)	Discount Factor	Discounted Cost
0 (Investment	$ 500	1.000	$ 500
1	70	0.943	66
2	60	0.890	53
3	90	0.840	76
4	60	0.792	48
5	65	0.747	49
6	110	0.705	78
7	65	0.665	43
8	70	0.627	44
9	80	0.592	47
10	100	0.558	56
Total	$ 1,270		$ 1,060

Note: 6% Discounting, 10 Year Life, Zero Salvage Value

where

A, B are independent variables

E = escalation rate

D = discount rate

Y = year (1, 2,. . .,K)

If costs vary from year to year, the expression $f(A,B)$ must be replaced with $f(A, B, Y)$ where the structure would express the variation in undiscounted, unescalated costs with time (as in the familiar U-shaped failure curve reflecting initial "bum-in", subsequent stability, and final service life effects on failure rates). In such an analytic cost model, E may vary among cost elements, expressing inflation in each cost element, but D should not, being a system, service, or economy-wide parameter.

TABLE 7-12.
COST CALCULATION FORM

INVESTMENT

Element	2.013	3.015	4	. . . . .	Total
cost	$	$	$	$	$

OPERATIONS AND SUPPORT

Year	Raw Cost Element 4.015	Escalation Factor At ____%	Escalated cost	Element	Escalation Factor At ____%	Escalated Cost
1						
2						
3						
4						
.						
.						
20						
Total						

Year	Element	Escalation Factor At ____%	Escalated Cost	Total Escalated Cost	Discount Factor At ____%	Discounted Cost
1						
2						
3						
4						
.						
.						
20						
Total						

7-5 COST ANALYSIS UTILITY

In planning system development we need to determine present and future financial implications of system decisions for budgeting purposes. Budgets also act in reverse—as a constraint on design. The principal M-related system variable having overall design input is availability. Par. 7-5.1 presents an overview of the problem of costing system availability, its effect on M-related design, and utility calculations relating budget constraints to performance and availability.

Par. 7-5.2 discusses the utility of analysis itself; given a series of system design alternatives, how far should each be carried in early feasibility studies before settling on a smaller number (perhaps one) of alternatives for detailed refinement?

Par. 7-5.3 discusses the problem of maintenance planning under varying conditions of manpower or skill utilization.

7-5.1 COSTING SYSTEM AVAILABILITY

In general, a system should be designed for maximum performance at specified cost, or for specified performance at minimum cost. In some special cases, where the budgetary outcome is known within a range of uncertainty for any feasible system, design may be directed to maximize a cost-effectiveness measure (utility per dollar), subject to side constraints on some performance measures. In the first and third cases, availability can be traded off for other performance characteristics if we know the "exchange rate" between a unit of availability and a unit of, say, speed. This requires the establishment of (military) utility functions which express such ratios in terms of an overall performance measure U. Suppose this is the case. Then we can calculate the marginal exchange rate of contribution to U per dollar spent on availability versus contribution to U per dollar spent on speed. Provided

$$\Delta U_A/\Delta C_A > \Delta U_V/\Delta C_V \qquad (7\text{-}37)$$

the extra benefit (marginal utility) per dollar spent on availability is greater than the extra benefit (marginal utility) spent per dollar on speed, we would design for an increment of availability rather than speed or, if we had designed up to a budget limit, we would reduce speed and increase availability until the ratios were equal. More generally, with several variables we would perturb all performance-related variables until their exchange ratios were equal at the specified system budget. At that point we would not gain any total performance by decreasing the amount spent on any performance factor and increasing the amount spent on another. Ref. 17 contains a detailed discussion of this subject with military examples, while Ref. 18 presents the economic theory in detail.

In the second case, (meeting fixed requirements with the least life-cycle cost), availability, as one of the performance parameters, will be specified. The objective is to reach the target availability at minimum cost. We would first find feasible (meeting all requirements) systems, and then search for least-cost alternatives between and within such feasible systems, while continuing to meet target parameter values.

Developing a system availability cost function requires the solution of three problems: finding the least-cost design and its cost for each reliability value over the range of interest and feasibility, finding the least-cost design and its cost for each maintainability value over the range of interest and feasibility, and finding the least-cost combination of reliability and maintainability to achieve each level of system availability over the range of interest and feasibility. Analytically, we can see that this is so since:

$$A_i = MTBF/(MTBF + MTTR) = 1/(1 + K) \qquad (7\text{-}38)$$

where

A_i = availability
MTBF = mean-time-between-failures
MTTR = mean-time-to-repair
K = maintenance time ratio ($MTTR/MTBF$)

For a specified availability, MTTR and $MTBF$ must be in a constant ratio. Increasing one requires increasing the other to maintain a required availability. If we assume that each level of MTTR has a particular cost (more about this in a moment) and similarly for $MTBF$, we have:

$$C_A = C_M + C_R = f(MTTR) + g(MTBF)$$

where

C_A = cost of availability
C_M = cost of maintainability
C_R = cost of reliability

and we wish to minimize C_A (find the least-cost system achieving specified availability) subject to

$$K = MTTR/MTBF = (1 - A)/A \qquad (7\text{-}39)$$

For a particular availability we wish to know the cost of the least-cost system. But C_M and C_R are themselves variable for a given MTTR and $MTBF$. There are many ways of achieving a particular *MTBF*, including redundancy. Each MTBF and similarly each ***MTTR*** has a least-cost method of achievement (e.g., changing number of technicians versus degree of automatic test). Par. 7-6 discusses these trade-offs to achieve specified maintainability, as well as the reliability-maintainability trade-off to achieve specified availability. For now, we postpone the discussion of particular hardware trade-offs.

Let us suppose we wish to cost system availability in early concept design. Then we need two sets of cost functions, C_R and C_M. These CER's must express system costs in terms of MTBF and ***MTTR,*** respectively. We therefore need a "prescriptive" cost model. The maintainability cost model would require **CERs** that reflect the cost of the least-cost system for **any** MTTR. Such expressions may be derived from historical data, statistically aggregated, or from preparation of a detailed model specifying and costing the least-cost maintenance doctrine, amount of automated test equipment, number of men at each level, and related variables for any level of M on a system-by-system basis. Another approach is to use state-of-the-art, military requirements and historical cost data on particular system-types and their availability to establish availability requirements and targets. Reliability and maintainability are next "allocated" at system, subsystem, and component levels based on state-of-the-art, cost, and next-level *R* and M requirements. Trade-offs are next used to make lower-level decisions, and to reallocate based on excursions from baseline designs. For example, if we allocate availability to R and M portions, and discover that the cost of increased *R* compared with that allocated is less than that of increased M, we would design in the direction of more R and less M, achieving a lower cost design for the specified availability. This "marginal trade-off" method (similar in concept to the marginal utility trade-off discussed at the beginning of par. 7-5.1) is discussed in par. 7-6.

Costing system availability is particularly important under constrained budgets. If the budget constraints are sufficiently low, we may not be able to achieve a desired level of availability. The lower availability can only be caused by reduced reliability or maintainability. Otherwise, we must sacrifice some performance parameter target(s) to "pay for" the required availability. Reduced reliability may be due to less costly components or reduced redundancy, for example. Reduced maintainability may be due to reduced test equipment and automatic checking, reduced maintenance manpower, or reduced sparing and other logistic support, for example. In sacrificing performance parameters we are likely to be tempted to give up an increment of performance having the highest marginal cost per unit of utility. This often results in small reductions of the most essential military factor.

7-5.2 SYSTEM DEVELOPMENT—ALTERNATIVES

During system development we often consider alternatives—from an M-related point of view such alternatives might include similar hardware except for the amount of built-in test equipment, the same hardware inspected at different intervals, the same hardware serviced by different maintenance systems (number and skill level mix of technicians, different levels for checkout, discard, or repair), or different amounts of redundancy. How long should such alternative systems be carried? How can we choose among them?

7-5.2.1 The Value of Information

As a general rule, when alternative systems are considered they should be developed in parallel provided such development is necessary and advantageous as follows: (a) At frequent intervals, the development process should be reviewed. (These intervals may be periodic, or a review may be triggered by an event such as the completion of a trade-off or analytic study.) (b) At these decision points, the first question to be asked is, "Do we know enough now to choose among alternatives?" If so, the number of alternatives is reduced (usually to one), using criteria of cost, effectiveness, and performance and mission requirements. If a choice cannot be made on these grounds, the second question to be asked is, "Do the benefits of continuing the development of alternatives outweigh the cost of so doing?" We compare the value of additional information with its cost. (c) If further work is not justified on these grounds (value need not be expressed in dollars, although it is helpful to do so), one or a reduced number of alternatives should be chosen on *a priori*, judgmental or other grounds and the work continued. Sometimes one alternative of high risk but extreme benefit is carried along with the baseline alternative. Often a reduction in the number of alternatives may be quickly achieved through trade-offs and analytic studies. Par. **7-5.2.2** introduces selection concepts and par. 7-6 deals extensively with the economics of trade-off decisions.

7-5.2.2 Selection of Maintenance Interval and Manpower Levels

One method for determination of maintenance interval and manpower levels has been suggested in Ref. 1, Chapter **2.** An item is assumed to be good, deteriorated, or failed at the start of any small increment of time between inspections. (Probabilities are specified for going between any of these states during the calculation increment, chosen so that the probability of more than one transition between states is negligible, but the increment is long enough to permit repair of the failed item.) The resulting Markov process leads to the calculation of the probability of needing a preventive repair P_i at the end of some interval ibetween inspections, (the interval i is much longer than the calculation increment) and to the calculation of E_i, the expected number of failures, at the end of that interval i, between inspections. A simple economic model is then specified. The costs of inspection, and of repair during an inspection, are specified in terms of the average times for these actions, the number of men available to perform them, and the cost of downtime. The cost of corrective repair is also calculated. Corrective repair costs can be large, particularly if we calculate the hourly economic loss during downtime as the total system lifetime cost divided by lifetime expected operating hour. Ref. 1 next calculates the total cost as a function of these individual costs, the probability of a preventive repair, number of corrective repairs needed, number of subsystems, and the set-up and manpower costs. The result is a cost function (presented in parts in Ref. 1):

$$C_T = N_A[NC_D(W_{1i} + W_{2i}P_i + MD_iE_i)/M + (SU)] + MC_M \tag{7-40}$$

where C_T the total cost, is to be minimized with respect to N the number of maintenance men M, and W the interval between inspections N. P_i and E_i are the probability of a preventive repair during inspection and the expected number of corrective repairs between inspections, respectively, and are functions of N, the interval between inspections. C_D is the hourly cost of system downtime and is related to the system cost and value. D_i, W_{1i}, and W_{2i} are the average elapsed time for emergency diagnosis and repair, and the average manhours for preventive repair, respectively. SU is set-up cost for preventive inspection, N_A the number of subsystems, and C_M the cost per maintenance man. Some of these elements may have to be calculated piecewise and summed. For example, if the subsystems are not all identical, a series of P_i, E_i, W's, and N_A must be considered separately. Again, in a realistic case, a series of manning and skill combinations and their impact on D, W, M, and C_M will have to be calculated separately. As a result, some complex and sophisticated maintainability/economic models have arisen, stepping through maintenance alternatives and calculating their consequences and costs. Par. **7-6.2** describes some examples of these models and others.

Some of the results of the model include:

a. As the time between scheduled inspections increases, the number of preventive repairs per year decreases, and the expected number of failures per year increases.

b. As the number of maintenance men increases, the cost per year of unit preventive repair (excluding set-up and parts costs) decreases as does the marginal cost of inspection; the marginal cost of inspection is much less than that of preventive repair. With enough maintenance men, the sum of the annual marginal costs of failures, preventive repairs, and inspections increases; cutting the number of maintenance men causes costs to have a minimum at a finite maintenance interval.

c. After all costs are included, the annual maintenance cost is a double U-shaped curve of maintenance men. The result is shown in Table **7-13** (adapted from Ref. 1).

Many sophisticated techniques for scheduling maintenance actions have been developed. One survey of that literature may be found in Ref. 19. Ref. 20 discusses stochastic and deterministic maintenance policies including economic criteria (cost per unit of time) and presents some illustrations, including product improvement.

7-6 ECONOMICS OF TRADE-OFF DECISIONS; APPLIED COST MODELS

This paragraph describes the economic approach to trade-off analysis. In conducting trade-offs (rational comparison of alternatives and selection among them), the economic, physical, performance, supply, logistic, development, and maintenance characteristics of equipment in an operational environment must be described in a way permitting comparison and choice. This requires description and quantification. (Qualitative comparison is often necessary for some factors; here we neglect such factors, although they must enter consideration when they are of major significance.) The most common approach to quantification is to build a "model"—a representation of reality. Some models have few equations and can be solved on inspection. More commonly, large numbers of equations and some mathematical decision rule are part of current maintainability models.

Par. **7-6.1** introduces the problem of comparison of alternatives and substitution to "set the stage" for an economist's view of the trade-off process, and includes a geometric and mathematical formulation for the most frequently encountered M-related trade-off decisions. Par. **7-6.2** discusses some trade-off models which illustrate model-types in use today, including programming models which permit "best" (optimal) solutions to be found when the underlying processes can be accurately represented in mathematical form.

In conducting trade-offs, we compare one or more "inputs" (e.g., cost incurred by equipment, replacement policy, maintenance level, manpower and skill levels, automatic test equipment, training and redundancy) to one or more resulting "outputs" (reliability, maintainability, availability, system effectiveness). By varying the quantity of an input, we can, usually, vary some desired output. When we vary more than one input simultaneously, the results may not be the sum of the effects of each input varied separately. Many different combinations of inputs may also give the same output. The remainder of this paragraph will show how to handle such effects quantitatively.

In conducting trade-offs we must express mission, deployment, logistic, and maintenance capability quantitatively. This involves reduction to averages (e.g., man-hours for a particular repair), profiles (percent of time in a particular year spent deployed, training, rotating; percent of time in each of these "modes" spent in a particular speed range), and decision tables (level for checkout or repair of a module). Par. 7-6.2 will describe these simplifications for some current trade-off models.

7-6.1 THE GEOMETRY OF TRADE-OFFS

Let us examine the effect on availability of different amounts of reliability and maintainability, and determine the least-cost combination of reliability R (obtained through equipment design) and maintainability M (obtained through test equipment, repair system, and sparing doctrine) which will produce a specified (mission-required) availability A. Since $A_i = MTBF/(MTBF + MTTR)$, a three-dimensional graph or "response surface" will show the relationship among R, M, and A. Note that $MTBF$ is a measure of R, but MTTR is a measure of $1/M$. We can show such a response surface, as in Fig. **7-4.** $A1$ through **A4** are increasing levels of availability; any particular availability, (say **A1**) may be obtained from many different combinations of R and M. Fig. **7-5** shows a two-dimensional projection of such a surface; for the rest of this paragraph we will use such projections as representatives of the response surface. Such projections may be thought of as contour lines on a map, just as we represent topographical features on a two-dimensional map. They are called "iso-availability" (constant availability) contours, or "isoquants" for short. The isoquants shown represent availability, increasing from $A1$ to **A4.** They say nothing about cost. Along any isoquant, many different costs are represented. The surface and its isoquants are convex to the original, showing decreasing marginal (incremental) returns (benefits). As reliability and maintainability are successively increased by a fixed increment, availability is increased by a decreasing increment. We can quickly see this from the formula $A_i = MTBF/(MTBF + MTTR)$.

TABLE 7-13.
ANNUAL MAINTENANCE COST OF A HYPOTHETICAL SYSTEM ($000)

		NUMBE OF MAINTENANCE MEN				
		1	2	4	8	12
Time	1	572	496	444	472	530
Between	2	534	436	406	452	520
Scheduled	4	504	438	434	492	566
Maintenance	8	492	464	482	550	626
(Months)	12	494	482	506	578	656

(Try assuming values for either *MTBF* or *MTTR*, holding the other constant. More powerfully, calculate the partial derivative of *A* with respect to either.) The individual isoquants approach asymptotes to either axis because larger and larger amounts of *R* or *M* are required to produce a fixed *A* as *M* or *R* becomes small.

R and *M* are "competitive substitutes" for each other, and the rate of such competitive substitution diminishes as we move along any isoquant. Again, reference to the formula shows the substitution effect. For a target availability *A* and the resultant constant *K*, if we take partial derivatives in Eq. *7-39*, remembering that *MTTR* is proportional to $1/M$, we obtain:

$$\frac{\partial M}{\partial R} = -\frac{K}{R^2} \qquad (7\text{-}41)$$

where $K1 = A/(1 - A)$. Thus, at any availability, as *R* increases *M* decreases proportionately with R^2. At very large *R*, the slope of an isoquant is almost zero, while at very small *R*, the slope is almost infinite (the isoquant becomes parallel to either axis as one goes out far enough). Note finally that the response surface approaches a point, *A5*. At some level of availability we have approached the "state-of-the-art" limit; no higher value of *R* or (perhaps) *M* is possible and a single combination of *R* and *M* is the only way to achieve such a high *A*. The isoquant has become a single point.

Fig. *7-6* shows a series of "budget" or "isocost" curves. Any curve represents combinations of *R* and *M* that can be purchased for a particular budget. Budgets increase from *B0* to *B4*. The curves intersect the axes at a point where all of a particular budget is spent on *R* or on *M* with nothing spent on the other. The isocost curves say nothing about availability; many different values of availability will likely occur along a particular isocost curve. The isocost curves are somewhat concave to the origin, showing increasing marginal costs. The last increment of *R* or *M* is usually much more costly than earlier increments. Thus, as we reduce the amount of *R* slightly, we can "buy" substantial *M*, and vice versa. This effect, diminishing returns to scale, is a well known economic phenomenon at high levels of output for many kinds of process. In this case, the highest increments of reliability or maintainability require special, costly techniques (gold-plating or redundancy, for example).

Note that in Fig. *7-6* any point (say, *C*) on an isocost curve represents the quantity (*qM*, *qR*) of *M* and *R* that can be purchased for the fixed budget (*B1* in this case). The unit cost, or price of *M* when all resources are spent on *M* is thus $B1/qM1$ (where *qM1* is the point of intersection of *B1* with the y-axis; the price of *R* is $B1/qR1$.)

Fig. *7-7* combines the isoquants of Fig. **7-5** with the isocost curves of Fig. *7-6*. Let us plot the lowest *B*-curve just tangent to each A-curve and denote the pair by the same number (*A1*, *B1* or *A2*, *B2* or *A3*, *B3*, etc.). The points of tangency, *P1* through *P4*, are the least cost combinations of *R* and *M* that will achieve each level of availability. Consider (*A1*, *B1*). If *B1* were any closer to the origin (any lower budget), it would not be tangent to *A1*. We could not buy availability *A1* for such a lower budget. Conversely, if *B1* were further from the origin than shown, it would overlap *A1* in several places; we would, however, be spending more than necessary to achieve availability *A1*. Thus, *B1* is the least cost to achieve *A1*; the tangent point *P1* shows the combination *M1* and *P1* of *M* and *R* that will achieve that availability *A1*.

P2, *P3*, and *P4* are the optimal combination points for *R* and *M* to achieve availability *A2*, *A3*, and *A4* at least cost. The curve connecting *P1* through *P4*, curve *QQ'*, is the "expansion path" and shows the increases in *R* and *M* that must be obtained to increase *A* at least cost in every case. This is the economic mechanism behind the trade-off between *R* and *M* to achieve *A*.

Note that at any tangent point (and only at such a point) the slope of both curves must be equal. Thus, for a given availability curve, at the optimum the slope $\Delta R/\Delta M$ must equal $\Delta C_M/\Delta C_R$, the slope of the budget curve. (The marginal rate of substitution of *R* for *M* must equal the inverse cost ratio.)

How should we search for an optimum (tangent) point? One way is to exploit the behavior of the availability isoquant. Fig. **7-8** reviews the cost behavior along such an isoquant. Beginning at a high *M*, low *R* combination yielding target availability *A1*, we move toward the low *M*, high *R* end of the same isoquant. As we do so we pass through isocost curves representing decreasing cost from *C5* through *C1* (the optimum) and then through curves of increasing cost *C2* through *C5*. Suppose we select a random point (design) on *A1*. By examining the cost of that design, and the cost of a more reliable, less maintainable design, we can tell whether we are moving away from the optimum point or toward it. If the costs increase, we are moving away from the optimum, and we would reverse our strategy by reducing *R* and increasing *M* to decrease costs. If, at some point, this strategy caused costs to increase, we will have passed through the optimum and must backtrack. On the other hand, if the trial design produced a lower cost for more *R* and less *M*, we would be above and to the left of the optimum and would continue to increase

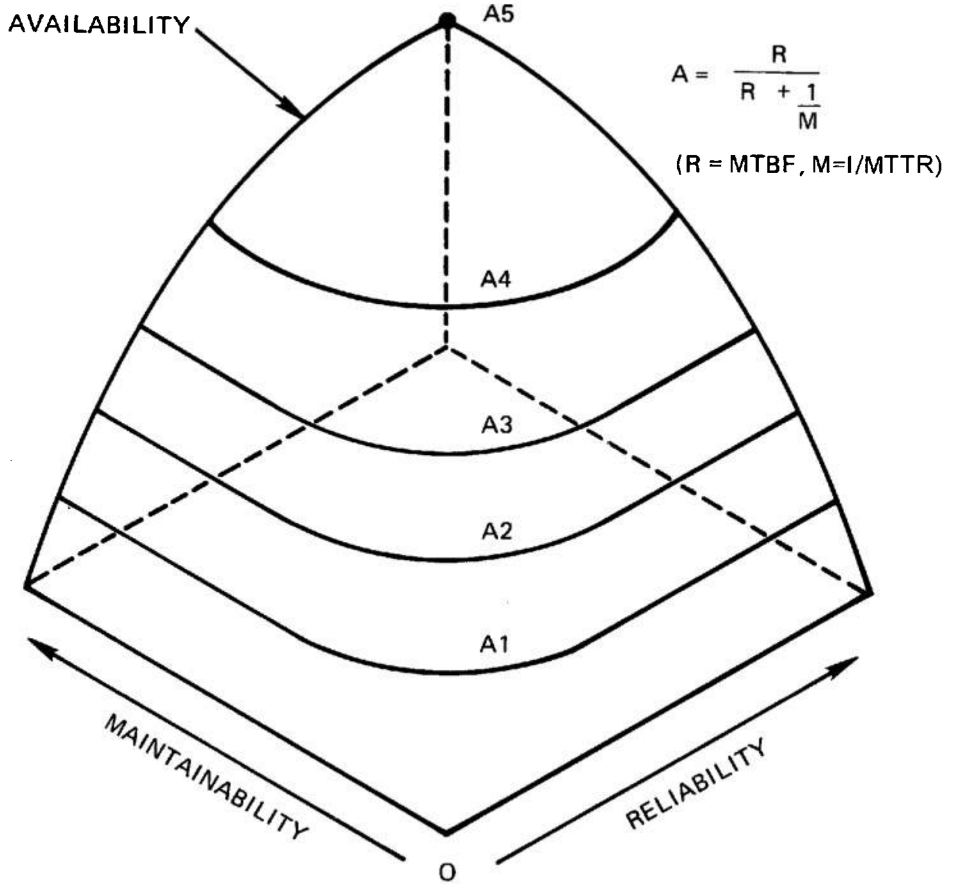


Figure 7-4. Hypothetical Availability Surface

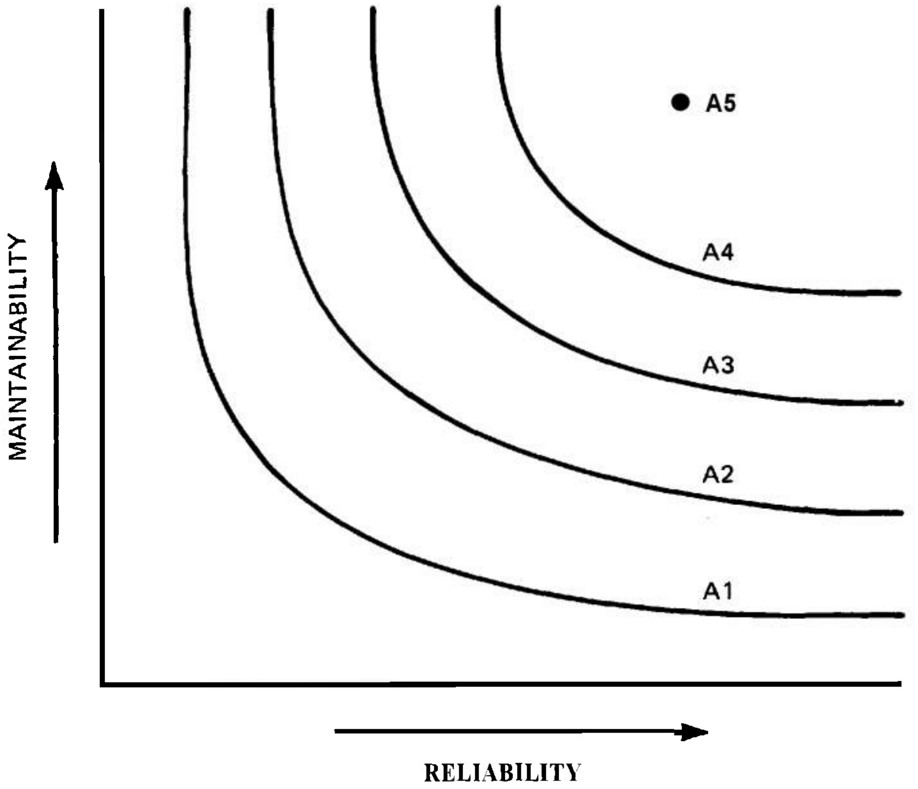


Figure 7-5. Two-dimensional Projection of Availability Surface

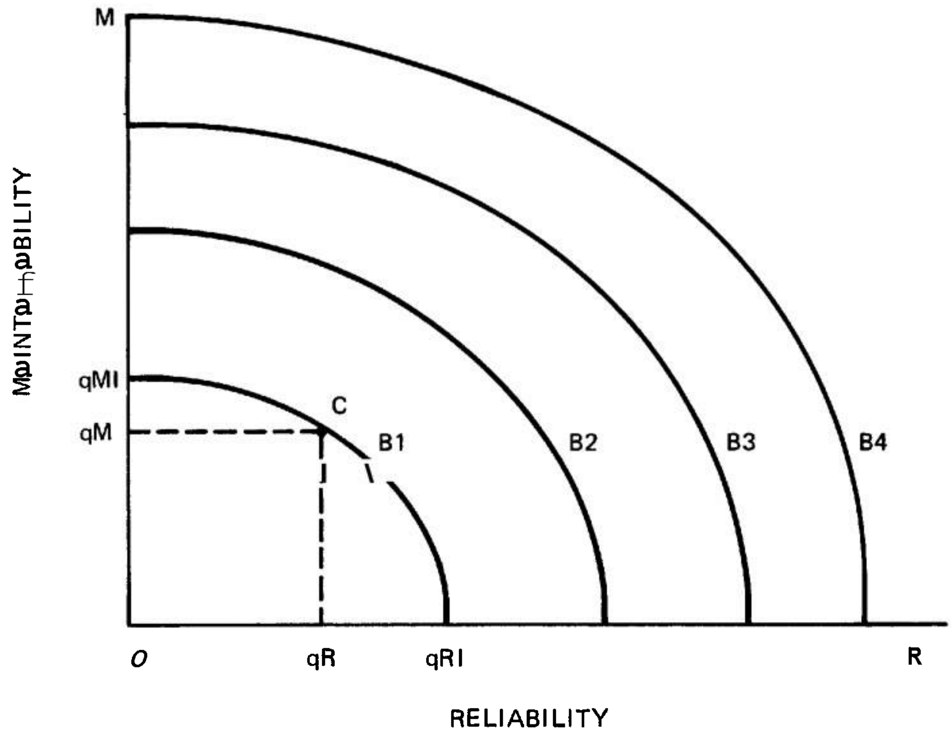


Figure 7-6. Hypothetical Budget Curves

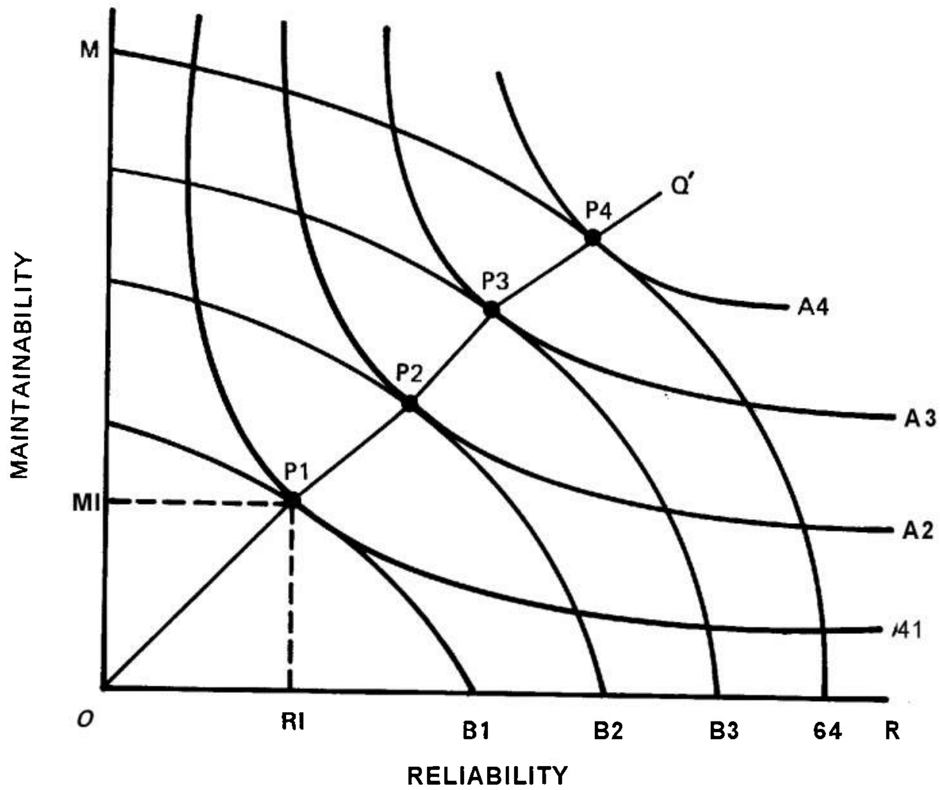


Figure 7-7. Optimal Combinations of Mand *R*

R and reduce M Fig. 7-9 shows this effect, and the familiar U-shaped cost curve that results.

This discussion centers at achieving a fixed availability at minimum cost. The problem of maximizing availability at a fixed cost is solved in similar fashion. A fixed isocost curve is chosen, and the availability curve just tangent is found. Any higher availability curve will not intersect the isocost curve at any point; the higher availability cannot be bought for the specified budget. Any lower availability curve than the tangent curve would represent less than maximum availability for the specified budget. Note, then, that the solution to the problem of maximizing benefits (effectiveness) for a fixed cost, and the solution to the problem of minimizing cost for a fixed benefit are formally identical. In theory, we may start with either and should find the same solution.

In this paragraph we have used availability as the measure of effectiveness, and *R* and *M* as elements. The techniques may be generalized to any measure of effectiveness and to more than two dimensions (measures) of that effectiveness.

For a detailed discussion of the mathematics and further details of the geometry, Ref. 18 is a definitive work in mathematical micro-economics; Ref. 21, particularly Chapter 9, presents a more engineering-oriented treatment of much of the same material.

7-6.2 MATHEMATICAL MODELS FOR TRADE-OFF DECISIONS

In par. 7-6.1 we have seen the basic economic mechanism, presented in continuous form, for development of a trade-off model for comparing elements of cost to produce effectiveness. *R* and M were the cost elements; A was a measure of effectiveness. In practical cases, the illustrative framework can be applied directly. In order to find an optimal point we need a mathematical expression (or a table of values in the range of interest) for the cost of any combination of *R* and M producing, say, a target availability. As we move along the target availability curve (as we examine different combinations of *R* and M producing the desired *A*), we cost each combination, searching for the minimum-cost solution (the tangent point of par. 7-6.1). We use one of a variety of search techniques. These techniques range from searching all possibilities in the (usually narrow) range of interest defined by the military problem, to selection of a single, feasible, combination of *R* and M which achieves the desired A, and then making excursions from that point on the A curve through system design modifications that reduce total cost until no further benefits (at least commensurate with the cost of continued analysis and design) can be found. Here we will examine a number of trade-offs between system elements and costs, and discuss typical techniques for the conduct of these trade-offs. In the trade-offs, variables may represent alternative ways of achieving a particular M or *R* separately, as well as alternative M-R combinations for achieving a given A. This is an important aspect of the general problem. Until now we have assumed some level of M and of *R* for each point on an availability isoquant. But the same value of M may be achieved in many ways—through different repair versus discard policies, piece part versus modular replacement, number and skill level of repairmen versus design considerations, automatic versus manual test equipment, alternative training methods and techniques, and diagnostic versus go/no-go techniques. In each case, the same level of maintainability can result from different combinations of resources. Before we can perform a global *R*-*M* trade-off for a given A, we must determine the least-cost method for achieving any particular M (and *R*). Only in this way can the implied isocost curves be established. The trade-off technique thus contains trade-offs within trade-offs. A restatement of the search algorithms in par. 7-6.1 will make this clear:

a. Pick a reasonable value of *R* and M which meets required A. Call this set of values $RM1$.

b. Find the least-cost system which provides that value of *R* and of M Call this $C1$.

c. Pick a nearby set of values $RM2$, such that $R2$ exceeds $R1$ and $M2$ is less than M1.

d. Find the least-cost system having $RM2$. Call it $C2$.

e. If $C2$ is less than $C1$, continue increasing *R* and decreasing *M*, finding the least-cost system in each case, as long as this cost continues to decrease.

f. Otherwise, decrease *R* and increase *M*, searching in that direction as long as each least-cost system costs less than the previous ones.

g. If the costs as a result of e. or f. start to increase once more, the least-cost point has been passed and the search should be reversed, using a smaller increment for *R* and M.

h. Continue the process until a minimum cost system is found or the cost of the search starts to exceed the decreases in cost obtained.

The remainder of par. 7-6.2 will discuss the individual trade-offs for M in order of increasing complexity within the larger process just described. The write-up is organized by approach rather than trade-off subject, since a particular trade-off may be structured in a variety of ways, and particular model-types may be

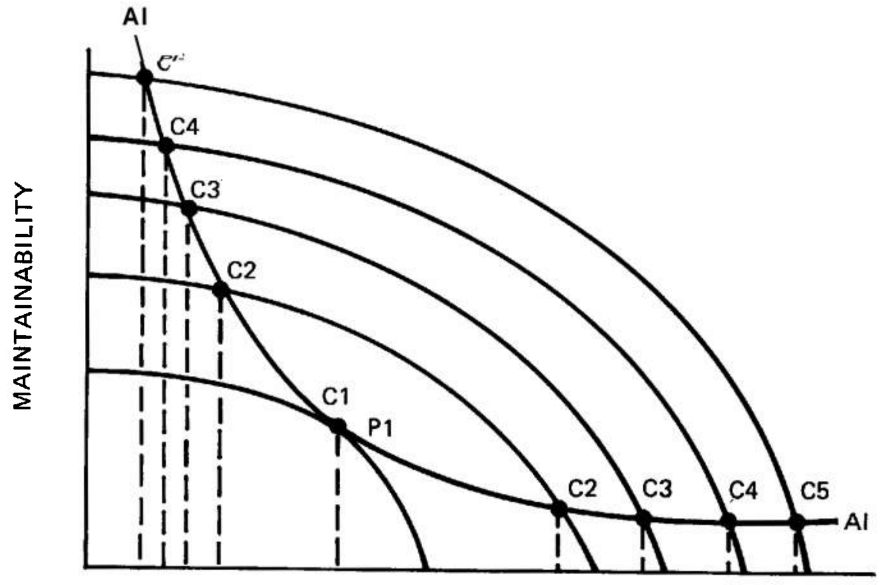


Figure 7-8. Cost Along Availability Isoquant

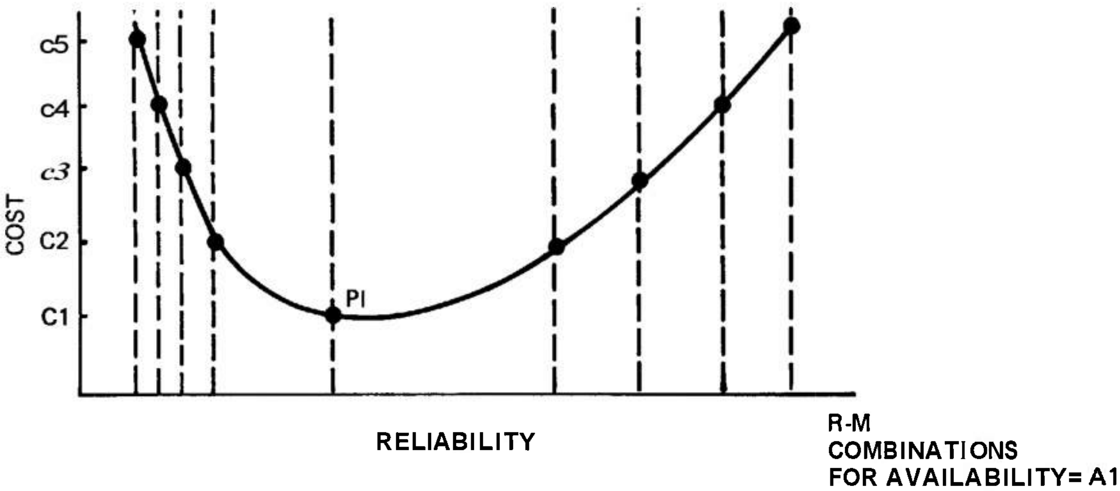


Figure 7-9. Cost Curve for Figure 7-8

used for a variety of applications. Par. **7-6.2.1** treats formulation of the trade-off; par. **7-6.2.2** discusses the "point trade-of€"; par. **7-6.2.3** describes the simple descriptive or "accounting" model. The descriptive model with decision is described in par. **7-6.2.4.** Finally, the prescriptive, or optimum seeking, model is described in par. **7-6.2.5.**

7-6.2.1 Formulating the Trade-off

The first step in formulating a trade-off is to clearly state the ***objectives*** of the trade-off. Are they to find a least-cost schedule of preventive maintenance meeting availability requirements, other things being equal? To find a preventive maintenance schedule maximizing M, at specified cost? To determine repair levels meeting required M at minimum cost? Specific objectives are necessary to appropriately formulate the remaining elements of a trade-off. Carefully specified objectives are necessary to be sure we pose and solve the right problem.

Once objectives have been specified, two or more ***alternatives*** are designed. Without altematives, there is no problem of choice and no trade-off is necessary. A range of alternatives which reflect different approaches ɔ the problem of choice (the objectives) is needed. They must be genuine alternatives, each of which represents a legitimate, feasible, and desirable choice from the decision-maker's point of view. Such "cooked alternatives" as everything, nothing, and the designers' favorite compromise are not viable. Alternatives may be discrete (two or three degrees of automatic test equipment) or semi-continuous (a finely divided continuum of manning and skill levels, or of $MTTR$).

The third element of a trade-off is an explicit statement of ***costs.*** All costs which vary among alternatives must be considered. Costs which would be incurred whether the system being considered were produced or not should be ignored. A detailed discussion of costs, cost elements, and cost structures was presented in pars. **7-2, 7-3,** and **7-4,** and will not be repeated here.

The fourth element of a trade-off is an explicit statement of ***benefits.*** These may be measured in dollars or in units of effectiveness. A benefit structure may consist of a single element (availability, say) or of a complex, tree-like structure similar to that of a cost structure.

The fifth element of a trade-off is ***criteria*** or a criterion function, the explicit, implementable rule for choice among alternatives. A criterion is usually a measure of the combination of costs and benefits applicable to the objectives of the particular trade-off. A criterion function may be simple (minimize costs for fixed effectiveness; maximize effectiveness for specified costs) or complex (maximize the ratio or difference between a weighted sum or product of costs and of benefitsj. Choosing appropriate criteria is fraught with pitfalls: Ref. **17** contains a discussion of appropriate and inappropriate criteria.

The final element of a trade-off is a ***model*** relating the criterion function to costs and benefits. In the remainder of this paragraph we will discuss "point trade-offs" in which the model is a direct accounting comparison of specific alternatives (par. **7-6.2.2),** the simple descriptive or accounting model (par. **7-6.2.3)** into which a variety of different alternatives may be "plugged in" in order to calculate their costs and benefits, the descriptive model with decision (par. **7-6.2.4)** where suboptimizations are made within each alternative (such as level of maintenance) but where each alternative still produces a separate output for analysis and comparison outside the model, and the prescriptive model (par. **7-6.2.5)** in which alternatives are generated and compared according to some predetermined set of rules, and the "best" solution is found.

7-6.2.2 The Point Trade-off

In the point trade-off we compare several explicit alternatives. It is convenient to organize such trade-offs according to a format, similar to that described in par. **7-6.2.1:**

a. ***Objectives.*** The objective of the trade-off are stated in an introductory paragraph. Suppose we are selecting a subsystem design for testing. Suppose further that the system already meets availability requirements; we seek the least-cost subsystem design for testing.

b. ***Alternatives.*** Suppose the alternatives are built-in test equipment which normally comes with the prime equipment, two types of semi-automatic fault detection equipment, and one type of automatic fault detection and isolation equipment. Each alternative would be described in terms of its basic design and principal features. A requirements matrix would be presented, showing whether each alternative met technical requirements, including (for example) $MTBF$, $MTTR$, size, weight, technical risk, and any other non-cost criteria. Alternatives which failed to meet all requirements would have to be redesigned or eliminated. The remaining candidates would be admitted to the next step of the trade-off.

c. ***Costs.*** The cost section should first define each element of development, investment, operating, and support costs applicable to the trade-off. After definition of each element the method for estimating the cost of that element for each alternative in the trade-off should be specified. Any cost estimating relationships, data or references should be presented at this point. Next, the costs should be calculated for each element and each design alternative. Finally, a series of convenient forms, one set per alternative, should present costs by element and year.

d. ***Benefits.*** Benefits (effectiveness) should be calculated and presented for each alternative. One common case, least cost for specified effectiveness, requires only a statement that each alternative meets the effectiveness level specified at this point. (If any alternative exceeds the required effectiveness in such a case, it should be redesigned until it just meets required effectiveness, if costs can be reduced by so doing.)

e. ***Criteria.*** The criteria for combining costs and benefits and selection of a "winner" are specified next. The specifications may range from complex weighting to a simple statement of least life-cycle cost for required effectiveness.

f. ***Model.*** Costs and benefits are combined as specified by the criteria. In a least-cost fixed-effectiveness case, for example, the individual costs would be combined for the total procurement size and delivery schedule, where applicable. The resultant system costs by year would next be discounted, and the total of undiscounted initial costs (with investment discounted for a multiyear procurement if appropriate) and discounted operating and support costs would be presented for each alternative.

g. ***Selection, Sensitivity Analysis and Discussion.*** The final section of the trade-off would use the model results to select the preferred alternative using the prespecified criteria. Sensitivity analyses should be presented, varying costs which have significant impact on the outcome, particularly when estimating uncertainty exists. A discussion should follow, particularly when sensitivity analysis makes the trade-off outcome inconclusive.

Table 7-12 shows one sketch of a cost calculation form. Individual forms for point trade-offs, however, must be tailored to the environment, total procurement, delivery schedule, and discounting and escalation rules applicable to the system being analyzed.

7-6.2.3 The Simple Descriptive (Accounting) Trade-off Model

The accounting trade-off model is a calculation convenience for costing one or more alternative systems. The model makes no decisions; a system must be completely specified in input parameters. The model applies averages, planning factors, and cost estimating relationships to the input data to produce costs directly. Most accounting models are suitable for a "quick look" at system costs, but are unsuitable for detailed trade-off analysis since they resort to many oversimplifications.

One series of accounting models is presented in Ref. 22. A combat vehicle model for tanks uses simple **CERs** and adds them. One hundred equations including accounting identities (sums of lower level elements) are presented. The model does scheduling and deployment, using average miles per year in CONUS and overseas. Maintenance costs are assumed to be given. One average is used for worldwide maintenance depot shipment, one for second destination transportation costs, etc. Most equations are simple sums and products of averages. For example:

(Total depot maintenance support costs) = (average annual civilian labor maintenance support cost per man year) × (average annual civilian man-years in maintenance support).

Both terms on the right are input to the model. The model produces many cost tables, but is basically a convenient arithmetic device and report generator using input data completely specifying costs. Ref. 22 contains four other models, similar to the first but with fewer equations. The last model is a 6-equation "quick and dirty" cost calculator, with a brief discussion of "accounting" sensitivity analysis. This sensitivity analysis varies individual parameters by 5, 10, 15, and 20 percent and plots the output cost variation. Another accounting model, for single barrel automatic weapons, is described in Ref. 23. Such models are still being developed, albeit for limited purposes.

7-6.2.4 The Descriptive Trade-off Model With Decision

The next step in the evolution of trade-off models was the addition of suboptimization and internal decision rules to descriptive models. Such rules were initially related to repair parts provisioning, inventory levels, and selection of manpower levels for maintenance activities imbedded within such models. The models produce a single "design" for each set of input parameters, with certain features of the design chosen for "best" performance. Often, inconsistent rules were used to suboptimize different sections of such models. For example, there are cases of sparing for a guaranteed probability of stockage and manning for least cost in the same model. Descriptive models with decision are usually much more detailed than simple descriptive models and sometimes have elaborate deployment and operational simulation sections.

Ref. 24 describes a model for the Armed Reconnaissance (SCOUT) Vehicle Support phase. The model accepts a deployment schedule, standard costs for Army operations, and maintenance engineering analysis data for "an interrelated group of ARSV assemblies" and calculates yearly costs incurred by the assembly group (using the Functional Group Code, or FGC) over the support phase of the system life cycle. Deployment is modelled by year and theater, including such theater characteristics as transit times and distances between supply levels within a theater. Standard Army costs are used for supply and maintenance operations. Maintenance planning data for an assembly group are taken from a maintenance engineering analysis. The model stores the costs developed in any run for a particular (input-data specified) FGC, so that subsequent runs may include the FGC as a single part and the entire system may be built up through successive runs. The model suboptimizes provisioning, supply storage, and distribution through policies in TM 38-715, TM 38-715-1, AR-710-2, and replenishment storage levels through policies in AR 710-1. Like most detailed models, failures and replacements are calculated from such factors as usage, failure rate (in this case based on miles), and population. Maintenance factors are based on number of failures per 100 end items per 1000 miles. Age is considered. Preventive maintenance is specified in man-hours by level per 1000 miles; corrective maintenance is specified in man-hours by level per repair action.

Another descriptive decision model for vehicles is described in Ref. 25. This model performs an extremely detailed vehicle fleet simulation. Reliability is a function of the age distribution of the vehicles; maintainability is calculated in the model based on maintenance requirements fulfilled during the simulation. Availability is a function of maintenance requirements and the model exercise, and is inherent (based on maintenance doctrine) and achieved (including the effects of preventive maintenance). Manpower is calculated within the model as a result of maintenance requirements, and repair parts provisioning is generated from parts usage in the simulation. The model deploys vehicles as a function of age; usage of vehicles is controlled by a mission-type (miles to travel) destination. Maintenance requirements are generated from "incidence rate" curves showing the relative frequency of occurrence. The model schedules maintenance on a first-in, first-out basis unless manpower or parts are below requirements. In such a case, random downtime is generated, with the maintenance action longest overdue having the highest probability of selection for execution. Many other aspects of *M* trade-off analysis are treated in the model, which produce detailed reports by cost elements, time, item, and location. While the model obtains data from TAERS (The Army Equipment Record System), it uses a large number of distributions because of its detail. The deployment portion, for example, uses $(2 + 2K)$ distributions for each location, 2 per vehicle category, showing the number entering or leaving overtime, and $2K$ (for K time periods) for each vehicle type. Each pair shows (for a given time period) the "quality" (age in miles) of vehicles entering and leaving a location. Use of deployed vehicles at each location requires further $(1 + K)$ distributions, one showing the number of vehicles required for mission overtime, and K (one for each time period) each showing usage in miles (mission) and percent of available vehicles experiencing that usage.

7-6.2.5 Prescriptive Trade-off Models

These models compare alternatives and select an "optimum" solution as defined by the criterion function used in each model. These models vary from simple screening rules, through full search and comparison, to linear and dynamic programming models.

Ref. 26 describes prescriptive criteria for repair versus discard decisions, and contains a useful approach to low-cost sequential analysis of such decisions, and a helpful bibliography of related models. The reference contains a series of simple screening rules, applicable quickly and at low cost, at each stage in the system life (development of design specifications, initial design or item selection, initial source coding for provisioning, coding/design review, repair action). Starting with the earliest stage, screening rules define an item as discard, repair, conduct a full economic analysis to decide, or postpone the decision until the next step. The basic screening strategy is to first assume repair until discard is justified and next analyze the highest level of assembly. Then, if a repair decision is made, the next lower level of assembly is analyzed. The approach advocates "integrated decision analysis" since the repair versus discard decision subsumes decisions about reliability vs unit cost, standardization vs nonstandardization, type of procurement and volume purchased, contractor vs military maintenance, preventive vs corrective maintenance, level of maintenance, and centralized vs decentralized maintenance. A number of useful considerations are suggested:

a. Standard costs are misleading; cost is related to the number of similar units of equipment subjected to the same operation simultaneously.

b. The economics are critically related to other decisions subject to change during the service life of equipment (e.g., varying parts cost over successive procurements).

c. Some complex, detailed decision models are costly to implement and the return doesn't always justify the cost.

d. Current data systems don't always provide quantitative data in appropriate detail.

e. Time often prohibits detailed analysis of all items. For cases typified by these considerations, screening rules can sometimes provide a solution.

The next prescriptive model, which specifies repair level and maintenance doctrine (repair or discard) for modular assemblies, is also oriented toward repair/discard decisions (Ref. 27). This electronic model was developed for discard-at-failure analysis, based on an earlier model containing a number of deficiencies. The report points out some important considerations for modeling:

a. Nonlinear relationships between maintenance workload and such services as manpower and test equipment must be modelled accurately.

b. Acquisition costs which vary with design alternatives must be included.

c. Common dimensions should be used. Investment in greater manpower or more replacement items should be considered in terms of operational readiness return, not, for example, manpower in workload terms and replacement items in terms of confidence against outage.

d. Unit cost differences may vary with absolute cost but total cost differences (unit differences X population) are what count.

e. Module cost is a nonlinear function of number of parts; in addition, other factors (standardization of module type and size, wiring techniques) may have more significant effect on cost.

f. Maintenance philosophy should be a function of resultant support cost and not established independently.

g. Proration of costs is very dangerous.

Ref. 27 specifies a cost model in detail; costs are a function of location of maintenance. The model is run for alternative module designs until all but one module is eliminated on a cost basis. The model considers 12 maintenance alternatives involving repair or discard of module group and subunits at different levels. (Check-out level is excluded from the decision process.) The model compares alternatives for repair or discard at different levels in sequence, to permit stepwise choice of the least-cost doctrine (repair or discard) and level. Simplified sparing calculations are made. The model is validated with a detailed applied case, the F105 bomb-toss computer. Details of sparing; manning, and other calculations are provided, as is a series of appendices providing details of field survey questionnaires, data, and model constants.

After screening rules and sequential selection among a limited number of alternatives, the next level of sophistication in prescriptive models is the pure search algorithm. This process, "marching through the solution space", is represented in Ref. 28; a model for trade-off of repair levels. This "Integrated Logistics Support Analysis Model (ILSAM)" calculates all possibilities asked for, and selects from among them. It can handle constrained optimization problems through adding up all individual values in a feasible solution (of maintenance cost or downtime or parts stocked), comparing the result with the constraint, and (if the constraint is not met) successively adding to and dropping elements from the feasible solution having the least effect on the objective function, to bring the result closer to the constraint value. If the objective were minimum maintenance cost subject to a constraint on downtime, for example, the model would first calculate the unconstrained least-cost solution (by trying all possibilities and sorting the results) and next check the downtime to see if the constraint was met. If not, the model would successively add alternatives to the unconstrained solution that would reduce downtime, adding them in order of least increase in cost first, until the constraint was satisfied. If the constraint could not be satisfied, the solution closest to satisfying the constraint would be presented.

The model of Ref. 28 determines repair policy, maintenance configuration, and parts stock levels for a weapon system, given the weapon configuration, tactical deployment, and system A, *M*, and R. It can be used for repair/discard analysis and to provide sensitivity analysis to failure rate, weight, and cost. It can be used for primary estimates of range and depth of replacement parts. It is intended for early general stages of weapon system design. If failure mode analysis is to be run, failure rates for each mode must be provided. The model will handle up to 100 assemblies of 500 parts each, with a maximum of five levels of maintenance and 50 locations at the lowest level, 25, 15, 6, and **4** a the second through fifth level, respectively. By codin the same assembly several different ways, different failure modes, repair policies and deployments, for example, may be compared. Ref. 28 provides a detailed discussion of the analytic and constrained optimum calculations, and should be referred to for such detail. Three main issues arise for this model: it may not find a feasible solution; it may not find an optimal solution; it uses a lot of computer time. It has been described here for heuristic purposes.

Ref. 29 describes the most recent prescriptive model to be discussed here; it is representative of recent optimizing models and is recommended for consideration. The optimizing approach used is similar to that of Dynamic Programming. The model is intended as a generalized electronic maintenance model ("GEMM"). It is an analytic model, not a simulation model, and considers 35 different maintenance allocation possibilities, using mean values for input data, and Poisson stockage rules (TM 38-715-1). Force structure data include number of equipments; number of organizational, direct, general, and depot support shops; and distance between shops. Life-cycle support costs and operational availability are calculated. Standard Army logistic support rules are applied. The effect of replacement parts policy, manpower, and test equipment, for example, on life-cycle cost and availability can be calculated. *R* and *M* trade-offs can be performed through the effect of design changes on *R* and M. Sensitivity analyses may be preferred.

The GEMM model selects the least-cost ("optimum") maintenance philosophy. First, the least-cost FIP-fault isolation to part (organizational [FIPO], direct support [FIPDS], general support [FIPGS], depot [FIPD], or throw-away module at organization [TAMO]) technique is selected for a part, assuming fault isolation to module is performed at organizational level (FIMO). The model next calculates and sums the least-cost FIP strategy for every other part, derived in the same way. This sum is the total FIMO cost of least-cost FIP or TAM strategies for each part, assuming FIMO. Next, the analysis "goes up" one level, and calculates transportation cost from FICO (fault isolation to component-organizational) to FIMO. Next the least-cost FIP strategy for each part is calculated assuming FIMDS. This process is repeated for FIMGS, FIMD, and TACO (FIM-general support, FIM-depot, and throw-away component at organizational level). This entire process is next repeated for each component. The result is the total "Fault Isolation to Module" cost at any level, assuming least-cost "Fault Isolation to Part" at that level and above. From these costs, the least-cost FIM strategy is selected (along with its associated FIP strategy). Each time GEMM compares alternate costs at a particular level, it saves the least-cost strategy only.

The GEMM process described is repeated to get the optimum maintenance philosophy specifying the level at which each module, component and part should be repaired or thrown away. It also determines the test equipment, manpower, and stockage requirements for this optimum (least-cost) maintenance philosophy since these costs are included in the model calculation process.

The model uses the fundamental theorem of Dynamic Programming for an optimum allocation policy; whatever the initial allocation is, the remaining allocation must be optimal with regard to the total possible allocation remaining after the first one has been made. GEMM begins by assuming that a fault has been isolated to a module in the initial allocation, and finds the optimal fault isolation to part strategy for each possible outcome of the initial FIM allocation doing this for all parts. This gives the optimal FIP strategy for any FIM, and it must be the optimal strategy for the optimal FIM (which we do not yet know). Now GEMM assumes that the initial allocation has established the optimal FIC, and finds the optimal FIM (subsuming the optimal FIC for each possible FIM, already determined) for each possible FIC. GEMM continues this process at higher and higher decision levels, until the network has been optimized.

The GEMM model, for speed of computation, uses *R* and *M* information for each end item, component, and part-class. Parts are assigned to classes containing, for example, about 25 parts, each similar in size, weight, and *R* and *M* parameters. (This assignment is up to the user and is a function of the size of the dimension statements his computer can accommodate.) GEMM also contains a subroutine to apportion reliability parameters to modules, components, and end items if parts reliability only is known. It uses mission profile data including operational hours, maintenance shop available hours, and force structure data including maintenance shop network and distance, and number of equipment supported per shop. Test equipment requirements are quantified by type and cost; maintenance personnel are quantified by MOS type, pay, and allowance per year. Attrition factors can reflect peacetime or wartime behavior. Stockage information includes an array of confidence levels to be investigated, turnaround, order and shipping times, length of replenishment periods, cost of replacement items, and economic life of equipment under study.

GEMM model outputs include operational availability based on the optimum (or any other of 35 possible) maintenance policies. GEMM produces life-cycle support costs by category; maintenance allocation for repair of all modules, components and the end item;

modules or components for throw-away maintenance (and level of throw-away); stock (by level); test equipment per shop (at each level); maintenance personnel requirements by MOS; and total force structure requirements. The model also produces graphical outputs, including availability or support cost or cost-effectiveness versus sensitivity to change in one or a set of variables.

Ref. 30 is a recent book on optimization and probability in system engineering. It presents a detailed but accessible mathematical treatment of methodology useful for system reliability, maintainability, availability, and dependability analysis, including calculus, linear programming, recursion, Markov and queuing approaches, and is strongly recommended as an advanced text or reference work for system engineers and analysts. Ref. 31 describes a useful optimization technique.

7-7 LIFE CYCLE COSTING

In par. 7-3 we discussed the use of cost structures for analysis, synthesis, and evaluation of systems. In par. 7-4 we discussed the development and use of cost estimating techniques and relationships. In par. 7-6 we treated comparison of design alternatives and selection between them to produce a desired system. The conclusion of the process is to present the life-cycle cost (LCC) of the selected system. If we have done the previous work properly, the system life-cycle costs for development, ownership and operation will reflect those for a least life-cycle cost system within performance and mission envelopes or, if the design task is so specified, the most cost-effective system. In performing this "final" life cycle costing we must carefully apply the same ground rules applied to the earlier trade-offs, so that the system we cost is assumed to be in the environment we designed it for. Without such an intimate connection between trade-off ground rules and final life-cycle rules, we might design a system using one set of criteria and evaluate it using contradictory ones. Particularly in a competitive design environment, such consistency is not only desirable, but required.

Life-cycle costing may also be used to compare or evaluate systems, however designed, for procurement purposes. Components and items may not be subjected to formal design trade-offs but may be "off-the-shelf". LCC comparison can be used to identify a preferred system, after required effectiveness has been assured. Alternatively, in cost-effectiveness comparisons, LCC should be used as the relevant cost measure.

When performing LCC analysis at the end of a design process, or for comparison of existing hardware, engineering cost estimates are used to estimate many element costs which, during design, were statistically costed or costed using aggregated techniques. Final LCC analysis considers designed systems; hardware and policies have been determined; and maintenance doctrine is specified in detail, supported by maintenance engineering analysis or prior policy. Logistic doctrine has been established—supported by ILS, sparing and provisioning policy, and a logistic system. A complete presentation of life-cycle costs for each alternative system (in a competitive environment), or for the selected system, results. If the final LCC estimates are to be used for selection, they must be carefully validated by examination of analyses, assumptions, and data used to derive them. Otherwise, there is the danger of "competition by assertion". (In a competition a common basis for costing of Government-controlled variables must be provided to competitors, or specified in a uniform way after competition-related LCC analysis is complete. When differing designs incur different Government-related costs, these consequences must be made clear to designers in advance.)

Once a single, validated set of LCC estimates has been derived for a system, it may be made the basis of incentives and penalties, and used to monitor "returned" (experienced) costs of the system as its life-cycle proceeds. This paragraph describes life-cycle cost management issues for the complete process from parametric analysis through trade-offs, including vendor procurements by system designers, "final" life-cycle costing, evaluation, and validation. It concludes with a discussion of some subelements of the life-cycle costing process which can be applied to small- and medium-scale design and procurement issues.

7-7.1 LIFE-CYCLE COSTING MANAGEMENT—LARGE-SCALE SYSTEM EXAMPLE

Life-cycle cost analyses and trade-offs have been performed using many management structures. In one illustrating major system design (Ref. 32), a single controlling group of life-cycle cost analysts —harged with explicit responsibility for analyses, trade-offs, and final life-cycle costs at the conclusion of system design, operating through a working group structure—proved effective.

7-7.1.1 System Parametric Analysis

Initially, system parametric analyses were performed, using a computer model which contained mission and performance, hardware design, and life-cycle cost submodels. Mean values were used for each mission and performance model, with tables of time percentage in each mode (e.g., deployed), and within mode in each submode (e.g., at each speed). The model was used to step through parameters sequentially. After initial variation of key parameters, a sequencing of parameters was established which permitted step-by-step design optimization. Sensitivity analysis also revealed those parameters which could be separately optimized, being unaffected by variation in other parameters within the ranges of interest. (Most system parameters are sharply constrained by mission and real-world factors.) In this process, the model represented a "rubber system" which could be pushed and pulled in different directions to establish a "best" baseline design. Detailed analyses and trade-offs of subsystems and policies (including maintenance and logistic plan) could then consider variations on the baseline designs. This parametric design stage used a single, project-level working group including performance and mission analysts, life-cost analysts, parametric engineering design experts, and ILS analysts. One or two individuals in each area of expertise were used, aided by two computer programmers.

7-7.1.2 Subsystem Design and Special Studies

The next step, subsystem design trade-offs and detailed special studies, used a larger number of working groups.

Each working group was responsible for life-cycle cost analyses and trade-offs in a specific area. During design analysis, working groups were identified to subsystems (e.g., propulsion, weapon, control, communication). Each working group included a subsystem design engineer, a reliability and maintainability analyst, an integrated logistic support specialist, a system requirement specialist (where necessary) and was chaired by a life-cycle cost analyst. The LCC analyst/chairman coordinated the inputs of each member and was formally responsible for preparation of the full analysis or trade-off. Hardware price estimates were obtained from an internal pricing group and from vendors.

A particular reliability and maintainability specialist, ILS analyst, and LCC analyst often participated in several such groups. The final analysis required the concurrence of the design engineer.

The least-LCC alternative (in trade-offs) was the selected one. If a higher-cost alternative was strongly supported, its selection required the explicit approval, in writing, of the program manager. Such exceptions were rare, and resulted from design and cost uncertainties, or technical risk, which made a few trade-off results ambiguous.

Some special working groups were formed to deal with analytic issues ("special studies") that cut across hardware subdivisions, including manning/automation decisions, overall maintenance doctrine, training plans, and integrated support facilities. Such groups had heavier participation from reliability, maintainability, and logistic specialists, and subsystem engineers from many areas.

7-7.1.3 Final Lifecycle Costing and Procurement Analysis

During final life-cycle costing, with design alternatives selected, a conventional price estimating group calculated acquisition costs for hardware and post-delivery support based on engineering design specification and ILS plans. Other elements of acquisition and operations and support costs were calculated by the LCC group, working with ILS analysts.

Extensive procurement analysis comparisons were made among vendors. At major component level, formal and detailed life-cycle cost trade-offs of competitive equipment, using the full working group approach, provided decision inputs, which were tracked to formal price quotations by these vendors. Intermediate components were compared using a simplified computer tabulation and comparison by vendor prices and significant design differences that could affect operating costs (e.g., *MTBF, MTTR,* replacement item requirements, power consumption). Where designs were similar in technical content and operating cost implications, the "low bidder" was selected. Simple "rules of thumb" were developed for selection of less costly items. The difference in life-cycle operating cost to offset a $100 difference in procurement cost was calculated. Such operating cost guidelines aided rapid comparison of alternatives having different prices, for thousands of lower-level procurement decisions where formal trade-offs were not justified.

A final life-cycle cost package, was prepared by the LCC group, integrating the "price" package of the procurement analysts with the Government costs to be incurred during procurement, installation, and operation of the system, provided by ILS analysts and by the Government.

While the details of the management mechanism may vary with the point of view (planners, designers, procurers, and operators) and the institution (contractor, Government), the basic process, elements, and working group management strategy can be applied to most large- and medium-scale system planning and design projects.

7-7.1.4 Life-cycle Cost Evaluation

After a design team (or a group of competitive teams) has completed its work, the final life-cycle costs presented must be evaluated. An evaluation team, made up of Government engineers (reliability, maintainability, ILS, and LCC analysts, and coordinated by the LCC analysts), is necessary. This team represents the procurement authority and must establish the correctness, reasonableness, and completeness of the LCC estimate. The evaluation team relies on detailed examination of the cost analysis rationale, the appropriateness of the financial data or other estimating basis used, the nature of those cost elements reflecting "prices" (firm fixed price, CPFF, or other basis), and the resultant implied price guaranteed by the offeror (if a procurement), the methods used to treat uncertainty and estimating error, and the degree of explicit detail provided. Independent data available to the evaluation team, consistency with appropriate historical data, and basic reasonableness of estimates all will be considered. During this process, insufficient or inadequate supporting rationale is often a cause for downgrading, and will often lead to formal questions from the evaluation team to the offeror, to clarify or support LCC estimates or predictions.

7-7.1.5 Life-cycle Cost Validation and Monitoring

During and after LCC evaluation, validation studies are often required. Demonstration measurements, operating data, and returned cost measurements after delivery may be required, and may be associated with significant incentives and penalties. Such data also will serve to establish capability for future programs. As a system proceeds through its life cycle, time serves as the ultimate validator.

The data produced through LCC performance measurement should be entered into an appropriate data base for future design use. Periodically, such data should be analyzed to provide LCC data by class of component or equipment, and by element cost. Inconsistencies or changes in such data can sometimes yield system condition or performance information which may be valuable for ongoing system evaluation, correction, and improvements.

7-7.2 LIFE-CYCLE COSTING MANAGEMENT—SUBTASKS

Par. 7-7.1 has sketched an integrated LCC process, applicable to large-scale systems. The team approach may be applied on a smaller scale, using fewer working groups. LCC modeling, subsystem trade-offs,-and special studies may be conducted by individual specialists, after appropriate training. At the competitive procurement level, life-cycle cost considerations should be used in comparing all costs to the Government from a particular procurement action, even when historical design criteria have not yet been modified to require formal LCC analysis in design. When offerors are told that they will be judged on all costs of their product, competitive forces will often produce lower total-cost designs over time.

At the components and parts level, simple LCC comparisons, perhaps using "rules of thumb", are also appropriate. The degree of formal analysis required or used should reflect the total procurement, not the unit value. Thus, careful study of LCC factors for standard electronic modules, for rifles, guns, and vehicles, and for many similar items should be made. Major opportunities exist for large cost avoidances (theoretical savings through selection of lower LCC alternatives). These cost avoidances may be applied to lowering budgets, increasing effectiveness, or improving a force mix. The basic goal is—"More bang for a buck", and it can be achieved through LCC analysis at all levels of planning, design, procurement, and operations.

REI E

1. A. S. Goldman and T. B. Slattery, ***Maintainability: A Major Element of System Effectiveness,*** John Wiley & Sons, N.Y. 1964.
2. DA PAM 705-1, ***Maintainability Engineering.***
3. ***Fast Deployment Logistic Ship System Program Breakdown and Life Cycle Cost Structure,*** Department of the Navy, Naval Ship System Command, 2 September 1966.
4. G. H. Fisher, ***Cost Consideration in System Analysis,*** American Elsevier, N.Y., 1971.
5. D. R. Earles, ***Integrated Logistics Support and Life Cycle Costing,*** Lecture Notes UCLA Course Engineering 885.10 and Raytheon

Company, Bedford, Mass., 1969.

6. R. J. Duffin, E. L. Peterson, and C. Zener, *Geometric Programming,* John Wiley & Sons, Inc., N.Y., 1967.
7. C. F. Christ, *Econometric Models and Methods,* John Wiley & Sons, Inc., N.Y., 1966.
8. E. J. Kane, *Economic Statistics and Econometrics,* Harper & Row, N.Y., 1968.
9. ***BMD*** *Computer Programs,* University of California Press, Berkeley, Calif., 1970.
10. N. Draper and H. Smith, *Applied Regression Analysis,* John Wiley & Sons, Inc., N.Y., 1966.
11. H. Theil, *Applied Economic Forecasting,* Rand McNally & Co., Inc., Chicago, Ill., 1966.
12. W. A. Spurr and C. P. Bonini, *Statistical Analysis for Business Decisions,* Irwin Publishing Co., Homewood, Ill., 1967.
13. J. G. Phillips, *Cost of Operations and Maintenance Activities (Army): Techniques for Analysis and Estimation,* Research Analysis Corporation, McLean, Virginia, January 1968.
14. D. Hoffman, *Throwaway or Repair Implications on Maintenance Cost Program,* Report No. BR-4298, Raytheon Company, Bedford, Mass., September 1967.
15. AR 37-18, *Weapons/Support Systems Cost Categories and Elements.*
16. N. I. Barish, *Economic Analysis for Engineering and Management Decision Making,* McGraw-Hill Book Company, N.Y., 1962.
17. C. Hitch and R. N. McKean, *The Economics of Defense in the Nuclear Age,* Harvard University Press, Cambridge, Mass., 1963.
18. J. M. Henderson and R. E. Quandt, *Microeconomics,* McGraw-Hill, N.Y., 1971.
19. J. J. McCall, "Maintenance Policies for Stochastically Failing Equipment: A Survey", Management Science **11**, No. 5 (1965).
20. J. J. McCall, *Some Recent Development in Maintenance Policies,* Report No. P-3239, The RAND Corporation, Santa Monica, Calif., October 1965 (AD-473 331).
21. A. D. Hall, *A Methodology for Systems Engineering,* D. Van Nostrand Company, Inc., N.Y., 1962.
22. F. M. Chakhour, *Automated Life Cycle Cost Models for Army Weapon Systems,* US Army Weapons Command, Technical Report CPD 69-11, May 1969 (AD-688 599).
23. *Automated Life Cycle Cost Model for Single Barrel Automatic Weapons,* US Army Weapons Command, Technical Report CPD 70-19, October 1970 (AD-714 223).
24. *Armed Reconnaissance Scout Vehicle Support Cost Model (SPCM) Handbook, Draft,* US Army Tank-Automotive Command, AMCPM-RSV-SM, 3 August 1971.
25. *Allison Life Cycle Operating Cost Model,* Report No. EDR 5957, Allison Division, General Motors Corporation, March 1969.
26. *Criteria for Repair Versus Discard Decisions,* Logistic Management Institute, Washington, D.C., May 1966 (AD-484 603).
27. R. E. Purvis and R. L. McLaughlin, *Validation of Discard at Failure Maintenance Model,* Technical Report No. RADC-TR-65-214, Rome Air Development Center, N.Y., February 1966 (AD-479 903).
28. K. D. Harvis, H. R. Gehle, and D. F. Meer, *USAMC Replacement Unit Repair Level Analysis Model,* US Army Management Engineering Training Agency, Rock Island, Ill., December 1970.
29. R. D. White and D. A. Tyburski, *Generalized Electronic Maintenance Model (GEMM),* US Army Electronics Command, May 1971.
30. J. G. Rau, *Optimization and Probability in Systems Engineering,* Van Nostrand Reinhold, N.Y., 1970.
31. R. P. Mandel and Leopold, "Optimization Methods Applied to Ship Design", Transaction Annual Meeting of Society of Naval Architect and Marine Engineers **74**, 30-35 (1966).
32. D. Sternlight, "Life Cycle Costing as an Economic Design and Decision Tool", Naval Research Logistics Quarterly **18**, No. 4 (1971).

CHAPTER 8

STATISTICAL MAINTAINABILITY
SECTION I

INTRODUCTION

8-1 GENERAL

8-1.1 USE OF STATISTICS IN TECHNOLOGY

Statistical analyses are very useful in technology as data analysis techniques for decision-making in design, operation planning, and contingency plan formulation. The following important specific engineering areas are strongly influenced by statistics:

1. ***Producibility.*** Statistics are very useful in manufacturing analyses. With data analysis such questions as the following can be answered. Can the equipment be manufactured within specified tolerances? What are the random factors affecting manufacturing variability, and can they be controlled sufficiently so that the repair/discard criterion for in-process items is kept within economical limits? What are the percentages of out-of-tolerance levels falling short of the given quality level requirements?

2. ***Equipment Calibration.*** Due to drift of components (from temperature, humidity, vibration, as well as from wearout effects), the reading error of measuring instruments (e.g., maintenance checkout instruments) tends to increase in time. What are the distributions of these errors; what are the drift rates; and what are the optimum times at which the instruments should be recalibrated?

3. ***Error Analysis.*** Error analysis is sometimes very important during the design phase when a high level of assurance is required to attain some performance parameter. For example, one might want to design a rocket take-off engine to have enough thrust to reach one mile upstream in X minutes. However, not too much thrust is wanted because this will entail unnecessary weight, volume, and cost penalties. Thus, an error analysis would be important to determine how the errors in component performance combine into an overall error for the engine thrust prediction. When measuring instruments used in component tests have errors in their readings, error analysis takes specific account of these in conjunction with the basic errors caused by manufacturing variability.

Other engineering disciplines which utilize statistics are maintainability, reliability, operations research, engineering acceptance testing, warehouse and inventory analysis, risk analysis, pollution monitoring and control, etc.

8-1.2 APPLYING STATISTICS TO MAINTAINABILITY

Statistics are necessary in the maintainability engineering tasks of prediction, demonstration, and field performance.

The two essential parameters that establish the maintainability of a system are time to perform maintenance tasks and frequency of the tasks. Both of these parameters are subject to statistical variation. Time of maintenance action is inherently a variable—since no one action is repeatable without variance, and no additive combination of actions is repeatable without variance. Frequency of maintenance action is variable due to the probabilistic nature of failure occurrence over a given span of usage. Therefore measures of central tendency and dispersion are necessary in establishing the design criteria for maintainability to be built into equipment. Statistical analysis is essential in determining whether design requirements have been achieved.

The time to repair an equipment depends on such design factors as

1. Accessibility
2. Latches and fasteners
3. Packaging
4. Labelling, marking, and coding
5. Functional testing

6. Work areas, protective devices, and personnel safety

7. Human factor consideration

Furthermore, the times to gain access, to test what has failed, and to repair the fault are variable—even for the same equipment and same type of failure—since these actions are not repeatable without variance.

Maintainability prediction involves the estimation of system maintainability characteristics from the maintainability characteristics of its components. Thus from layout diagrams, access and attachments, functional testing, diagnostics and other design aspects, one predicts the system maintainability characteristics of the system.

Maintainability demonstration involves the estimation of maintainability parameters from physical tests using statistical techniques, and, once the right conditions have been set, it is essentially a pure statistical problem. That is, one first designs maintainability tests and then analyzes the test results to estimate the achieved maintainability characteristics of interest.

Field performance involves the estimation of maintainability characteristics of systems operating in field environments. The basic design and performance characteristics, combined with logistic support provided for the field operation in a given geographic area, also require a statistical analysis approach.

The application of statistics to the three tasks of prediction, demonstration, and field performance is somewhat involved because of the many random variables affecting system maintainability performance. The most obvious instance of this is the prediction of field performance. For example, the true downtime of a system in the field is not only a function of design but also a function of the times to obtain the repair parts, queuing of work loads at the maintenance levels, and technician efficiencies. Since most of these logistic support factors are mainly extraneous to the system and are mostly administrative delay times not directly related to a specific design feature, the maintainability engineer concentrates his attention on the prediction of active maintenance times and their frequencies. These direct maintenance tasks still characteristically involve many random variables.

For example, the amount of downtime a system is expected to require for repair depends on the type of maintenance action to be performed. If it is known to be of a preventive type at a specific maintenance level and at a regular frequency, then a certain degree of accuracy can be achieved. However, if the maintenance action is to repair a subsystem which fails at random for different causes, then the accuracy which can be achieved in predicting the expected downtime is much lower, since the expected downtime is very much a function of which subsystem fails and what kind of failure occurs within that subsystem. The accuracy of prediction is much better for a specific subsystem and a specific type of failure. However, even then a random element is inherent in the expected downtime estimate, because the time to perform any specific task by a technician is a random variable. This partially is explained by the fact that human performance varies randomly and that sometimes one finds an appropriate schematic diagram or achieves a correct fix faster than at other times. Furthermore, different individuals, with different levels of experience, will perform at different speeds. Thus, the time measurements of maintenance efficiency and the analysis of maintainability have to proceed by recognizing the variability of times to complete a maintenance task. This is accomplished by means of probability distributions.

A probability distribution is a function $F(t)$ which, for a given time t, gives the percentage of instances (trials or attempts) in which the maintenance task in question is performed within a span of time no longer than t. Fig. 8-1 is an example of a probability distribution for downtimes of a hypothetical system. It indicates that 50 percent of the actions are completed in time $t \leq t_1$, and that 90 percent of the maintenance actions are completed in time $t \leq t_2$. The theory of the probability distributions is given in par. **8-4.**

8-1.3 STATISTICAL ANALYSIS OF MAINTAINABILITY DATA

Statistics is the science of estimating parameters and /or testing hypotheses concerning a distribution based on random samples. The reason for using samples is a result of such factors as *cost, convenience, time,* and *energy.* However, sometimes this restriction to samples is in the very nature of the measurement problem. This is especially easy to understand when the only tests possible are destructive tests (e.g., testing electric fuses, ordnance fuzes, the kill radius of explosive projectiles, or the number of vibration cycles which an aircraft wing or landing gear can survive before failure). However, the use of samples is also necessary during the development and prototype stage of product genesis.

Design decisions can be made only from historical or experimental data, not from future data.

8-1.3.1 Statistical Tests

Tests of statistical hypotheses are very useful in decision making. Examples of the types of questions

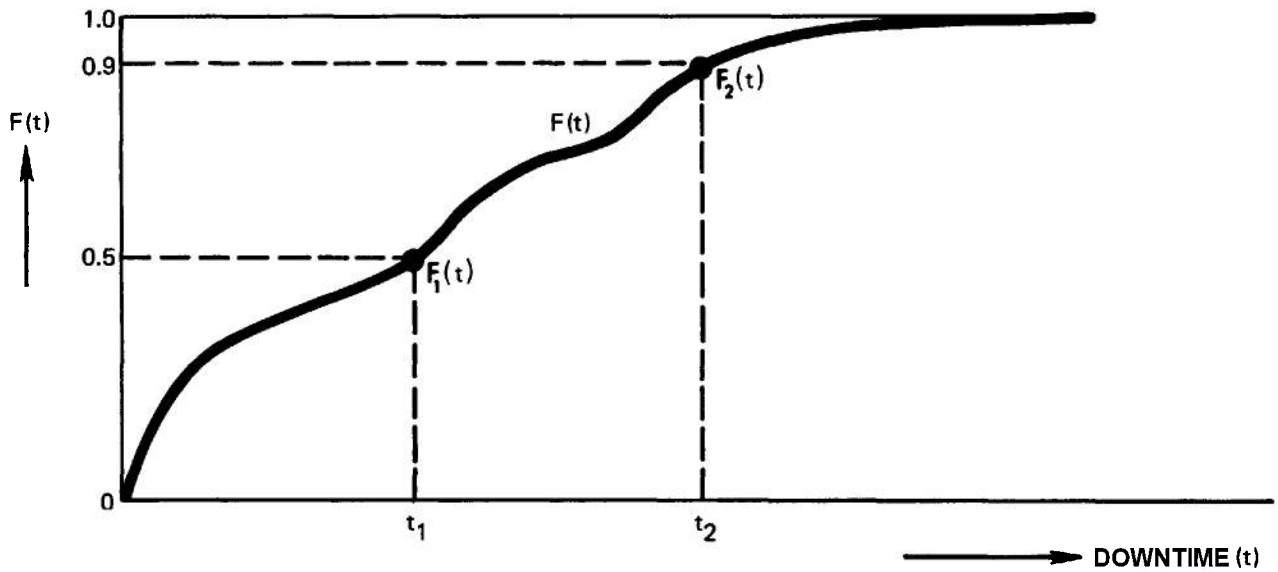


Figure 8-1. Probability Distribution of Downtime

(hypotheses) for which special statistical formulas or theories have been developed are (Refs. 1 and 2):

1. Does the average of a new product differ from the standard, i.e., does the average downtime of a new design differ from the old one?

2. Is the average of a new product greater (less) than the standard?

3. Do products A and B differ in average performance, i.e., do they differ on the average?

4. Does the average of product A exceed that of product B?

5. Does the variability (standard deviation) of a new product differ from the standard?

6. Does the variability (standard deviation) of a new product exceed the standard?

7. Do products A and B differ in performance variability (standard deviation)?

8. Does the variability (standard deviation) of product A exceed that of product B?

9. What is the correlation between two variables; i.e., if the outcome or value of one is known, how much more certain are we of the other?

10. What is the statistical relationship between two variables, e.g., between years of maintenance experience and time to diagnose failures, or what is the effect of alternative diagnosis procedures on the time to diagnose?

8-1.3.2 Consumer and Producer Risks

The first eight questions of par. 8-1.3.1 are special examples of what can be answered by a very general statistical technique known as hypothesis testing. A discussion of the context in which this very general technique is used follows.

The general form of the probability distribution of the items to be tested is known (say, by experience). However, some of the parameters of this distribution are not known (e.g., because they vary with the season, or from batch to batch). A hypothesis, known as the null hypothesis H_0, is asserted to be true; e.g., the mean of the distribution is asserted to be a definite number m. The results of a statistical test are used to determine whether the null hypothesis is true, or whether another hypothesis should be accepted. The types of errors which may result from this test (and which are inherent in the very nature of statistical tests) are of two types. Either the test can result in the rejection of the H_0 hypothesis when in fact it is true, or it can indicate that the H_0 hypothesis is correct when in fact it is false. These two errors are called errors of the ***first kind*** and errors of the ***second kind***, respectively. The probability of an error of the first kind is usually denoted by α, the producer risk; and of the second kind by β, the consumer risk. The producer risk α is also called the *significance level.* Commonly used significance levels are 1% and 5%. In these cases, the decision-maker will reject a true hypothesis only 1% and 5% of the time. Of course, the specific test used must be chosen so that the specified values of α and β are not exceeded.

In simple cases, the choice of test is equivalent to the choice of the sample size to be used. Importance of the sample size is illustrated by the following case. Suppose that the cost of testing n units is nX. This cost becomes large as n becomes large. However, if n is too small, the likelihood of making errors of the first or second kind will be large, and this is also expensive. If the test designer has a choice in the selection of the α and β risks, the type of analysis that follows may be used to determine an optimum cost-effective value of n.

Suppose that the cost of making errors of the first kind is C_1, and that of making errors of the second kind is C_2. Then the expected total cost (testing cost plus decision cost) is $C_1\alpha + C_2\beta + nX$. If C_1 and C_2 can be estimated, then the test designer has to select the triple (α, β, n) so that the total cost is minimal. This leads to the following procedure:

1. Select a number of different pairs (α,β) and for each fixed pair (α,β) determine the minimum sample size $n_m(\alpha,\beta)$ which gives α,β. Ref. 3 contains tabulations of the common cases.

2. If n_m must be an integer (which is not always the case because, for example, in reliability testing a single unit may be tested for any length of time and not just for an integer multiple of its MTBF), then α and β must be re-evaluated in terms of $\alpha(n_m)$ and $\beta(n_m)$ to account for the need of n_m to be an integer.

3. For each of the triples $[\alpha(n_m),\beta(n_m),n_m]$ calculate the expected total cost $C_1\alpha(n_m) + C_2\beta(n_m) + n_mX$.

4. Choose that triple $[\alpha(n_m),\beta(n_m),n_m]$ which gives the smallest total cost calculated by the procedure in Step 3.

The real problem in following this theoretically rigorous approach for determining the optimum sample size n_m is the estimation of the costs C_1 and C_2. These costs depend on the intended use of the equipment being tested, on what will be done if the equipment is rejected in the test, and on the consequences of accepting substandard equipment. For instance, if the whole batch is discarded when the test is not passed, the cost of the error of the first kind C_1 is the cost of the batch. If the whole contract is lost, C_1 is the expected profit which was not realized. If redesign and redevelopment are necessary to save the contract, C_1 is the cost of such redesign and redevelopment. The C_1 costs can often be estimated in a way similar to estimating manufacturing and engineering costs.

The estimation of the C_2 costs is much more difficult. It is related to the setting of the original minimum acceptable standard, i.e., to the accept/reject criteria of the statistical test. Possibly, the activity which set this standard could help estimate C_2. To estimate C_2 is actually equivalent to determining the cost of accepting equipment which passes the test but in fact is of inferior quality that does not meet the contractual performance requirements. The cost of this error of the second kind depends on the length of time the equipment will be fielded and on the incurred degraded performance.

For instance, if it were known that the MTTR instead of being m_1 hours was m_2 hours, with $m_2 > m_1$, then, in a lifetime of T operating hours, the cost C_2 would amount to $C_2 = (m_2 - m_1) T\lambda c$, where λ is the failure rate of the item, and c is the average maintenance labor cost per hour of item maintenance. A very good general discussion of error cost considerations is provided in Ref. 4, which also considers the actual probability density functions of the cost of errors.

8-1.3.3 Sequential Testing

The test design methodology developed from the theory of hypothesis testing for fixed sample sizes is not the most economical available, since it does not utilize all of the data provided in the sample. For example, when the null hypothesis H_0 is "the length of an item is m" but the first few items drawn at random from a sample of size n have been found to have an average length of 5 m, then the testing of the next item of the sample will most probably not affect the test result of rejecting the null hypothesis. Hence, testing of the next sample is, in such a case, a waste of energy, time, and money, and the test can be stopped before all n items are tested. The novel feature in this situation is an elementary use of the individual results immediately, in sequence, as they are obtained for each item of the sample being tested. These ideas have been thoroughly developed into sophisticated techniques of sequential test design, evaluation, and hypothesis acceptance or rejection.

A main characteristic of such tests is that the sample size is not fixed in advance but, instead, a criterion is used to decide whether to accept the sample, reject it, or continue testing. Such decisions can be made continuously, item by item, or sample by sample. In the long run, sequential tests are much more economical than fixed-length tests. Specific details of sequential test design are discussed in par. 6-7.

8-1.4 SAMPLING PROBLEMS IN MAINTAINABILITY DEMONSTRATION

Some of the sampling difficulties in maintainability demonstration are:

1. The maintenance task population is highly heterogeneous and highly stratified. In each stratum the individuals are homogeneous. For example, a maintenance task on a unit with a high failure rate will probably occur relatively often in the system operation and will provide a relatively large segment (stratum) of data with a specific homogeneity (Ref. 5).

2. The test results obtained from a test specimen can be highly misleading with respect to the actual maintenance times experienced during the operational phase of the system (Ref. 5). For this reason, special sampling techniques and test precautions must be taken (e.g., simulation of operating conditions).

3. The maintenance tasks fall into various categories which have different effects on downtime and on the utilization of support equipment. These categories are:

   a. Preventive Maintenance Tasks
   b. Corrective Maintenance Tasks
   c. Servicing Maintenance Tasks.

4. The skill levels required to perform the various maintenance tasks vary considerably and require technicians with different qualifications.

5. The statistical distribution of maintenance time for each of the given categories of maintenance tasks is not only dependent on the MTTR's of each component type, but also on:

   a. The quantity of the components of each type
   b. The location of the component within the system, since location influences the access time and thus will differentiate even between identical component types in different locations.
   c. The frequency with which the maintenance actions become necessary, since this is directly related to the *reliability* of each particular component in its particular location (as affected by ambient vibrations, temperatures, humidity, pressures, etc.).

Since theoretical prediction techniques for any maintenance category (corrective, preventive, or servicing) can vary widely in accuracy, the sampling technique should ensure that representative samples from each category are chosen according to an accurate statistical procedure.

Representative sampling is also important in the selection of different technician skill levels required, because the theoretical predictions may be highly accurate for some types of skill levels but not for others. Furthermore, estimation of the maintenance times within each of the strata (i.e., maintenance categories and skill levels) should take into account location within the system (i.e., access time), and relative frequency of occurrence.

8-2 MAINTAINABILITY FIGURES OF MERIT

8-2.1 DOWNTIME

The three basic types of maintainability figures of merit are those related to cost, manpower, and maintenance time. An example of a cost-related figure of merit is maintenance cost per year. Other figures of merit are maintenance manhours per operating hour (*MMH/OH*), MTTR, *ERT*, M_{MAX}, etc. They are discussed in Chapter 1. These figures of merit are interrelated, and each is important in different phases of the life cycle to different degrees. Of course, the most fundamental aspect of maintainability design is that of maintenance time. Of major importance, as a figure of merit, is the amount of downtime spent performing a given maintenance task.

The downtime for a maintenance task is the amount of time needed to complete the maintenance task. Different tasks will in general result in different amounts of downtime. Moreover, the downtime incurred for the same maintenance action will vary according to the types of tools and maintenance manuals used, the skill of maintenance technicians, and the environment. Furthermore, downtime for the same maintenance action performed under the same conditions by the same individual at different times will in general also be different—thus downtime is a random variable.

Almost all maintenance actions of interest can be divided into smaller subtasks, the time of each of these being a random variable. When the subtasks of a maintenance action are performed *in series*, then the sum of the times to perform the subtasks is equal to the downtime of the composite maintenance action. However, the time to perform each subtask is a random variable and, therefore, the time to perform the composite maintenance action, being the sum of random variables, also will vary from case to case. This means that the probability density function (*pdf*) of the composite maintenance time is a convolution of the *pdf's* of **performing** the subtasks. The mean times of these *pdf's* of course, remain additive, per theorem of **Eq. 8-26.** If we define

D_1 = mean time to restore system to operating condition
D_2 = mean time to diagnose the fault
D_3 = mean time to gain access to faulty module
D_4 = mean time to obtain repair part
D_5 = mean time to replace faulty module
D_6 = mean time to checkout if correct repairs and installation have been made
D_7 = mean time to close up access panels

then the mean downtime equals

$$D_1 = D_2 + D_3 + D_4 + D_5 + D_6 + D_7 \quad (8\text{-}1)$$

The term downtime has different quantitative meanings to the activities involved in a maintenance action. As an example, consider the scenario that follows:

A helicopter has landed with its engine about to fail due to excessive vibration. Symptom analysis shows that it will take about 8 hr to repair the engine in a workshop. However, a spare engine is available and it takes 1 hr to replace the faulty engine. What is the downtime caused by this failure? From the pilot's viewpoint it is 1 hr at the organizational level. From that of the depot maintenance personnel administrator it is 9 hr (1 hr for organizational replacement plus 8 hr for depot repair). The actual calendar downtime of the removed faulty engine may be several days, due to administrative delays or waiting for repair parts. The amount of time lost in performing a mission may depend on other circumstances. For example, if the aircraft were scheduled to be serviced and the service turnaround time were scheduled to exceed 1 hr, no mission time would necessarily be lost. But if the pilot were forced to return because of the faulty engine before accomplishing his mission objective, considerable loss of mission time may be involved. If the mission objective was accomplished, little or no mission time was lost. But if aircraft performance was so degraded because of the excessive engine vibration that enemy damage was sustained requiring 12 hr of fuselage repair, such enemy-caused downtime could be attributed to the engine.

Ref. 6 provides an excellent account of how maintainability requirements (both *resource* and *operational*) can be used during the design of a major system. It shows that with proper analytical computer-oriented techniques, great depth and detail of investigation can be accomplished to provide a basis for design decisions at a system level.

Other aspects of downtime which are relevant to a maintainability analysis and which should be carefully considered are:

1. Corrective maintenance does not necessarily cause system downtime, because the correction may (with proper design) be accomplished without any interruption in system operation (e.g., if replacement parts are carried in the system, or if standby components are provided and repair provisions exist in the vehicle).

2. Preventive or service maintenance may cause unscheduled downtime if the servicing equipment malfunctions during servicing, or if the servicing personnel make mistakes in servicing.

3. Administrative downtime (specifically downtimes caused by sickness or injury of personnel, or by higher priority orders, etc.) may increase the maintenance downtime of equipment. Such downtimes should be recorded during demonstration tests but not included in the acceptheject decision of the equipment maintainability demonstration tests.

8-2.2 SPECIFICATIONS

Downtime is a fundamental figure of merit. It greatly impacts on equipment availability. The soldier under battle conditions does not care how much a maintenance action costs (in terms of new parts, special tools, training level of maintenance technicians, or sophistication of diagnostic procedures), but he is very much concerned with how quickly his equipment can be restored to proper operating condition so that he can use it to carry out his mission. However, since national resources are not unlimited, costs always must be considered.

Cost factors of interest in the context of planning and accounting are labor, parts, training, and power consumption. These factors are of special interest to design and maintenance personnel. Since downtime in the field does in fact depend on the availability of repair parts, the logistic support procedures (e.g., types and amounts of repair parts stored at various maintenance levels, types and amounts of repair tools, etc.) do affect the downtime of operational equipment. Thus, the logistic support sophistication should be considered at least generally, even during design. In fact, the guideline that follows should be employed by every maintainability engineer during design.

Other pertinent things being equal, choose parts or components for a particular location which have been most proven in other areas and/or other equipment. Commonality of parts will reduce the total number of different repair parts and will facilitate training. From the increased familiarity that will result, a higher level of expertise will be maintained. Warehousing will be simplified, and repair parts availability will be greatly improved.

A disadvantage of using downtime as the only figure of merit is that a long downtime might be tolerated if it occurs very infrequently. But unless the designer knows the frequency of occurrence of long downtimes he would not be able to judge their impact on system effectiveness. A means of circumventing this difficulty is to use availability as a measure of system effectiveness. See par. **1-7.3** for a discussion of availability factors.

In the long run, inherent, steady-state availability A is good design guideline and is given by

$$A = MTBF/(MTTR + MTBF) \tag{8-2}$$

where

$MTBF$ = mean time between failures
$MTTR$ = mean time to repair

Inherent availability is thus the probability that a system is not in a state of active repair after it has been in use for some time. Availability reflects both reliability and maintainability. It can be large (close to 1.0) even if the $MTTR$ is large, provided the $MTBF$is much larger. However, even though the inherent availability can be very useful as an auxiliary design guideline, the $MTTR$ still needs to be specified since it is a fundamental measure of system maintainability design characteristics. However, specifying availability and $MTTR$ by themselves is not satisfactory. It is better to specify some points on the distribution repair times. Thus, one could specify that 90% or 95% of the maintenance actions be completed within a time $t \leq M_{MAX}$, where M_{MAX} is the corresponding percentile point (see par. **8-3.4.3**) of the distribution which has the $MTTR$ as its mean, given by:

$$MTTR = \int_0^\infty G(t)\,dt \tag{8-3}$$

where $G(t) = 1 - M(t)$ is the probability that repair will not be completed in time t, while $M(t)$ is the probability that repair *will* be accomplished by time t $M(t)$ is also called the Maintainability Function and

may assume many different shapes. For instance, if repair time is exponentially distributed, $M(t) = 1 - \exp(-$ pt), where μ is the repair rate.
Then

$$MTTR = \int_0^{\infty} \exp(-\mu t)\, dt$$

$$= -\mu^{-1}[\exp(-\mu t)]_0^{\infty}$$

$$= -\mu^{-1}[0 - 1] = 1/\mu \qquad (8\text{-}3a)$$

8-3 PROBABILITY DISTRIBUTIONS IN MAINTAINABILITY

To develop the theory of probability distributions in maintainability, some basic probability concepts are first discussed.

8-3.1 BASIC LAWS OF PROBABILITY

Consider the possibility of occurrence of an event. The event might be one which is certain to occur, one which is impossible to occur, or one which has some possibility of occurring. The degree of certainty, the likelihood, or the odds that the event will occur is measured by its probability. Several alternative definitions and developments of this concept are possible. The definition chosen here is not the most rigorous possible, but is presented mainly because of its closeness to the concepts of statistics.

The probability of an event E, written as $P(E)$, is a measure of the frequency (percentage) of the occurrence of the event out of a given number of possible experiments or observations. The measure is normalized to equal 1 when the event is certain to occur in each experiment, and to equal 0 when the event cannot occur in any experiment. Thus, the probability of any event E occurring is a real number between 0 and 1, i.e., $0 \leq P(E) \leq 1$. Note that 100 $P(E)$ is equal to the long-run percentage of times that the event E occurs out of all the possible chances it had for occurring.

The basic laws which the probabilities of any set of events must satisfy are (Refs. 2 and 7):

1. If in an experiment two events E_1 and E_2 can occur either singly or together, then the probability that either E_1 or E_2, or both E_1 and E_2, occur is

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2) \qquad (8\text{-}4)$$

This is the addition theorem of probabilities. The symbol $\cup$ stands for union of events, and $\cap$ stands for intersection of events (Ref. 2, pp. 10-13).

2. If the two events E_1 and E_2 are mutually exclusive, i.e., both cannot occur in a single experiment (they do not intersect), the probability that either E_1 or E_2 occurs is

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) \qquad (8\text{-}5)$$

This is a special case of the previous addition theorem in that for mutually exclusive events $P(E_1 \cap E_2) = 0$.

3. If in addition to being mutually exclusive, the two events are also complementary, i.e., one of them must occur in any experiment, the probability that either E_1 or E_2 occurs is unity

$$P(E_1) + P(E_2) = 1 \qquad (8\text{-}6)$$

For example, if the outcome of an experiment can be only a success E_1 or a failure E_2, then the probability that either event E_1 or event E_2 occurs in that experiment must equal 1. Such two events are said to be complementary.

4. If in an experiment two events E_1 and E_2 can occur, and it is known that one of the events (say, E_2) has already occurred, the probability of occurrence of the other event E_1 is

$$P(E_1 | E_2) = P(E_1 \cap E_2)/P(E_2) \qquad (8\text{-}7)$$

This is called the conditional probability of occurrence of E_1, given that E_2 has already occurred (written $E_1 | E_2$). Eq. 8-7 may be rewritten as

$$P(E_1 \cap E_2) = P(E_1 | E_2)P(E_2) \qquad (8\text{-}8)$$

Eq. 8-8 states that the probability of both E_1 and E_2 occurring is equal to the conditional probability of the occurrence of event E_1, given that E_2 has already occurred, *multiplied* by the probability that event E_2 actually occurs. This is sometimes called the generalized product law of probabilities. Of course, the sequence of occurrence of E_2 and E_1 may be reversed, i.e.,

$$P(E_1 \cap E_2) = P(E_2 | E_1)P(E_1) \qquad (8\text{-}8a)$$

As an example, one may ask what is the probability of drawing two aces in succession from a deck of 52 playing cards, if the first card drawn is not replaced (Ref. 2, p. 47). Event E_1 is drawing an ace at the first draw, and event E_2 is drawing an ace at the second draw. Thus, $P[E_1 \cap E_2]$ the probability that both E_1 and E_2 will occur as specified is

$$P(E_1 \cap E_2) = (3/51)(4/52) = 1/221 \qquad (8\text{-}8b)$$

where $P(E_2 \mid E_1) = 3/51$ is the conditional probability of drawing an ace on the second draw given that an ace was drawn at the first draw (i.e., 3/51 because only 3 aces are left at the second draw and the number of cards is only 51) and $P(E_1) = 4/52$ is the probability of drawing one ace out of four from 52 cards at the first draw. More on conditional probabilities will be found in par. 8-3.5.

5. If the events E_1 and E_2 are independent of each other, i.e., $P(E_1 \mid E_2) = P(E_1)$ and $P(E_2 \mid E_1) = P(E_2)$, then Eq. 8-8 assumes the form

$$P(E_1 \cap E_2) = P(E_1)P(E_2) \qquad (8\text{-}9)$$

Stated in words, the probability that two independent events E_1 and E_2 occur in an experiment is the product of $P(E_1)$ and $P(E_2)$. As a matter of fact, two events can be defined as *independent* of each other only if Eq. 8-9 holds.

In terms of the previous card example, this case would arise if one would replace the first drawn card. In that case $P(E_1) = 4/52$ and also $P(E_2) = 4/52$ because the card replacement makes the two events independent, and then $P(E_1 \cap E_2) = (4/52)(4/52) = 1/169$.

8-3.2 CUMULATIVE DISTRIBUTION AND PROBABILITY DENSITY FUNCTIONS

Associated with random variable X is its ***cumulative distribution function* (*CDF*)**, $F(x)$ which is defined to be $F(x) = P(X \leq x)$. When several different random variables (e.g., X, Y) are being discussed, a suffix is attached to distinguish one *CDF* from another (e.g., $F_X(x)$). Since it is a probability, $F(x)$ obeys the laws of probability. Thus $F(-\infty) = 0$ and $F(+\infty) = 1$, and if $x < y$, then $F(x) \leq F(y)$ (i.e., $F(x)$ is a nondecreasing function of x). When $F(x)$ is continuous, then X is said to be a ***continuous*** random variable. When $F(x)$ varies only by jumps and is constant between jumps, then X is said to be a ***discrete*** random variable.

An example of a discrete random variable occurs when the probability that X will be equal to a_n [we write $P_n(X = a_n)$] is equal to P_n and is 0 elsewhere, and when $(a_n)_{n=1}^{N}$ is a sequence of N numbers. In increasing order of their numerical values we have

$$P(X = x) = \begin{cases} P_n, & \text{if } X = a_n \text{ and where } 1 \leq n < \infty \\ 0, & \text{otherwise} \end{cases} \qquad (8\text{-}10)$$

By the definition of the *CDF*,

$$F(x) = \sum_{n=1}^{N_x} P_n \qquad (8\text{-}11)$$

where N_x is that unique integer which is the largest integer that is not bigger than X. Since one of the events must occur (i.e., X must assume some value), a condition of the values of the P_n is that

$$\sum_{n=1}^{\infty} P_n = 1 \qquad (8\text{-}12)$$

The ***probability density function* (*pdf*)** $f(x)$ of a random variable X is the probability

$$f(x) = \lim_{\Delta x \to 0} P[X \text{ in } (x, x + \Delta x)](\Delta x)^{-1} \qquad (8\text{-}13)$$

If the limit does not exist for any X, then the random variable X will not have a *pdf*. A more intuitive definition of a *pdf*, $f(x)$, is that

$$f(x)\,dx \approx P[X \text{ in } (x, x + \Delta x)] \qquad (8\text{-}14)$$

where $\approx$ denotes "approximately equals, when Δx is small enough". By definition of the CDF, this can also be written as

$$f(x)\,dx \approx F(x + dx) - F(x) \tag{8-15}$$

since

$$f(x) = \lim_{dx \to 0} \frac{F(x + dx) - F(x)}{dx} \tag{8-15a}$$

which results in

$$f(x) = \frac{dF(x)}{dx} \tag{8-16}$$

Thus $f(x)$, when it exists, is a measure of how often the random variable X will take on a value very close to x.

Also, by Eq. 8-16, we have

$$F(x) = \int_{-\infty}^{x} f(y)\,dy \tag{8-17}$$

and

$$1 - F(x) = \int_{x}^{\infty} f(y)\,dy \tag{8-18}$$

The *CDF* or *pdf* of a random variable may be determined by the construction of histograms. A histogram is a graphical summary of data arranged in such a way as to approximate either the *CDF* or the *pdf*. A more formal definition is facilitated by the following terminology:

1. The ***range*** of a random variable is the set of values it can take.

2. A ***partition*** π of the range is a set of numbers, $a_1 < a_2 < a_3 < \ldots$, all within the range of the random variable.

3. The ***class intervals*** for the partition π are the sets (a_i, a_{i+1}). For example, the *j*th class interval is the set of all numbers between a_j and a_{j+1} from a partition –.

4. A ***relative frequency*** for a given class interval (a_j, a_{j+1}) and for a given set of statistical data for the random variable X is the quantity $f_j = n_j/n$, where n_j is the total number of times X falls within the class interval and n is the total number of times X was observed to occur (i.e., sample size). Note that f_j is an estimate for the probability $P[X \text{ in } (a_j, a_{j+1})]$.

5. A ***cumulative frequency*** F_j for the class interval (a_j, a_{j+1}) and for a given set of statistical data is equal to N_j/n, where N_j is the number of times that X had a value less than or equal to a_{j+1}, and n is as before. Note that F_j is an estimate for the probability $P(X \le a_{j+1})$. The notations of cumulative frequency and relative frequency are related. In fact,

$$F_j = N_j/n = \sum_{i=1}^{j} n_i/n = \sum_{i=1}^{j} f_i \tag{8-19}$$

and

$$f_j = F_j - F_{j-1} \tag{8-20}$$

These relations correspond to those cited earlier between the *CDF* and the *pdf*. One can prove that, under general conditions,

$$\lim_{n \to \infty} F_j = F(a_{j+1}) \tag{8-21}$$

and

$$\lim_{n \to \infty} f_j = F(a_{j+1}) - F(a_j) \tag{8-22}$$

where the *F*s on the right side denote the *CDF* values at their respective arguments. This shows that the more samples or trials one attempts, the more accurate will be his predictions. However, cost and time considerations will limit the number of attempts. The factors influencing selection of an appropriate sample size n are discussed in Chapter 6 of this handbook.

An example for constructing histograms is given at the end of this chapter. To conclude the discussion on histograms, we may define the cumulative frequency histogram as the curve obtained when the cumulative frequencies for a given set of class intervals are plotted as a function of the range of a sample of independent observations of the random variable of interest.

8-3-3 MEASURES OF CENTRAL TENDENCY

The cumulative distribution function $F(x)$ completely describes a random variable X, but this information is often too diffuse to be of direct use in determining the distribution parameters. Data reduction techniques can simplify this process. As a first example of such data reduction techniques, consider the ***measures of central tendency*** of the random variable X. Such measures are the mode, the median, and the mean, i.e.,

1. The ***mode*** of random variable X is that value of x at which $f(x)$ is maximum (i.e., that value of X which is more likely to occur than any other single value). Not all distributions have a unique mode. In fact, for the ***uniform distribution*** defined by (Ref. 2, p. 126)

$$F(x) = \begin{cases} 1/(b+a), & \text{for } X \in (a,b) \\ 0, & \text{for } X \notin (a,b) \end{cases} \quad (8\text{-}23)$$

each value of $X \in (a, b)$ is a mode E is standard notation for "in" and $\notin$ for "not in".

2. The ***median*** of a random variable X is a value such that

$$F(a) = 1 - F(a) = 1/2 \quad (8\text{-}24)$$

Stated in words, the probability that the value of X is less than a is 0.5, and the probability that X is greater than a is 0.5 also. The median might not be unique. As an example consider the discrete random variable with *CDF*:

$$F(x) = \begin{cases} x, & \text{for } 0 \leq x \leq 1 \\ 0, & \text{otherwise} \end{cases} \quad (8\text{-}24a)$$

In this case the median is unique, and is equal to 0.5. However, the random variable with $F(0) = 0.5$, $F(1) = 0.5$ does not have a unique median because for *any* $0 < x < 1$, we have $F(x) = 0.5$.

3. The ***mean*** or the expected value $E[X]$ of a nonnegative random variable X is defined as

$$E[X] = \int_0^\infty x f(x)\,dx = \int_0^\infty [1 - F(x)]\,dx \quad (8\text{-}25)$$

whenever the *CDF*, $F(x)$ has such a form that

$$\lim_{x\to\infty} x[1 - F(x)] = 0 \quad (8\text{-}25a)$$

To prove Eq. **8-25**, we shall integrate this improper integral by parts. First we integrate from 0 to a, then we take the limit $a\to\infty$:

$$\int x f(x)\,dx = \left\{ x \int_0^x f(x)\,dx \right\}_0^a - \int_0^a \left[\int_0^x f(x)\,dx \right] dx$$

$$= \{-x[1 - F(x)]\}_0^a + \int_0^a [1 - F(x)]\,dx$$

$$= -a[1 - F(a)] + \int_0^a [1 - F(x)]\,dx \quad (8\text{-}25b)$$

Since in the limit the first term is

$$\lim_{a\to\infty} a[1 - F(a)] = 0 \quad (8\text{-}25c)$$

the result is Eq. 8-25.

For a continuous random variable X with range from $-\infty$ to $+\infty$, the mean is defined as

$$E[X] = \int_{-\infty}^{+\infty} x f(x)\,dx \quad (8\text{-}25d)$$

The mean is the most common measure of location (also called the average value). The mean is often denoted by m or μ, or by $E[X]$.

Two important theorems on expected values are:

1. The expected value of the sum of independent random variables is the sum of their expected values, or

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] \quad (8\text{-}26)$$

(see the case described by Eq. 8-1).

2. Theorem on Total Expectation

$$E[X] = E_Y[E[X|Y]] \quad (8\text{-}27)$$

where $E[X|Y]$ is the expected value of X, given the actual value of Y, which is known to have occurred, and the subscript Y in E_Y is a reminder that the expected value is being taken with respect to Y (not X).

As a trivial example of this theorem consider the case of a system which has two components A and B with expected system mean downtimes $E(D_A)$ and $E(D_B)$, according to which component fails. Let the conditional probability of component A failure be $P(A)$, given that the system has failed. Then this theorem (Eq. 8-27) says that the expected system downtime, $E(D)$ is

$$E[D] = P(A)E(D_A) + (1 - P_A)E(D_B) \quad (8\text{-}27a)$$

where we use the fact that $E(D|A$ failed$) = E(D_A)$ and $E(D|B$ failed$) = E(D_B)$, and take Y to be the two-valued random variable

$$P(Y = y) = \begin{cases} P(A), & \text{if } y = a \\ P(B), & \text{if } y = b \neq n \\ 0\ . & \text{otherwise} \end{cases} \quad (8\text{-}27b)$$

where Y = a indicates A failed, and Y = b indicates B failed. Then Eq. 8-27 gives Eq. 8-27c, since for any function $g(Y)$

$$E[g[Y]] = P(A)g(A) + P(B)g(b) \quad (8\text{-}27c)$$

8-3.4 OTHER MEASURES

8-3.4.1 Measures of Dispersion

8-3.4.1.1 Variance

Another important item of information obtainable from the cumulative distribution function is the measure of the dispersion of the random variable (i.e., of its tendency to spread from its average value). The most common measure for the dispersion of X is the variance, denoted by VAR[X]. VAR[X] is defined as the average or mean of the square of the deviation from the mean m. When X is nonnegative:

$$\text{VAR}[X] = \int_0^\infty (x - m)^2 f(x)\,dx \quad (8\text{-}28)$$

The expression $(x - m)^2 = x^2 - 2xm + m^2$, so that

$$\text{VAR}[X] = \int_0^\infty x^2 f(x)\,dx - 2m\int_0^\infty x f(x)\,dx + m^2\int_0^\infty f(x)\,dx$$

$$= E[X^2] - 2m^2 + m^2 = E[X^2] - m^2 \quad (8\text{-}29)$$

8-3.4.1.2 Standard Deviation

The standard deviation of a random variable X is the square root of the variance of X. The standard deviation is used frequently in the literature and is denoted by σ. It is given by

$$\sigma = \sqrt{\text{VAR}[X]} = \sqrt{E[X^2] - m^2} \quad (8\text{-}30)$$

8-3.4.1.3 Coefficient of Variation

The coefficient of variation V is defined as

$$V = o/m \quad (8\text{-}31)$$

generally expressed as a percentage. Because of the linear relationship between a and V, V sometimes replaces o in some textbooks.

8-3.4.2 Covariance

When two random variables are being considered in a single context, a very important concept is that of covariance. Written COV(X, Y), the covariance of two random variables X and Y with means A and B, respectively, is defined as

$$\text{COV}(X, Y) = E[(X - A)(Y - B)] \quad (8\text{-}32)$$

When X and Y are independent

$$\text{COV}(X, Y) = E[(X - A)]E[(Y - B)] = 0 \quad (8\text{-}33)$$

i.e., the covariance of two random variables that are independent of each other is 0.

8-3.4.3 Percentiles

The 100pth percentile of an $F(x)$ is defined as the value x_p such that

$$F(x_p) = F(X \leqslant x_p) = p \qquad (8\text{-}34)$$

8-3.5 CONDITIONAL PROBABILITY AND TOTAL PROBABILITY

As mentioned in par. 8-3.1, the conditional probability of a random event A, given that a random event B has occurred, is denoted by $P(A|B)$. A heuristic explanation of the probability law of **Eq. 8-7** is that if we know that event B has occurred, we can immediately restrict our attention to the case when the events A and B occur together. The probability of that event is denoted by $P(A \cap B)$ (read: Probability of A and B). Now out of all the cases when B occurs, which has probability $P(B)$, A *given* B occurs only when they both occur, and $P(A|B)$ is measured in the limit by the ratio of the relative occurrences of these two cases. Hence

$$P(A|B) = P(A \cap B)/P(B) \qquad (8\text{-}35)$$

Example: In n observations, event B occurred n_B times, and event $A \cap B$ occurred n_2 times. Thus the estimate for $P(B)$ is n_B/n and for $P(A \cap B)$ is n_2/n. For $P(A|B)$ the estimate is n_2/n_B. However, this also is the estimate for $P(A \cap B)/P(B)$, because

$$\frac{P(A \cap B)}{P(B)} \simeq \frac{n_2/n}{n_B/n} = \frac{n_2}{n_B} \qquad (8\text{-}36)$$

Thus

$$P(A \cup B) = P(A|B)P(B) \qquad (8\text{-}37)$$

Now combining this with the law of **Eq.** 8-5 leads to the Theorem of Total Probability

$$F(X) = P(X \leqslant x) = \int_{-\infty}^{\infty} P(X \leqslant x \mid Y = y) f(y)\, dy \qquad (8\text{-}38)$$

For discrete random variables, this theorem reads:

$$F(X) = \sum_{n=1}^{w} P(X \leqslant x \mid N = n) P(N = n) \qquad (8\text{-}39)$$

Example: Consider a three-component system with conditional downtime *CDF*s, $F_1(x)$, $F_2(x)$, and $F_3(x)$. $F_i(x)$ is the *CDF* of downtime when it is known that the ith system fails first. Then the discrete version of this theorem, by **Eq. 8-39** shows that the unconditional downtime *CDF* is:

$$F(X) = P_1 F_1(x) + P_2 F_2(x) + P_3 F_3(x) \qquad (8\text{-}40)$$

8-4 CLASSES OF PROBABILITY DISTRIBUTIONS

8-4.1 THE NORMAL CLASS OF FUNCTIONS

The normal *pdf* has a very distinguished place among probability distribution functions and for this reason is presented separately rather than in par. 8-4.3 grouped with other exponential functions. The main reason for such a position is that under very general conditions the sums of a large number of random variables are approximately normally distributed (see Central Limit Theorem, par. 8-8.2). Thus, for maintainability studies the *pdf* of the time to perform a specific maintenance task tends to be normally distributed when the maintenance task is composed of many subtasks (e.g., determine size of bolt to fit, find the corresponding wrench, etc.). Because of these limiting properties described in the Central Limit Theorem, the normal *pdf* also is used often in large sample statistics. The following definitions and properties of normal distributions may be helpful:

1. Definition. The normal pdf is by definition

$$f(x) = (\sqrt{2\pi}\sigma)^{-1} \exp[-(x - m)^2/(2\sigma^2)] \qquad (8\text{-}41)$$

where x is the variable, m is the mean, and σ is the standard deviation. The corresponding *CDF* is

$$F(x) = (\sqrt{2\pi}\sigma)^{-1} \int_{-\infty}^{x} \exp[-(t-m)^2/(2\sigma^2)]\,dt \qquad (8\text{-}42)$$

where t is a dummy variable.

2. Mean and Variance. One can verify by change of variables in the integrand that $E[X] = m$ and $VAR[X] = \sigma^2$. Or, the mean of a normally distributed random variable X with a *pdf*, as in Eq. 8-41, is m and its standard deviation is σ.

3. Notation. In terms of modern notation, a normal variable X with mean m and variance σ^2 is denoted by $N(X|m,\sigma^2)$. $N(X|m,\sigma^2)$ also can denote the *CDF* in Eq. 8-42. The corresponding *pdf* in Eq. 8-41 is then denoted by $n(X|m,\sigma^2)$.

4. The Standard Normal. The random variable with a CDF of $N(X|0,1)$ with 0 mean and standard deviation 1, is called the standard normal variable or unit normal variable. One can easily prove by a straightforward change of variables in $F(x)$ that if $N(X|m,\sigma^2)$, then $N[(X-m)\sigma|0,1]$. The variable $(X-m)/\sigma$ is often denoted by U (for unit normal), or by z, i.e., a measure of the deviation from the mean in units of standard deviation.

Now

$$P(X \leq x) = P[(X-m)/\sigma \leq (x-m)/\sigma]$$
$$P(X \leq x) = P[U \leq (x-m)/\sigma] \qquad (8\text{-}43)$$

where

$$U = (X-m)/\sigma.$$

5. Tables. One can obtain the cumulative probabilities of any random variable $N(X|m,\sigma^2)$ from tables of the standard normal random variable U by the use of Eq. 8-43. Thus all numerical information can be plotted as a function of one parameter U instead of three parameters, (x, m, and σ) for *all* normal distributions—a tremendous advantage. $N(U|0,1)$ and $n(U|0,1)$ are tabulated in Ref. 8 to 15 decimal places for $U = 0.00(0.02)\ 3.00$, and to 10 decimal places for $U = 3.00(0.05)\ 5.00$. Note that $U = a(b)c(d)e$ means that the function is tabulated for values of U between a and e, at increments of b between a and c, and increments of d between c and e. Ref. 9 gives similar tables for values $U = 0.0(0.0001)\ 1.0000(0.001)\ 7.800(0.01)\ 10.00$.

Some very commonly used values of $N(U|0,1)$ are tabulated in Table 8-1.

6. Shapes. The dependence of the shapes of $f(x)$ and $F(x)$ on the parameters μ and σ is illustrated in Figs. 8-2 and 8-3.

7. The Normal Addition Theorem

If $N(X|m_1,\sigma_1^2)$ and $N(Y|m_2,\sigma_2^2)$

then $N[(X+Y)|m_1+m_2,\sigma_1^2+\sigma_1^2]$ (8-44)

In words, the sum of normally distributed random vari-

TABLE 8-1.
COMMON PERCENTILES OF $N(U|0,1)$

U	$N(U\|0,1)$
0.000000	0.5
0.253347	0.6
0.524401	0.7
0.674490	0.75
0.841621	0.8
1.281552	0.9
1.644854	0.95
1.959964	0.975
2.326348	0.990
2.575829	0.9950
2.807034	0.9975
3.090232	0.9990

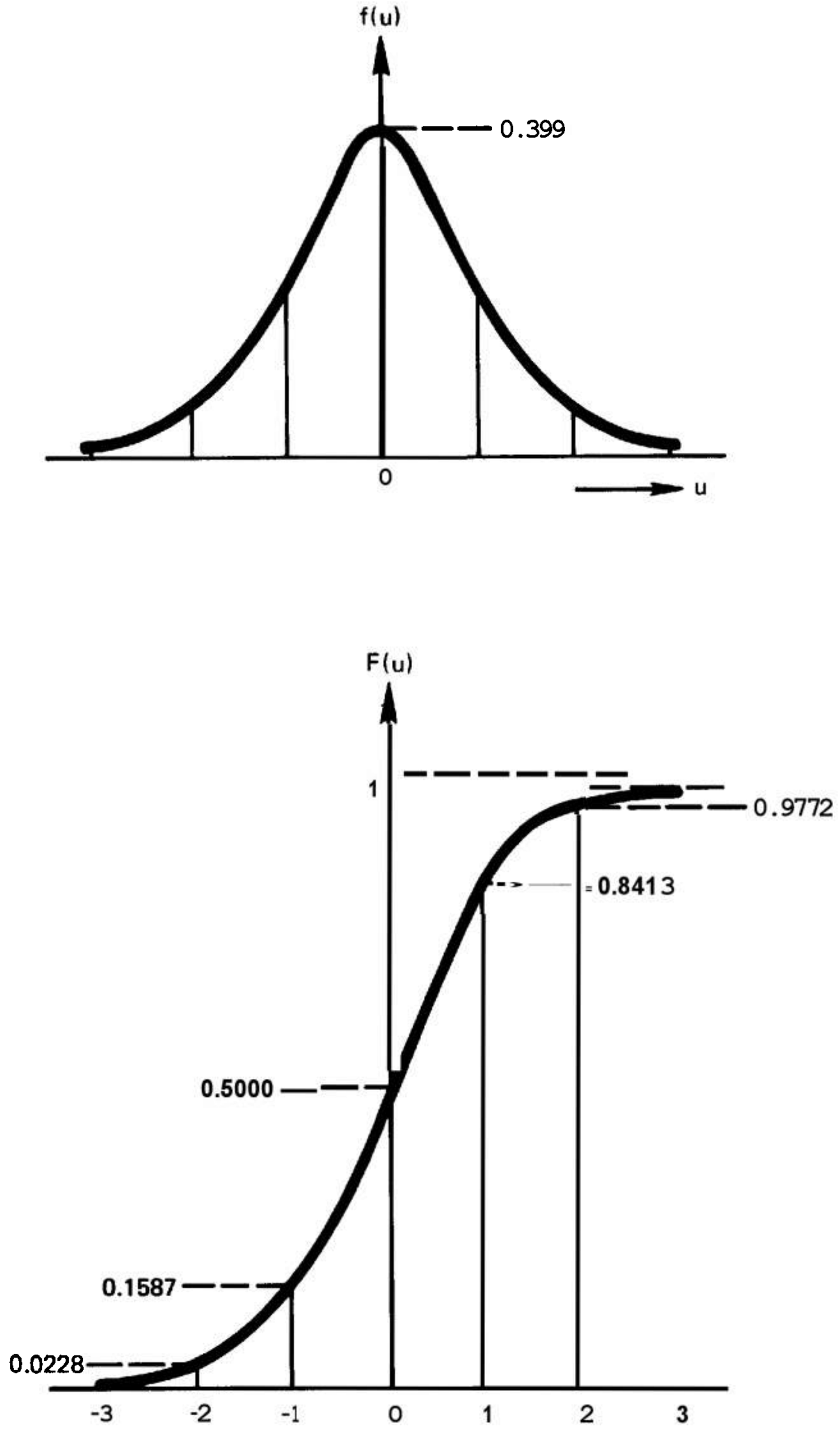


Figure 8-2. The Standard Normal Distribution $N(X|0, 1)$

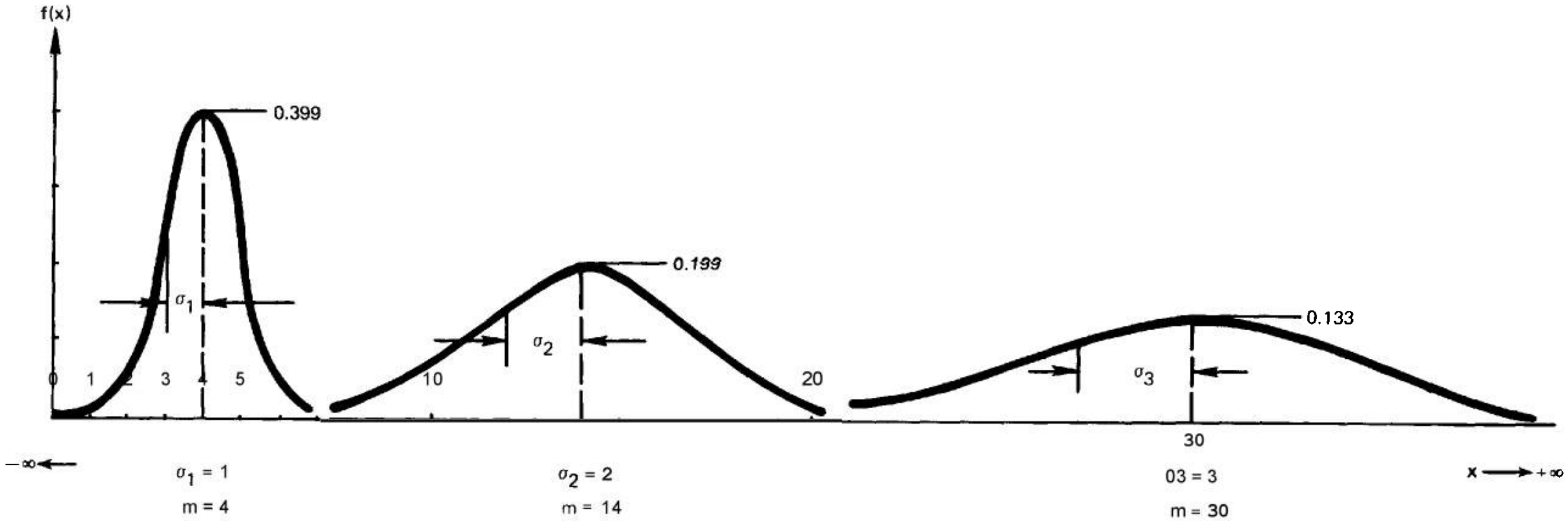


Figure 8-3. Shapes of Normal Densities for Different σ's

ables is again normally distributed with a mean the sum of their respective means, and a variance the sum of their respective variances. If the sum of two independent random variables (X, Y) is normally distributed, then so are X and Y, i.e., if $N(X + Y|m, \sigma^2)$ then if X, Y are independent,

$$N(X|m_1, \sigma_x^2)$$

and

$$N(Y|m_2, \sigma_y^2)$$

for some m_1, m_2 so that

$$m_1 + m_2 = m$$

and for some

$$\sigma_x^2, \sigma_y^2$$

so that

$$\sigma_x^2 + \sigma_y^2 = \sigma^2.$$

8-4.2 THE LOGNORMAL CLASS OF FUNCTIONS

The lognormal distribution frequently is applied in maintainability, since the physical laws of search and classification often obey the assumptions leading to the lognormal distribution (Refs. 10, 11, 12). That is, the time taken to find classified items (such as tool crib, a maintenance procedure in a service manual, or a document in a library) is more strongly dependent on the information structure of the classification system and less so on user characteristics. Since the fault location and correction processes are classification processes (e.g., classifying symptoms, responses to test signals, selecting tools, or selecting repair parts), they also will often be lognormally distributed. A definition of the lognormal distribution and descriptions of its properties follow:

1. Definition. A positive random variable Z is said to be lognormally distributed with parameters μ and v if $Y = \ln Z$ is normally distributed with *CDF* $N(\ln Z|\mu, \sigma^2)$, where μ is the mean and σ^2 is the variance of $Y = \ln Z$. The logarithm is the natural logarithm to the base $e \approx 2.718281828$.

The lognormal *pdf* is given by

$$f(z) = (\sqrt{2\pi}\sigma z)^{-1} \exp\{-(1/2)[(\ln z - \mu)/\sigma]^2\} \quad (8\text{-}45)$$

and the corresponding *CDF* is

$$F(z) = (\sqrt{2\pi}\sigma t)^{-1} \exp\{-(1/2)[(\ln t - \mu)/\sigma]^2\} dt \quad (8\text{-}46)$$

2. Notation. That a random variable Z is lognormally distributed with parameters μ, σ^2 is denoted, by convention (Ref. 13), as $\Lambda(Z|\mu, \sigma^2)$, so that $\Lambda(Z|\mu, \sigma^2) = N(\ln Z|\mu, \sigma^2)$. The corresponding *pdf* is denoted as $\lambda(z|\mu, \sigma^2) = n(\ln z|\mu, \sigma^2)$.

3. Mean, Variance, Median, and Mode. The mean and variance can be obtained easily once the kth moment of Z is known. This kth moment is given by

$$E[Z^k] = \int_{-\infty}^{\infty} e^{ky}(\sqrt{2\pi}\sigma)^{-1} \times \exp\{-(1/2)[(y - \mu)/\sigma]^2\} dy \quad (8\text{-}47)$$

where $y = \ln z$.

Integrating Eq. 8-47 (Ref. 13, Appendix A) gives

$$E[Z^k] = \exp(k\mu + k^2\sigma^2/z) \quad (8\text{-}48)$$

Thus, the first moment (mean) is

$$m = E[Z] = \exp(\mu + \sigma^2/2) \quad (8\text{-}49)$$

and the second moment (variance) is

$$\mathrm{VAR}[Z] = E[Z^2] - E^2[Z] = m^2[\exp(\sigma^2) - 1] \quad (8\text{-}50)$$

The median M and the mode m_0 are respectively:

$$M = \exp(\mu) \quad (8\text{-}51)$$

$$m_0 = \exp(\mu - \sigma^2) \quad (8\text{-}52)$$

4. Tables. The *CDF* of $\Lambda(Z|\mu, \sigma^2)$ can be read directly from the tables of the standard normal distribution:

$$\Lambda(Z|\mu, \sigma^2) = N(\ln Z|\mu, \sigma^2) = N[(\ln Z - \mu)/\sigma|0, 1] \quad (8\text{-}53)$$

5. Reproductive Property. If X is distributed as

$$\Lambda(X|\mu_1, \sigma_1^2)$$

and Y is distributed as

$$\Lambda(Y|\mu_2, \sigma_2^2)$$

then

$$Z = aX^bY^c \text{ is } \Lambda(Z\,|\ln a + b\mu_1 + c\mu_2, \sigma_1^2 + \sigma_2^2) \quad (8\text{-}54)$$

This follows since $\ln Z = \ln a + b(\ln X) + c(\ln Y)$.

6. Limiting Property. From the Central Limit Theorem (par. 8-8.2) and the reproductive property, when Z_i is a sequence of random variables and when the conditions for the Central Limit Theorem are satisfied by their logarithms, then the pdf of the product X_N = N Π 1 Z_i tends to the lognormal distribution.

7. Specification of the Lognormal Distribution. To uniquely determine the lognormal, any one of the following four combinations can be used.

a. Specify μ and σ. This is the direct method.

b. The mean μ and variance VAR[Z] of Z are specified. Then Eqs. 8-49 and 8-50 enable μ and σ^2 to be determined as

$$\sigma^2 = \ln[(m^2 + \text{VAR}[Z])/m^2] \quad (8\text{-}55)$$

$$\mu = \ln m - \ln[m^2 + \text{VAR}[Z]/m^2]/2 \quad (8\text{-}56)$$

c. The mean m and the 100pth percentile value z_p are specified. Employing the definition of the lognormal and transforming to the normal, we obtain

$$P(Z \leqslant z_p) = P[(\ln Z - \mu)/\sigma \leqslant (\ln z_p - \mu)/\sigma] = p \quad (8\text{-}57)$$

where

$$\ln z_p = \mu + a N_p$$

and N_p is the 100pth percentile for the standard normal random variable Y = ln Z.

Combining this with Eq. 8-49 and solving for σ gives

$$\sigma_1 = N_p + \sqrt{N_p^2 - 2\ln(z_p/m)} \quad (8\text{-}58)$$

$$\sigma_2 = N_p - \sqrt{N_p^2 - 2\ln(z_p/m)} \quad (8\text{-}59)$$

Once σ is obtained, μ *can* be oalculated by

$$\mu = \ln m - \sigma^2/2 \quad (8\text{-}60)$$

If m and z_p are to define a legitimate normal distribution, the two possible variances must be real and positive. First, the quantity within the square root must be non-negative. This requirement gives

$$z_p = m\exp(N_p/2) \quad (8\text{-}61)$$

To assure positivity of σ, two cases occur:

(1) Case I. p I 1/2

Then, $N_p \leq 0$ arid $z_p < m$

Subcase I.A. If p I 1/2 and $z_p < m$, then there is one unique lognormal *pdf* which satisfies these values, and

$$\sigma_1 = N_p + \sqrt{N_p^2 - 2\ln(z_p/2)} \quad (8\text{-}62)$$

Subcase *I.B.* If $N_p \geq M$, then no lognormal distribution can satisfy the requirements.

(2) Case 2. p > 1/2. Then $N_p > 0$.

Subcase *2.A.* If $z_p \geq m$, then one unique lognormal distribution exists that satisfies the requirements.

Subcase *2.B.*

$$m < z_p \leqslant m\exp(N_p^2/2) \quad (8\text{-}63)$$

Then two lognormal populations exist satisfying these values, and their standard deviations are given by Eqs. 8-58 and 8-59.

Subcase 2.C $z_p > m \exp(N_p^2/2)$.

Then no lognormal distribution can satisfy the requirements m and z_p simultaneously.

d. Two percentile values z_{p1} and z_{p2} for 100p_1 and 100 p_2 are given.

From Eq. 8-57:

$$\begin{aligned} \mu + \sigma N_1 &= \ln z_1 \\ \mu + \sigma N_2 &= \ln z_2 \end{aligned} \quad (8\text{-}64)$$

Thus,

$$\sigma = (N_2 - N_1)^{-1} \ln (z_2/z_1) \quad (8\text{-}65)$$

$$\mu = (N_2 - N_1)^{-1} \ln (z_1^{N_2}/z_2^{N_1}) \quad (8\text{-}66)$$

8. Three-parameter Lognormal. A random variable Z is said to have a three-parameter lognormal distribution when there exists a constant *a* so that $Y = Z - a$ is a lognormal variate, i.e., $\Lambda(Y - a|\mu,\sigma^2)$. Z can be regarded as a displaced two-parameter lognormal variable.

To explore how the lognormal distribution behaves, let us consider the three lognormal curves of Fig. 8-4, which may be thought of as representing the maintainability of three system designs, composed of line replaceable units for line maintenance. All three configurations have the same median maintenance time but vary greatly in the spread of duration of their maintenance actions.

All three curves have a median maintenance time or geometric mean M of 15 min. This was purposely so chosen. A median of 15 min corresponds to a mean of $\mu = \ln M = \ln 15 = 2.71$ log-minutes of the normal transform. The three curves have different standard deviations of the logarithms of t, i.e., $\sigma_1 = 0.1\mu = 0.271$ log-minutes, $\sigma_2 = 0.3\mu = 0.82$ log-minutes, and $\sigma_3 = 0.5\mu = 1.36$ log-minutes. The equations of the three density curves are:

$$f_1(t) = (\sigma_1 t\sqrt{2\pi})^{-1} \exp\{-(1/2)[(\ln t - \mu)/\sigma_1]^2\}$$
$$= (0.271\, t\sqrt{2\pi})^{-1} \exp\{-(1/2) \times [(\ln t - 2.71)/0.271]^2\} \quad (8\text{-}67)$$

$$f_2(t) = (\sigma_2 t\sqrt{2\pi})^{-1} \exp\{-(1/2)[(\ln t - \mu)/\sigma_2]^2\}$$
$$= (0.82\, t\sqrt{2\pi})^{-1} \exp\{-(1/2) \times [(\ln t - 2.71)/0.82]^2\} \quad (8\text{-}68)$$

$$f_3(t) = (\sigma_3 t\sqrt{2\pi})^{-1} \exp\{(-1/2)[(\ln t - \mu)/\sigma_3]^2\}$$
$$= (1.36\, t\sqrt{2\pi})^{-1} \exp\{(-1/2) \times [(\ln t - 2.71)/1.36]^2\} \quad (8\text{-}69)$$

From these equations the three density curves of Fig. 8-4(A) can be plotted by calculating $f(t)$ for various values of t, or by reading from normal tables the normalized densities (u or z) corresponding to various values of t, as shown later, and dividing z by σt to obtain $f(t)$ at t.

Fig. 8-4(B) shows the *CDF*, $F(t)$, corresponding to the three density curves. The ordinates of an $F(t)$ curve are found from normal tables as follows. As stated before, we write:

$$F(t) = \int_{-\infty}^{x=\ln t} (\sigma\sqrt{2\pi})^{-1} \exp\{-(1/2) [(x - \mu)/\sigma]^2\} dx \quad (8\text{-}70)$$

which is the same as

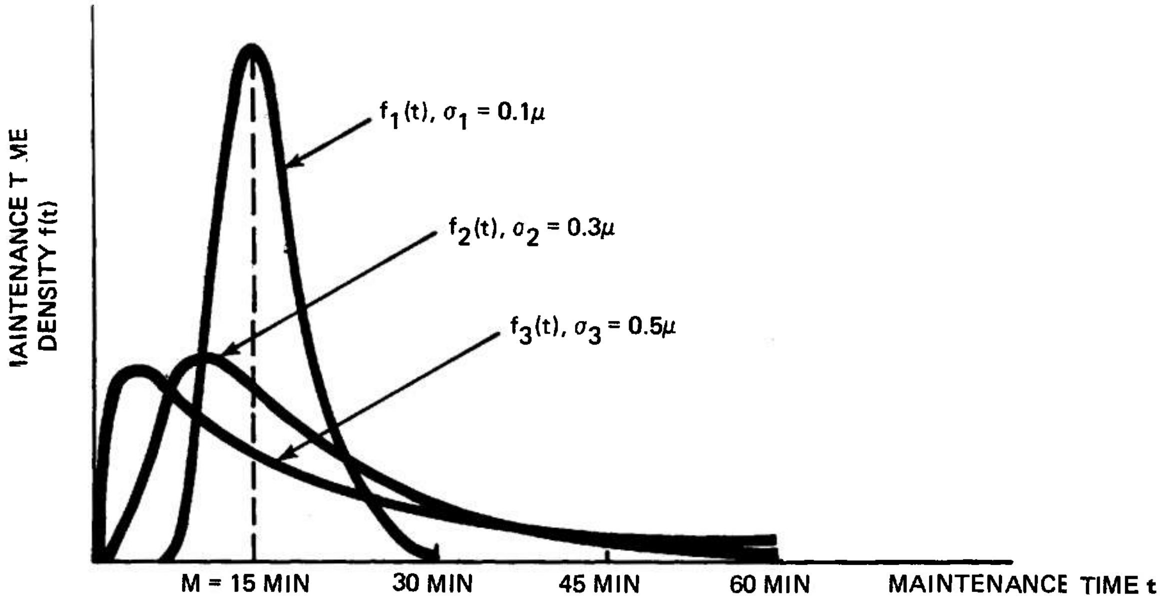
$$F(t) = \int_{-\infty}^{z} (2\pi)^{-1/2} \exp(-u^2/2)\, du \quad (8\text{-}71)$$

where

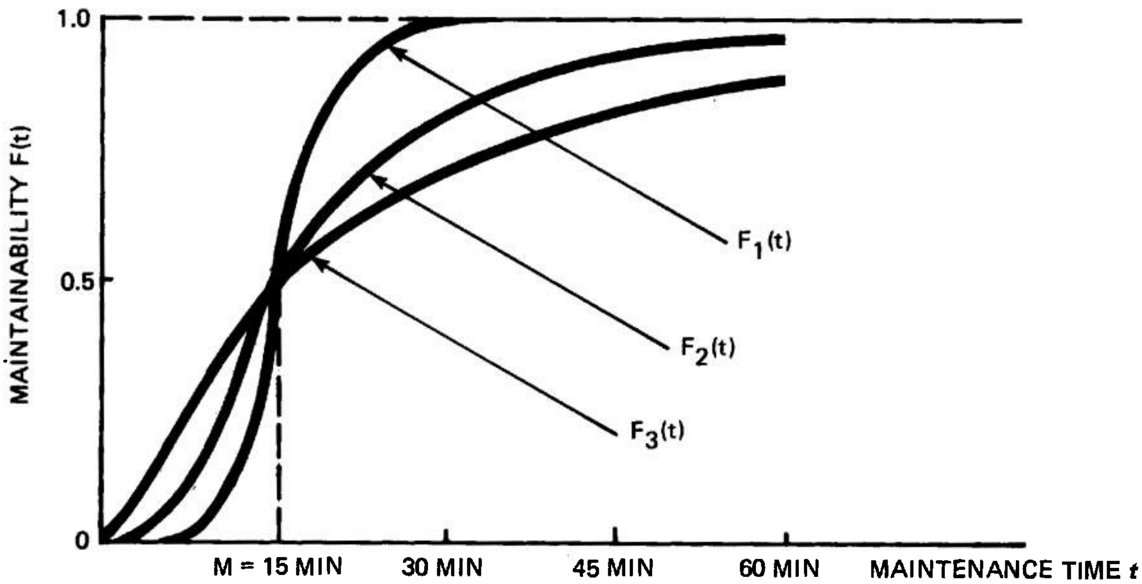
$$z = (\ln t - \mu)/\sigma = (x - \mu)/\sigma$$

Say we want a point on $F(t)$ corresponding to a definite value of t Knowing μ and σ, we calculate for the required t the numerical value of $z = (\ln t - \mu)/\sigma$ and look up in standardized normal tables the area from $-\infty$ to z. This gives us directly $F(t)$ at t. For instance, taking the curve $F_3(t)$, we find for $t = 60$ min, with $\sigma_3 = 1.36$ and $\mu = 2.71$, $z = (4.095 - 2.71)/1.36 = +1.02$. In Normal tables we find the area from $-\infty$ to $+1.02$ to be **0.85** or 85 percent, and this is $F_3(t)$ at $t = 60$ min.

It is interesting to note that with a median M of 15 min and if maintenance times are spread with $\sigma_3 = 1.36$ log-minutes, the probability of completing maintenance in 60 min is only 0.85. The reason is obviously the excessive skewness of $f_3(t)$. If we would like to find the 90 percent point of $F_3(t)$, usually referred to as M_{MAX}, we find in normal tables that an area of 0.9 corresponds to $z = +1.28$. From this we get directly



(A) DENSITY FUNCTIONS



(B) MAINTAINABILITY FUNCTIONS

Figure 84. Three Lognormal Distributions With the Same Median M = 15 min

$\ln M_{MAX} = z\sigma + \mu = (1.28)(1.36) + 2.71 = 4.45$ log-minutes and $M_{MAX} = 85.5$ min.

Now let us compute the mean maintenance times m, the modes m_0, the variances VAR(t), and the standard deviations $\sigma(t)$ of the three curves. Note here that $\sigma(t)$ is not the same $\sigma(\ln t)$ of the logarithms of maintenance time shown in Fig. 8-4(A) for the three curves to be 0.1μ, 0.3μ, and 0.5μ.

The means are obtained according to Eq. 8-49 as

$$m_1 = \exp(\mu + \sigma_1^2/2) = \exp(2.71 + 0.037) = 15.5 \text{ min}$$

$$m_2 = \exp(\mu + \sigma_2^2/2) = \exp(2.71 + 0.3362) = 21 \text{ min}$$

$$m_3 = \exp(\mu + \sigma_3^2/2) = \exp(2.71 + 0.925) = 37 \text{ min} \quad (8\text{-}72)$$

This is quite a spread in the mean maintenance times, considering that the median is only 15 min for all three curves.

The variances and modes follow from Eqs. 8-50 and 8-52. The modes are

$$m_{01} = \exp(2.71 - 0.073) = 14 \text{ min}$$

$$m_{02} = \exp(2.71 - 0.67) = 7.7 \text{ min} \quad (8\text{-}73)$$

$$m_{03} = \exp(2.71 - 1.85) = 2.4 \text{ min}$$

and the variance VAR(t) and their corresponding standard deviations $\sigma(t) = \sqrt{\text{VAR}(t)}$ are

$$\text{VAR}_1(t) = 19.2 \text{ min}^2, \quad \sigma_1(t) = 4.4 \text{ min}$$

$$\text{VAR}_2(t) = 420 \text{ min}^2, \quad \sigma_2(t) = 20.5 \text{ min} \quad (8\text{-}74)$$

$$\text{VAR}_3(t) = 7850 \text{ min}^2, \quad \sigma_3(t) = 89 \text{ min}$$

This again reflects the various degrees of skewness of the three curves.

The lognormal distribution, even though somewhat cumbersome to work with, can be handled satisfactorily with the help of normal tables. Design requirements in the lognormal case are usually given in terms of a desired median maintenance time M and a statement that an upper limit of maintenance time M_{MAX} shall not be exceeded with at least a given probability $F(M_{MAX})$. Here, M and M_{MAX} with a given $F(M_{MAX})$ fully define the parameters μ and σ of the lognormal distribution. The mean is $\mu = \ln M$ and the standard deviation is given by

$$(\ln M_{MAX} - \mu)/z_F = [\ln(M_{MAX}/M)]/z_F \quad (8\text{-}75)$$

where z_F is the value of the standard normal variable z at $F(M_{MAX})$, found in normal tables. For instance, when $F(M_{MAX}) = 0.9$ then $z_F = 1.28$, and when $F(M_{MAX}) = 0.95$ then $z_F = 1.65$. (See Table 8-1 for values of z - denoted by U.)

Assume the specification says $M = 20$ min, $M_{MAX} = 60$ min and $F(M_{MAX}) = 0.95$. Then $\ln(60/20) = \ln 3 = 1.10$, and since $z_F = 1.65$, $\sigma = 1.10/1.65 = 0.67$ logminutes and $\mu = \ln 20 = 3.00$ logminutes.

8-4.3 THE EXPONENTIAL CLASS OF FUNCTIONS

The exponential class of distribution functions is a large family which contains many commonly used specific distributions, (see Table 8-2). The reason for studying this family in general is that when one becomes familiar with the general theory, then the treatment of the many specific distributions in Table 8-2 becomes possible (see Ref. 14, Section 3.5).

A *pdf* $f(x, \theta)$ with only one parameter θ is said to belong to the exponential family if it can be written in the following form:

$$f(x, \theta) = B(\theta)\, h(x) \exp[Q(\theta) R(x)] \quad (8\text{-}76)$$

where B, h, Q, and R are arbitrary functions of the indicated index. The parameter θ may be p, λ, m, etc., as defined. A fairly extensive theory of statistical estimation has been worked out for *pdf*'s in the exponential class (Ref. 7).

8-4.3.1 The Gamma *Distribution*

As seen in Table 8-2, the gamma distribution belongs to the exponential class. The gamma distribution is very important and has the pdf of

$$f(x) = [\lambda^n/\Gamma(n)]\, x^{n-1} \exp(-x\lambda) \quad (8\text{-}77)$$

where λ is the scale parameter, n is the shape parameter, and $\Gamma(n)$ is the Gamma Function.

In general, the kth moments of the gamma distribution are

TABLE 8-2.
SOME DISTRIBUTIONS OF THE EXPONENTIAL FAMILY

Name	$f(x,\theta)$	$B(\theta)$	$Q(\theta)$	$R(x)$	$h(x)$	Normal Parameters
Binomial	$\binom{n}{x} P^{x} \left(1-P\right)^{n} -$ if n is considered fixed along with P	$\left(1-P\right)^{n}$	$\ln\left(\frac{P}{1-P}\right)$	x	$\binom{n}{x}$	P, n
Poisson	$\frac{m^{x}}{x!}\left[\exp(-m)\right]$	$\exp(-m)$	$\ln m$	x	$\frac{1}{x!}$	m
Normal* $N(0,\sigma)$	$\left(2\pi\sigma^{2}\right)^{-1/2} \exp\left[-x^{2}/\left(2\sigma^{2}\right)\right]$	$\left(2\pi\sigma^{2}\right)^{-1/2}$	$-\left(2\sigma^{2}\right)^{-1}$	x^{2}	1	σ
Gamma	$\lambda^{n} x^{n-1}\left[\frac{\exp(-x\lambda)}{\Gamma(n)}\right]$	$\frac{\lambda^{n}}{\Gamma(n)}$	$-\lambda$	x	x^{n-1}	λ, n
Exponential	$\lambda \exp(-\lambda x)$	λ	$-\lambda$	x	1	λ

*The Normal distribution belongs to the exponential class but, because of its great importance and extensive theory, is covered separately in the text (par. 8-4.1).

General form: $f(x,\theta)=B(\theta)\, h(x) \exp\left[Q(\theta)\, R(x)\right]$

$$E[X^k] = [\Gamma(n+k)]/[\lambda^k \Gamma(n)] \quad (8\text{-}78)$$

When n is an integer, this reduces to:

$$E[X^k] = (n+k-1)!/[(n-1)!\,\lambda^k] \quad (8\text{-}79)$$

Then the mean and variance are $m = n/\lambda$, and $\mathrm{VAR}[X] = n/\lambda^2$.

The gamma distribution satisfies the following addition theorem. If two gamma random variables X_1 and X_2 have *pdf's* with the same h, but possible different n parameters (say, n_1 and n_2), then their sum $Y = X_1 + X_2$ is again gamma distributed with parameters λ and $n_y = n_1 + n_2$ (Ref. 15). This theorem is useful in problems of summing maintenance times for a fixed number of repairs.

Some shapes of the gamma pdf's (Eq. **8-77**) are shown in Fig. **8-5**. Many tables are available. Ref. 16 gives standardized (h = 1) tables to nine decimal places for $n = -0.5(0.5)\,75(1)\,162$ at $\mu = 0.0(0.1)50$, where n and u are defined by the Incomplete Gamma Function I (u, n), given by

$$I(u,n) = [\Gamma(n)]^{-1} \int_0^{u\sqrt{n}} t^{n-1} \exp(-t)\,dt \quad (8\text{-}80)$$

A special case of the gamma distribution occurs when the shape parameter $n = 1$. This results in the exponential distribution $f(x) = \lambda \exp(-\lambda x)$, per Eq. **8-71**.

8-4.3.2 The Chi-square Distribution

The chi-square distribution is also a special case of the gamma distribution. A random variable Xis said to be chi-square (χ^2) distributed with ν degrees of freedom when it is gamma distributed with scale parameter $\lambda = 1/2$ and shape parameter $n = \nu/2$. The usefulness of the χ^2 distribution arises from the fact that when X is normally distributed according to $N(X|0, \sigma^2)$, then χ^2 has the chi-square distribution with one degree of freedom (Ref. 17). Also, the *CDF* of χ^2 follows the addition theorem property of the gamma distribution. These two facts explain its use in Goodness-of-fit Tests (par. **6-7.5**). Ref. **16** tabulates the χ^2 CDF and its percentile points.

8-4.3.3 Exponential Distribution

A random variable Xis said to be exponentially distributed if there is a number $b \geq 0$, so that,

$$P(X \leq x) = F(x) = \begin{cases} 1 - \exp[-\lambda(x-b)], & \text{if } x > b \\ 0, & \text{if } x \leq b \end{cases} \quad (8\text{-}81)$$

The *pdf* is

$$f(x) = \begin{cases} \lambda \exp[-\lambda(x-b)], & \text{if } x \geq b \\ 0, & \text{otherwise} \end{cases} \quad (8\text{-}82)$$

(Note: this is a gamma *pdf* with parameters λ and $n = 1$)

The moments around zero are

$$E[X^k] = \sum_{i=0}^{k} \binom{k}{i} b^{k-i} \lambda^i (i!) \quad (8\text{-}83)$$

and therefore

$$\mathrm{MEAN} = E[X] = 1/\lambda + b \quad (8\text{-}84)$$

$$\mathrm{VAR}[X] = E[X^2] - E^2[X] = 1/\lambda^2 \quad (8\text{-}85)$$

8-5 ESTIMATION

8-5.1 POINT ESTIMATION PROBLEMS IN GENERAL

Sampling is undertaken to estimate some parameters of the population, e.g., the mean and variance. These parameters are to be estimated from the values of a sample of size n. Suppose these are $x_1, x_2, \ldots, x_n$.

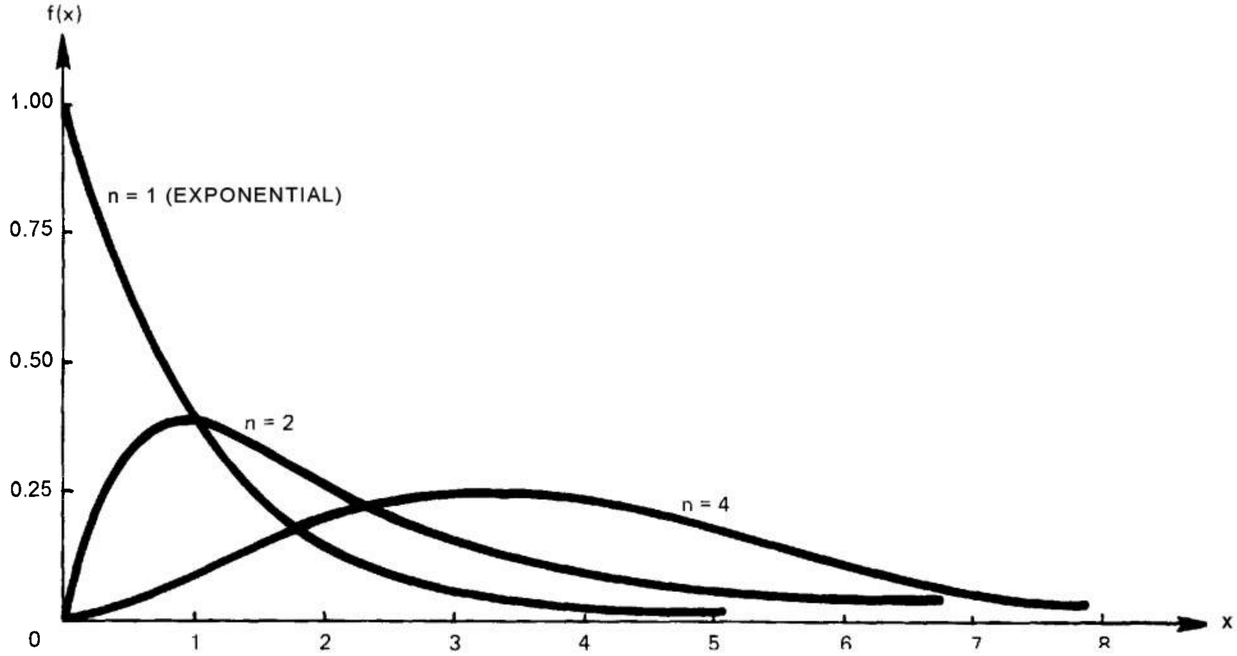


Figure 8-5. Shapes of Some Gamma *pdf's* for $\lambda = 1$

By definition, an ***estimator*** or ***statistic*** is a function $T(\vec{x})$ of the sample values, $\vec{x} = (x_1, x_2, \ldots, x_n)$, which is used for estimating some quantity T of interest.

As an example

$$\bar{x} = \sum_{i=1}^{n} x_i/n \tag{8-86}$$

is an estimator of the population mean and is called the sample average. There are several useful properties which any estimator $T(\vec{x})$ should have, if possible. These are:

1. An estimator is said to be ***unbiased*** if

$$E[T(\vec{x})] = T \tag{8-87}$$

2. An estimator is said to be ***consistent*** if for any $a > 0$

$$\lim_{n\to\infty} P[\,|T - T(\vec{x}| > a] \tag{8-88}$$

3. An estimator is said to be ***sufficient*** if it utilizes all the information given by the sample values $\vec{x} = (x_1, \ldots, x_n)$.

Note that if $\lim n \to \infty$ VAR$[T(\vec{x})] = 0$, then $T(\vec{x})$ also is consistent.

4. When $T_1(\vec{x})$ and $T_2(\vec{x})$ are two unbiased estimators, then the ***relative efficiency*** of $T_1(\vec{x})$ with respect to $T_2(\vec{x})$ is the ratio

$$[\text{VAR}\, T_2(\vec{x})]/[\text{VAR}\, T_1(\vec{x})] \tag{8-89}$$

5. An estimator $T(\vec{x})$ is said to be a ***best unbiased*** estimator or a ***most efficient*** estimator when

$$\text{VAR}[T(\vec{x})] \leqslant \text{VAR}[T_1(\vec{x})] \tag{8-90}$$

for any other unbiased estimator.

These properties are important because our estimate is most likely to be close to being correct when the estimator has these properties. Thus for large enough samples, we are practically assured of being very close to the true value if we are using a consistent estimator. Sufficiency is important from a practical point of view. For example $[1/(n-2)] \sum_{i=3}^{n} x_i$ is not sufficient because it neglects the first two observations and thus loses accuracy. When T_1, T_2 are consistent estimators and the relative efficiency of T_1 with respect to T_2 is greater than 1, then T_1 is preferred to T_2. Thus the notation of relative efficiency is useful in choosing between estimators.

8-5.1.1 Maximum Likelihood Estimates

If the true value of the quantity to be estimated by the estimator $T(\vec{x})$ is T, then some vectors $\vec{x}$ will be more likely to occur than others. In fact if $f(x|T)$ is the pdf of $\vec{x}$ given T, then the pdf of x is

$$g(\vec{x}\,|T) = f(x_1|T)f(x_2|T)\cdots f_n(x|T) \tag{8-91}$$

The maximum likelihood method finds that value of T which maximizes $g(\vec{x}|T)$ for a given $\vec{x}$, i.e., it solves

$$\frac{d[g(\vec{x}\,|\,T)]}{dT} = 0 \tag{8-92}$$

to obtain the estimator $T(x)$. Maximum likelihood estimators are most efficient in the limit as n approaches infinity. They are also sufficient estimators whenever such exist.

Further, if $T(\vec{x})$ is a maximum likelihood estimate for T, and if $h(T)$ is a function with a single-valued inverse, then $h(T(\vec{x}))$ is a maximum likelihood estimate for $h(T)$.

8-5.1.2 Method of Moments

The kth sample moment is defined by

$$\bar{m}_k = \sum_{i=1}^{n} x_i^k/n \tag{8-93}$$

which can be used as an estimator for $E[X^k]$ the kth moment of the random variable X.

8-5.2 INTERVAL ESTIMATES

The estimators discussed in the preceding paragraphs give only point estimates of the quantities they are estimating. Naturally, the smaller the variance of a point estimator, the greater its "precision". The way to make this notion of "greater precision" better understood is to use an interval estimator in such a way that the estimated interval covers the largest amount of the probability mass concerning the point estimator.

Suppose first there is a variable $T(\vec{x}, \theta)$ which is a function of the sample values and of a parameter *8.* $T(\vec{x}, \theta)$ is a random variable and has a *CDF.* If $F(T)$ is its *CDF,* then given any number $0 \leq a \leq 1$, we can find numbers T_1 and T_2 so that

$$P[T_1 \leqslant T(\vec{x}, \theta) \leqslant T_2] = F(T_2) - F(T_1) = 1 - \alpha \quad (8\text{-}94)$$

Suppose that the inequality for $T(\vec{x}, 8)$ *can* be solved for *8* so that Eq. 8-94 can be written as

$$P(\theta_1 \leqslant \theta \leqslant \theta_2) = F(\theta_2) - F(\theta_1) = 1 - \alpha \quad (8\text{-}95)$$

This equation provides a way for an interval estimate which gives the probability that the true value of *8* is in the interval (θ_1, θ_2). When *8,* and *8,* are determined from the random variables T_1 and T_2 of Eq. 8-94 so that

$$P(\theta_1 \leqslant \theta \leqslant \theta_2) = 1 - \alpha \quad (8\text{-}96)$$

then the random interval (θ_1, θ_2) is called a two-sided $(1 - a)$ confidence interval.

A $(1 - a)$ confidence interval is said to be upper one-sided if $8, = -\infty$, and lower one-sided if $\theta, = +\infty$.

One-sided intervals are sometimes more natural than two-sided ones. For example, we are not concerned with how low the MTTR is but want to be sure that it is not too high. Thus we would want to find *8,* so that $P(\theta \leq \theta_2) = 1 - a$.

A detailed discussion of confidence intervals is given in par. 6-8. Numerical examples are presented in par. 8-7.

SECTION II

PRACTICAL APPLICATIONS

8-6 POINT ESTIMATORS FOR SPECIFIC DISTRIBUTIONS

Although point estimates should not be used when interval estimates are possible, point estimators are necessary because point estimates can serve as best guidelines to selecting meaningful intervals. This paragraph is thus an exposition of obtaining point estimates. The terminology introduced in par. 8-5.1 is used whenever appropriate.

8-6.1 ESTIMATING FROM NORMAL POPULATIONS

1. *Estimating the Mean.* The sample mean is

$$\bar{x} = \sum_{i=1}^{n} x_i/n \tag{8-97}$$

This is the maximum likelihood estimate of the true mean μ, and is unbiased.

When the sample data are censored (lost or otherwise unavailable) estimation can be based on a central block of order statistics.

If the values $x_1, x_2, \ldots, x_n$ from an n-sized sample are ordered so that $x'_1 \leq x'_2 \leq x'_n$, then for any integer j, $0 \leq j \leq n/2$, the sample points,

$$x'_{j+1}, x'_{j+2}, \ldots, x'_{n-j}$$

are referred to as a central block of order statistics. For any central block of order statistics, the jth Winsorized mean is defined as

$$m_j = \left[jx'_{j+1} + \sum_{i=j+1}^{n-j} x'_i + jx'_{n-j} \right] \Big/ n \tag{8-98}$$

The efficiency of the Winsorized mean is very good and never falls below 99.9% for $n \leq 20$ when taken with respect to the best linear unbiased estimator from the same data points (Ref. 18).

The sample median $\bar{a}$ is a central block when n is odd. The efficiency of the median relative to the sample mean $\bar{x}$ is 63.7%. Its variance approximately equals (Ref. 19):

$$\mathrm{VAR}[\bar{a}] \approx \pi\sigma^2/(2n) \tag{8-99}$$

2. *Estimating the Standard Deviation*

If the true mean μ is not known, then the maximum likelihood estimate s of the standard deviation σ (corrected for bias) is (Ref. 19),

$$s = a_n \sqrt{\sum_{j=1}^{n} (x_j - \bar{x})^2/n} \tag{8-100}$$

where for $n > 10$,

$$a_n \approx 1 + (3/4)(n - 1)^{-1} \tag{8-101}$$

For $n \leq 10$ the values of n are given in Table 8-3.

The variance of the estimate of s is

$$\mathrm{VAR}[s] = [a_n^2(1 - 1/n) - 1]\sigma^2 \tag{8-102}$$

And if the true mean μ is known, which is almost never the case, the estimate s' of the standard deviation is

$$s' = b_n \sqrt{\sum_{j=1}^{n} (x_j - \mu)^2/n} = b_n \sqrt{s^2 + (\bar{x} - \mu)^2} \tag{8-103}$$

where s is as before and b_n is obtained as follows:

a. For $n > 10$, b_n is approximately

$$b_n \approx 1 + (n - 1)^{-1}/4 \tag{8-104}$$

TABLE 8-3.
VALUES OF a_n

n	a_n
2	1.77245
3	1.38198
4	1.25331
5	1.18942
6	1.15124
7	1.12587
8	1.10778
9	1.09428
10	1.08372

TABLE 8-4.
VALUES OF b_n

n	b_n
2	1.25331
3	1.12838
4	1.08540
5	1.06385
6	1.05094
7	1.04235
8	1.03624
9	1.03166
10	1.02811

b. For $n \leq 10$, the values of b_n are given in Table 8-4.

When the data are censored, then Dixon (Ref. 20) gives easy to use estimators for σ.

3. ***Estimating the Variance.*** An unbiased estimator for σ^2 is

$$S^2 = [n/(n-1)](s/a_n)^2 = (n-1)^{-1}\sum_{j=1}^{n}(x_j - \bar{x})^2 \qquad (8\text{-}105)$$

where s is given by *Eq.* 8-100, and a_n by *Eq.* 8-101, or Table 8-3. The variance of S^2 is then (Ref. 19)

$$\mathrm{VAR}[S^2] = 2\sigma^4/(n-1) \qquad (8\text{-}106)$$

4. ***Estimating the Percentiles.*** To estimate the 100*p*th percentile N_p of X, note that

$$p = P(X \leq N_p) = (2\pi)^{-1/2}\int_{-\infty}^{U_p} \exp(-u^2/2)\,du \qquad (8\text{-}107)$$

where

$$U_p = (N_p - \mu)/\sigma \qquad (8\text{-}108)$$

so that

$$N_p = \mu + \sigma U_p \qquad (8\text{-}109)$$

where U_p is the 100*p*th percentile of the standard normal.

N_p can be estimated from Eq. 8-109 by using estimators $\bar{x}$ and s for μ and σ, respectively. In particular, one can use

$$\bar{N}_p = \bar{x} + sU_p \qquad (8\text{-}110)$$

Then $\bar{x}$ and s are independent, and the variance of $\bar{N}_p$ is

$$\mathrm{VAR}[\bar{N}_p] = \mathrm{VAR}[\bar{x}] + U_p^2 \mathrm{VAR}[s] \qquad (8\text{-}111)$$

5. ***Estimating P[X ≤ x].*** Using par. 8-5.1.1, we see that the maximum likelihood estimator of $P[X \le x]$ is

$$\overline{P[X \le x]} = (2\pi)^{-1/2} \int_{-\infty}^{Z} \exp(-u^2/2)\,du \qquad (8\text{-}112)$$

where

$$Z = (x - \bar{x})(s/a_n)^{-1} \qquad (8\text{-}113)$$

8-6.2 ESTIMATION FROM LOGNORMAL POPULATIONS

The notation used in this paragraph is that introduced previously in par. 8-4.2. Specifically, a random variable Z is said to be lognormally distributed, when $Y = \ln Z$ is normally distributed. The mean of the distribution of Z is denoted by m and the variance by VAR(Z). The corresponding estimators are denoted by $\bar{m}$ and by $\overline{\mathrm{VAR}(Z)} = \bar{V}$. The mean of the distribution of $Y = \ln Z$ is denoted by μ and the variance by σ^2. Frequently, $\bar{\mu}$ and $\bar{\sigma}$ are referred to as "mean of the logs" and "standard deviation of the logs". Their estimates are denoted by $\bar{\mu}$ and $\bar{\sigma}$ or s.

Assume that the sample values $z_1, z_2, \ldots, z_n$ are drawn independently from a lognormal population. From this sample one wants to get estimates of m, VAR(Z), and any percentiles. First one must obtain estimates of the parameters μ and σ.

1. ***Estimating μ and σ***

Since $\mu = E[\ln Z]$, and $\ln Z$ is normally distributed, the maximum likelihood estimate for μ is

$$\bar{\mu} = \sum_{i=1}^{n} \ln z_i / n \qquad (8\text{-}114)$$

Similarly, an unbiased estimator for σ is

$$s = \sqrt{\sum_{i=1}^{n} (\ln z_i - \bar{\mu})^2/(n-1)} \qquad (8\text{-}115)$$

Having obtained these two estimates one now proceeds as follows.

2. ***Estimating the Population Mean***

The maximum likelihood estimate of the mean $\bar{m} = \exp(\mu + \sigma^2/2)$ of the lognormal distribution (see Eq. 8-49) is given by

$$\bar{m} = \exp(\bar{\mu} + s^2/2) \qquad (8\text{-}116)$$

Since this estimate is biased, a better, minimum variance unbiased estimate of $\bar{m}$ is

$$\bar{m} = e^{\bar{\mu}} \psi_n(s^2/2) \qquad (8\text{-}117)$$

where (Ref. 12, p. 45 and Ref. 13, p. 2-4)

$$\psi_n(t) = 1 + \frac{(n-1)t}{n} + \frac{(n-1)^3 t^2}{n^2(n+1)2!} + \frac{(n-1)^5 t^3}{n^3(n+1)(n+3)3!} + \ldots \qquad (8\text{-}118)$$

Asymptotically, $\psi_n(t)$ converges to (Ref. 12, p. 46):

$$\psi_n(t) = e^t[1 - t(t+1)/n + t^2(3t^2 + 22t + 21)/(6n^2)] + O(1/n^3) \qquad (8\text{-}119)$$

Tables of $\psi_n(t)$ are given in Appendix A of Ref. 12. Also Ref. 13 (Section 3) describes computer programs for calculating $\psi_n(t)$ and various lognormal parameters.

3. ***Estimating the Population Variance***

The maximum likelihood estimate $\bar{V}$ of VAR[Z] is

$$\bar{V} = \exp(2\bar{\mu} + s^2)[\exp(s^2) - 1] \qquad (8\text{-}120)$$

The minimum variance unbiased estimator of VAR[Z] is

$$\bar{V} = \chi_n(s^2)\exp(2\bar{\mu}) \tag{8-121}$$

where

$$\chi_n(t) = \psi_n(2t) - \psi_n[t(n-2)/(n-1)] \tag{8-122}$$

Tables of $\chi_n(t)$ are given in Appendix A, Table A3, of Ref. 12. Ref. 13 describes a computer program for calculating $\chi_n(t)$.

4. *Estimating the pth Percentile.* It follows from the fact that $Y = \ln Z$, that the pth percentile Z_p of the lognormal *CDF* is related to the pth percentile N_p of the normal *CDF* by

$$\bar{Z}_p = \exp(\bar{\mu} + sN_p) \tag{8-123}$$

This is the maximum likelihood estimator and, for large sample sizes, it approaches unbiasedness and normality.

8-6.3 POINT ESTIMATES OF THE EXPONENTIAL DISTRIBUTION

For the exponential distribution with a CDF of $F(x) = 1 - \exp(-\lambda x)$, the maximum likelihood estimator for the mean $\bar{m} = 1/\lambda$, obtained from an uncensored sample $x_1, x_2, \ldots, x_n$ is

$$\bar{m} = \sum_{i=1}^{n} x_i/n \tag{8-124}$$

This also means that the maximum likelihood estimators for the parameter $\bar{\lambda}$ and for the CDF of $\bar{F}(x)$ are

$$\bar{\lambda} = 1/\bar{m} \tag{8-125}$$

$$\bar{F}(x) = 1 - \exp(-\bar{\lambda}x) \tag{8-126}$$

For estimating the mean from censored samples see par. 6-8.3.

8-7 CONFIDENCE INTERVALS FOR SPECIFIC DISTRIBUTIONS

Par. 6-8 contains a detailed discussion of confidence intervals for the parameters of the normal, lognormal, and exponential distributions. Here we give a tabulation of various equations of confidence intervals for $1-\alpha$ confidence.

8-7.1 MEAN OF NORMAL DISTRIBUTION

1. ***When σ is known:***

a. Two-sided Confidence Interval:

$$\bar{m} - K_{\alpha/2}\sigma/\sqrt{n} \le m \le \bar{m} + K_{\alpha/2}\sigma/\sqrt{n} \tag{8-127}$$

b. One-sided Upper Confidence Interval:

$$m \le \bar{m} + K_\alpha\sigma/\sqrt{n} \tag{8-128}$$

c. One-sided Lower Confidence Interval:

$$m \ge \bar{m} - K_\alpha\sigma/\sqrt{n} \tag{8-129}$$

2. ***When σ is unknown:***

a. Two-sided Confidence Interval:

$$\bar{m} - (t_{\alpha/2;n-1})s/\sqrt{n} \le m \le \bar{m} + (t_{\alpha/2;n-1})s/\sqrt{n} \tag{8-130}$$

b. One-sided Upper Confidence Interval:

$$m \le \bar{m} + (t_{\alpha;n-1})s/\sqrt{n} \tag{8-131}$$

c. One-sided Lower Confidence Interval:

$$m \ge \bar{m} - (t_{\alpha;n-1})s/\sqrt{n} \tag{8-132}$$

In Eqs. 8-127 through 8-132, $K_{\alpha/2}$ is the

$100(1 - \alpha/2)$ percentile and K_α is the $100(100 - a)$ percentile of the standard normal distribution; s is the estimate of the standard deviation from a sample size n; $t_{\alpha/2;\,n-1}$ is the $100(1 - \alpha/2)$ percentile; and $t_{\alpha,\,n-1}$ is the $100(1 - a)$ percentile of Student's t distribution (Ref. 2, p. **201**) with $n - 1$ degrees of freedom. Fig. **8-6** illustrates the meaning of a two-sided **(1** $-$ a) confidence interval, while Fig. **8-7** illustrates a one-sided upper (1 $-$ a) confidence interval for the mean m. Both figures depict the standard normal distribution. Par. **8-7.4** provides practical applicatons.

8-7.2 STANDARD DEVIATION OF THE NORMAL DISTRIBUTION

1. Two-sided Confidence Interval for σ:

$$s\sqrt{(n-1)/(\chi^2_{\alpha/2;n-1})} \leq \sigma \leq \sqrt{(n-1)/(\chi^2_{1-\alpha/2;n-1})} \quad (8\text{-}133)$$

where $\chi^2_{\alpha/2;\,n-1}$ $\chi^2_{1-\alpha/2;\,n-1}$ $100(\alpha/2)$ and $100(1 - \alpha/2)$ percentiles of the chi-square distribution with n $-$ 1 degress of freedom.

2. One-sided Upper and Lower Confidence Intervals:

$$\sigma \leq s\sqrt{(n-1)/\chi^2_{1-\alpha;n-1})} \quad (8\text{-}134)$$

$$\sigma \geq s\sqrt{(n-1)/(\chi^2_{\alpha;n-1})} \quad (8\text{-}135)$$

8-7.3 THE LOGNORMAL DISTRIBUTION

The procedure is identical with that for the normal distribution when computing confidence intervals for the parameters μ and σ, i.e., the mean and the standard deviation of the random variable $Y = \ln Z$ which is normally distributed.

1. Two-sided Confidence Interval for $\bar{\mu}$:

$$\bar{\mu} - K_{\alpha/2}s/\sqrt{n} \leq \mu \leq \bar{\mu} + K_{\alpha/2}s/\sqrt{n} \quad (8\text{-}136)$$

where s is an estimate of σ obtained from a sample size n and $K_{\alpha/2}$ is the **100(1** $-$ **α/2)** percentile of the standard normal distribution. This equation is identical with **Eq. 8-127,** except that μ, $\bar{\mu}$, and s are log values. One-sided intervals for μ are computed in the same way as computed for **m** in par. **8-7.1.**

2. Two-sided Confidence Interval for σ:

$$s\sqrt{(n-1)/(\chi^2_{\alpha/2,n-1})} \leq \sigma \leq s\sqrt{(n-1)/(\chi^2_{1-\alpha/2;n-1})} \quad (8\text{-}137)$$

This equation is identical with Eq. ,**8-133,** except that σ and s are log values here. The one-sided confidence intervals are computed as in par. **8-7.2.**

3. Confidence Intervals for the Mean m:

The mean of the lognormal distribution is, by Eq. **8-49,**

$$m = \exp(\mu + \sigma^2/2) \quad (8\text{-}138)$$

The two-sided confidence interval on m, when σ is known, may be approximated by

$$\exp(\bar{\mu} - K_{\alpha/2}\sigma/\sqrt{n} \leq m \leq \exp(\bar{\mu} + K_{\alpha/2}\sigma/\sqrt{n}) \quad (8\text{-}139)$$

and when σ is unknown and only an estimate s of σ exists,

$$\exp[\bar{\mu} - (t_{\alpha/2;n-1})s/\sqrt{n}] \leq m \leq \exp[\bar{\mu} + (t_{\alpha/2;n-1})s/\sqrt{n}] \quad (8\text{-}140)$$

where $t_{\alpha/2;n-1}$ is the **100(1** $-$ **α/2)** percentile of Student's t distribution.

It should be stated, however, that no exact confidence intervals can be obtained either for m or for VAR[Z], as stated in Ref. **12,** p. 50.

8-7.4 EXAMPLES

1. ***Example No. 1***:

Assume that from a sample of $n = 100$ from a normal population, the estimate of the mean $\bar{m}$ was computed to be $\bar{m} = 2$ hr, and the standard deviation σ is known to be $\sigma = 1$ hr. Required are the two-sided and the upper one-sided confidence intervals for $1 - a = 0.95$, **i.e.,** $a = 0.05$, $\alpha/2 = 0.025$, $1 - \alpha/2 = 0.975$.

First one reads from standard normal tables (see Table **8-1** where K is denoted by U) the values of $K_{\alpha/2} = K_{0.025} = 1.96$, which correspond to the $100(1-\alpha/2) = 97.5$ percentile, and $K_\alpha = K_{0.05} = 1.645$ which correspond to the $100(1-\alpha/2)' = 95.0$ percentile. Then, using **Eq. 8-127,** one computes the two-sided confidence interval as

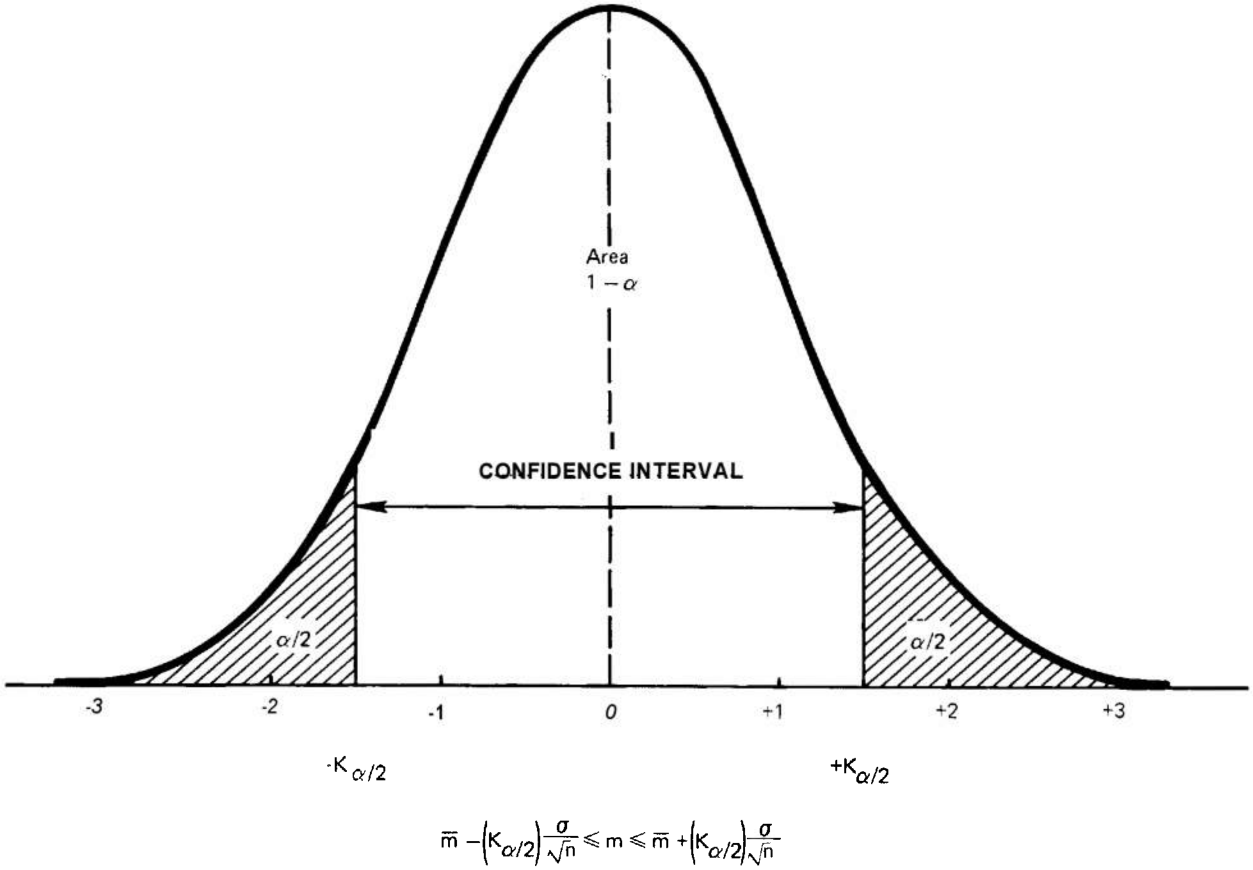


Figure 8-6. Two-sided 1 — a Confidence Interval

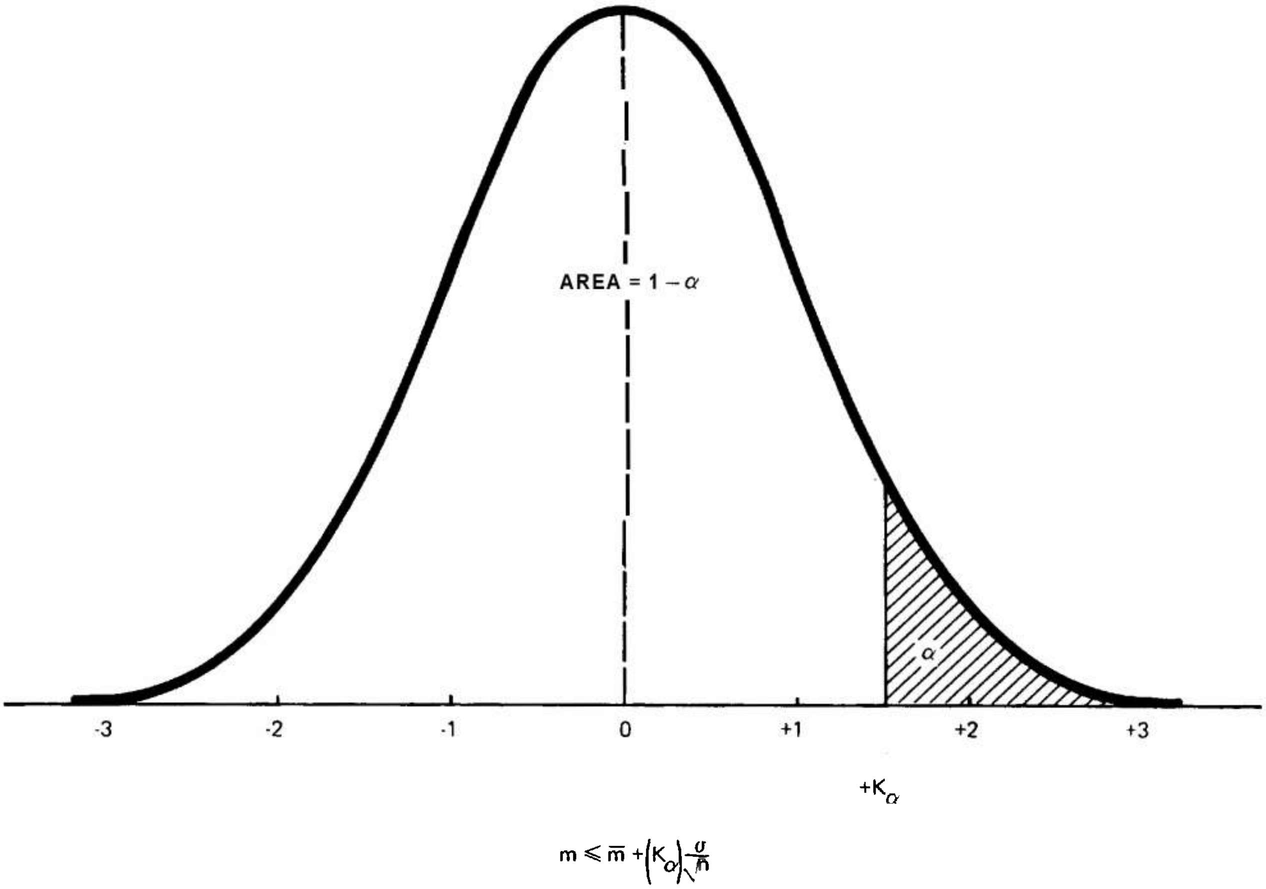


Figure 8-7. One-sided Upper 1 − a Confidence Interval

$$[2 - 1.96(1/10)] \leq m \leq [2 + 1.96(1/10)]$$

$$1.804 \leq m \leq 2.196 \qquad (8\text{-}141)$$

The statement then reads "There is a $(1 - \alpha) = 0.95$ probability that the calculated limits of 1.804 hr and 2.196 hr will include the true mean m". And using Eq. 8-128, one computes the one-sided upper confidence interval as

$$m \leq 2 + 1.645(1/10) = 2.1645 \qquad (8\text{-}142)$$

which reads: There is a $(1 - \alpha) = 0.95$ probability that the true m is less than or equal to 2.165 hr. Of course, in this example m could be the MTTR of a system.

2. ***Example No. 2:*** Continuing with Example No. 1, we assume that the standard deviation σ is not known, but an estimate of $s = 1$ hr was computed from the sample of $n = 100$. What are the confidence intervals? Using Eqs. 8-130 and 8-131, we first find the approximate values of $t_{\alpha/2;n-1} = t_{0.025,99} = 1.971$ and $t_{\alpha,n-1} = t_{0.05,99} = 1.664$, from such tables as in Ref. 3, Table A-4. Then we compute

$$[2 - 1.97(1/10)] \leq m \leq [2 + 1.97(1/10)]$$

$$1.803 \leq m \leq 2.197 \qquad (8\text{-}143)$$

and

$$m \leq 2 + 1.664(1/10) = 2.166 \qquad (8\text{-}144)$$

These results are almost identical with the previous ones except that the intervals are a little larger because of the use of percentiles of the t distribution. The difference would be significantly larger for small sample sizes n.

3. ***Example No. 3:*** For the same case as in Example No. 1, let us compute the one-sided upper confidence interval for σ when the estimate from the sample of $n = 100$ is $s = 1$ hr. From tables (Ref. 3, Table A-3) we find the approximate value of the chi-square percentile $\chi^2_{1-\alpha} = \chi^2_{0.95}$ for $n - 1 = 99$ degrees of freedom. This value is 77.9 and using Eq. 8-134, one obtains

$$\sigma \leq 1 \times \sqrt{99/77.9} = 1.13 \qquad (8\text{-}145)$$

This reads, that with 95% confidence the true standard deviation is less than or equal to 1.13 hr. Attention is drawn to the fact that some chi-square tables give the $(1 - \alpha)$ percentage points while others give the α percentage points.

Computations of the confidence intervals for the log-normal distribution are done in the same way as in the preceding normal examples, except that the log values must be used.

8-8 LIMIT THEOREMS

Limit theorems are widely used in statistics because they give the approximate behavior of large samples. In fact, they form the basis of many statistical procedures. They serve a practical purpose because limiting behavior is much simpler than exact analysis; the use of limit theorems simplifies any requisite analysis. As an introduction to the Central Limit Theorem, the Law of Large Numbers will be discussed.

8-8.1 LAW OF LARGE NUMBERS

This states that if the mean or expected value m of a random variable is finite, then the probability that the arithmetic average of the sample will differ from m approaches zero as n approaches infinity, or for any $\epsilon > 0$

$$\lim_{n \to \infty} \left(\left| m - \sum_{i=1}^{n} X_i/n \right| > \epsilon \right) = 0 \qquad (8\text{-}146)$$

where n is the number of observations in the sample, and X_i is the value of the ith observation. This law is of use in determining appropriate point estimates. For example, besides assuring us that

$$m = \sum_{i=1}^{n} X_i/n$$

is a good estimator for the mean, it also provides us with a good estimator for the probability of an event. Thus, let A be an event and consider a sequence of independent trials such athat $P(A)$ is the same in each trial. In order to get a probability estimate from

$$\sum_{i=1}^{n} X_i/n, \text{ let}$$

$$X_i = 0, \text{ if A does not occur at the } i\text{th trial}$$
$$= 1, \text{ otherwise}$$

then

$$\bar{P} = \sum_{i=1}^{n} X_i/n = \frac{\text{number of successes}}{\text{number of trials}} \quad (8\text{-}147)$$

where $\bar{P}$ is a point estimate for $P(A)$. Now the law of large numbers tells us that

$$\lim_{n\to\infty} P[\bar{P} \neq P(A)] = 0 \quad (8\text{-}148)$$

The reason $P(A)$ can be substituted for m is that,

$$m = E[X] = 1[P(A)] + 0[1 - P(A)] = P(A) \quad (8\text{-}149)$$

As an example, suppose we want to estimate the probability that a maintenance action will be completed within a hours, i.e., we want to estimate the probability $P(t \leq a)$. Then, by defining X_i as in the preceding paragraph, we know that a good point estimate of $P(t \leqslant a)$ is the mean,

$$\sum_{i=1}^{n} X_i/n$$

Since a is arbitrary, this really says that the cumulative histogram approaches the true distribution $F(a) = P(t \leqslant a)$ of maintenance time.

To answer the question of how good such point estimates are, we need another limit theorem. This limit theorem really follows directly from the more fundamental result that

$$P(|X - m| \geqslant k) < \sigma^2/k^2 \quad (8\text{-}150)$$

where m is the mean, σ the standard deviation, and $k > 0$. Note that the preceding estimates the probability in the tails $(-\infty, m - k)$ and $(m + k, \infty)$. Thus it also estimates the probability concentrated symmetrically about the mean m, i.e.,

$$P[X \in (m - k, m + k)] > 1 - \sigma^2/k^2 \quad (8\text{-}151)$$

From this follows

$$P\left[\left|\left(\sum_{i=1}^{n} X_i/n\right) - m\right| \geqslant a\right] \leqslant \sigma^2/(na) \quad (8\text{-}152)$$

for any $a > 0$, and any $n \geq 1$.

This gives us a good indication of how close our estimate for m is when we know n. Conversely, it can serve as a means for choosing n when we want a certain confidence in the estimate. For more sophisticated estimation techniques involving consumer and producer risks, see par. 8-1.3.2.

8-8.2 CENTRAL LIMIT THEOREM

This states that under some commonly attained conditions, the average of a sum of random variables X_i becomes normally distributed as the number of summands becomes large. Since the normal distribution is easy to use, and since for the conditions of the theorem the exact form of the *pdf* for the X_i is not needed, this theorem is very useful in estimation. More specifically, let X_i be identically distributed random variables with mean m and standard deviation σ, then the *pdf* of

$$\left[\left(\sum_{i=1}^{n} X_i/n\right) - m\right] \Big/ (\sigma/\sqrt{n}) \quad (8\text{-}153)$$

tends to the $n(X|0, 1)$ *pdf*. Equivalently, the *pdf* of

$$\sum_{i=1}^{n} X_i/n$$

tends to the $n(X|m, \sigma^2/n)$ *pdf*.

8-9 HISTOGRAMS

8-9.1 USE

In many instances the assumption that the repair time of systems is distributed according to some of the theoretical distributions (exponential, gamma, normal,

Class Interval Number j	Class Midpoint	Class Width	No. of a_j Observations	Frequency F_j	Cumulative Frequency
1	x_1	At	a_1	F_1	M_1
2	x_2	At	a_2	F_2	$M_2 = F_1 + F_2$
3	x_3	At	a3	F_3	M3 $= F_1 + F_2 +$ F3
m	x_m	Δt	a_m	F_m	$M_m = \sum_{j=1}^{m} F_j = 1$

lognormal, etc.) may be an oversimplification of the facts. One finds very frequently that the repair time has a multimodal distribution. In such cases, the use of theoretical distributions and attempts to combine them so as to approximate the real life multimodal distribution of system maintenance time become very cumbersome and, at times, mathematically nontractable.

Recourse to distribution-free methods is then necessary, and the histogram/polygon method to graphically display the probability density and the cumulative distribution of the repair time of both simple and complex systems becomes very practical in the maintainability prediction process. From such graphs of the maintainability function, it is possible to read directly estimates of the MTTR, the median time at the 50th percentile, or M_{MAX} at the 90th or 95th percentile.

8-9.2 FREQUENCY HISTOGRAM AND POLYGON

Let us assume that a system can fail in n different modes which occur at different frequencies and that each of these failures require a different time to repair the system. To construct the frequency histogram, one needs as inputs estimates of the failure rates λ at which the failure modes occur, and estimates of the system repair times t for all the possible failure modes. First, one ranks the repair times for the various failure modes in an ascending order: $t_1, t_2, t_3, \ldots, t_i, \ldots, t_n$, with t_1 being the shortest repair time and t_n the longest one. With each repair time is associated a failure rate or, more precisely stated, a steady-state renewal or replacement rate (see par. 2-3.2.5), at which that particular failure mode occurs: λ_1, A, $\lambda_3, \ldots, \lambda_i, \ldots, \lambda_n$. These failure rates are obtained from reliability analyses. For a serial system the total failure rate is the sum $\lambda = \lambda_1$ t λ_2 t $\lambda_3 + \ldots + \lambda_n$. For each t_i we compute the relative frequency of its occurrence F_i as the ratio λ_i/λ. As a next step we choose a suitable time interval which includes all time estimates from t_1 to t_n (such as from $t = 0$ to $t \leq t_n$) and divide it into m class intervals, possibly (but not necessarily) of equal width Δt, as shown in Fig. 8-8. Now we denote the midpoints of the class intervals by $x_1, x_2, \ldots, x_j, \ldots, x_m$ and count the number of observations, i.e., the number of repair time estimates which fall into each Δt time interval. Thus in the jth time interval we shall have a_j time estimates of different length t, each of which must satisfy the inequality:

$$x_j - \Delta t/2 < t \leq x_j + \Delta t/2 \qquad (8\text{-}154)$$

Now the relative frequency F_j of repairs (failures) occurring in the jth time interval is the sum of the F_i's falling into that time interval, i.e., summed for the a_j observations in the interval. We tabulate the result as shown in Table **8-5**.

Having performed the tabulation, we are ready to draw the frequency histogram and the cumulative frequency polygon as shown in Figs. 8-8 and 8-9 for 100 failure modes, with the time scale divided into **13** class intervals. The choice of the length of class intervals is important for the proper presentation of the data (Ref. **21**, p. **51**). When fitting smoothed curves, we obtain approximations of the probability density function and of the cumulative distribution $M(t)$, for the maintenance time of the analyzed system. The area above the cumulative frequency curve is, per Eq. 8-3, an estimate of the MTTR (see Fig. 8-9).

When the whole system consists of line replaceable

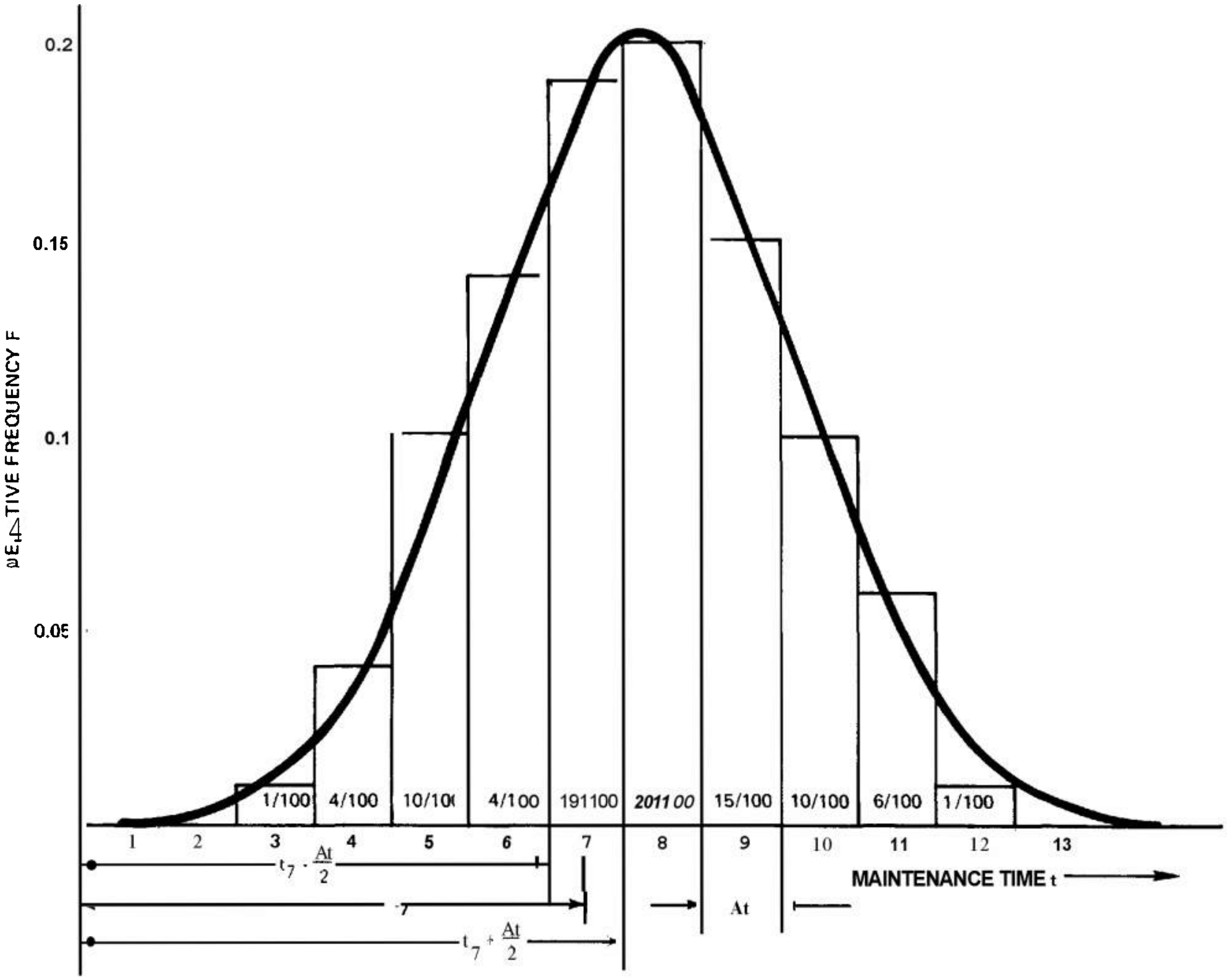


Figure 8-8. Relative Frequency Histogram

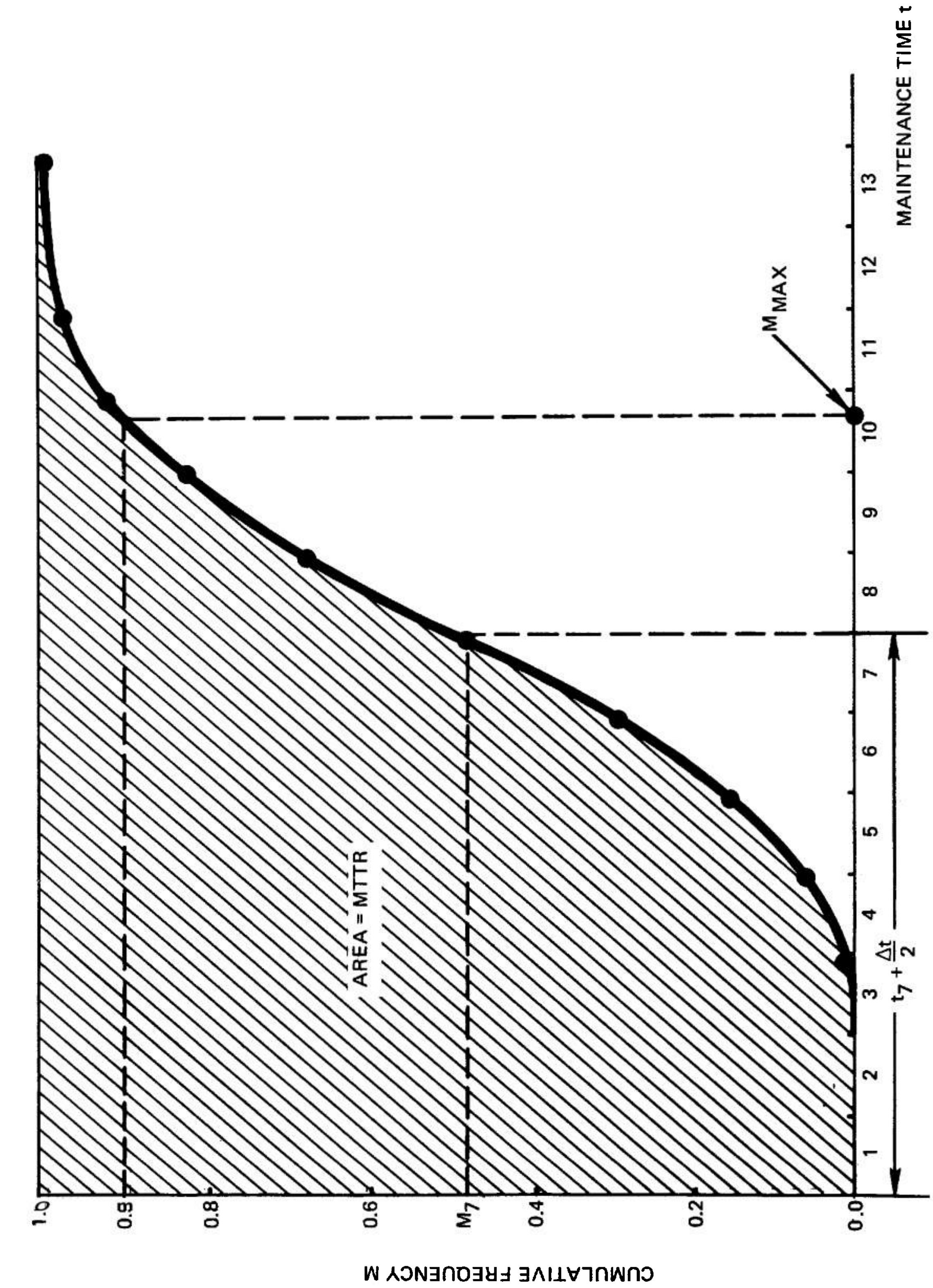


Figure 8-9. Cumulative Frequency Diagram

units (LRU's) and there are only a few (X) of them (say, up to **20),** we have Xsystem failure modes and X associated system repair times (by LRU replacement). In **such a case it is** advisable to **skip the** grouping **into** class intervals and develop directly the cumulative frequency polygon, **as** shown in Ref. **21,** p. **54.**

Using the statistical tools developed in this chapter will aid the maintainability engineer in performing the tasks of maintainability prediction, demonstration, and estimation. The types and sources of maintainability **data** needed to perform these tasks **are** discussed **in the** following chapter.

REFERENCES

1. AMCP **706-**110 through **-114,** Engineering Design Handbook, ***Experimental Statistics, Sections I through 5.***
2. John E. Freund, ***MathematicalStatistics,*** Prentice-Hall, Inc., Englewood Cliffs, N.J., **1962.**
3. ***Experimental Statistics,*** *op. cit.,* ***Section 5 (Tables).***
4. J. Dennis Aigner, ***Principles of Statistical Decision Making,*** Macmillan Decision Series, The Macmillan Co., Canada, **1968.**
5. Hunter M. Sohn, "Proven Techniques for Maintainability Predictions and Demonstration", presented at 10th Reliability and Maintainability Symposium, ASME, Anaheim, Calif., **1971** (unpublished, available from Lockheed-Georgia Company, Marietta, Georgia).
6. Donald D. Gregor, ***Operational Capability Analysis,*** Nov. **1970,** Document NB **70-251,** Northrop Corporation-Aircraft Division.
7. B. ***W.*** Lindgren, ***Statistical Theory,*** Macmillan, Canada, **1968.**
8. Ed. **M.** Abromowitz and I. A. Stegun, ***Handbook of Mathematical Function4*** **(p. 925-995,** "Probability Functions"), **US** Department of Commerce, **1965.**
9. National Bureau of Standards, ***Tables of Normal Robability Functions, US*** Department of Commerce, ***Applied Mathematics Series,*** **23, 1953.**
10. R. L. Bovaird and H. I. Zagor, "Lognormal Distribution and Maintainability in Support of Systems Research", ***Naval Logistics Quarterly*** **, 8, No. 4,343-356,** (December **1961).**
11. William J. Horvath, "Applicability of the Lognormal Distribution to Servicing Times in Congestion Problems", ***Operations Research,*** **7,** No. 1, **126-128,** (Jan.-Feb. **1959).**
12. J. Aitchison and J. A. Brown, ***The Lognormal Distribution,*** Cambridge University Press, **1969.**
13. C. E. Pangburn and C. Arabadjis, ***The*** Log-***Normal Distribution,*** General Electric Company, Report No. **R67S1PD12,** April **1967.**
14. Lindgren, ***op. cit.***
15. Freund, ***op. cit.,*** Sections **7.3** and **6.4.2.**
16. H. **L.** Harter, ***New Tables of the Incomplete Gamma Function Ratio and Percentage Points of the Chi-square and Beta Distributions,*** U.S. Government Printing Office, Washington D.C., 1964.
17. Freund, ***op. cit.,*** Sections **5.4** and **8.2.2.**
18. A. **E.** Sarhan and G. B. Greenberg, "The Best Linear Estimates for the Parameters of the Normal Distribution", ***Contributions to Order Statistics,*** John Wiley and Sons, Inc., N.Y., **1962** (pp. **206-269).**
19. Norman L. Johnson and Samuel Kotz, ***Continuous Univariate Distributions - I,*** Houghton Mifflin Company, Boston, **1970,** (pp. **59-63).**
20. W. J. Dixon, "Simplified Estimation from Censored Normal Samples", ***Annals of Mathematical Statistics,*** **31, 385-391, 1960.**
21. A. Hald, ***Statistical Theory with Engineering Application,*** John Wiley and Sons, Inc., N.Y., **1952.**

CHAPTER 9

MAINTAINABILITY DATA

SECTION I

LIFE CYCLE DATA

9-1 INTRODUCTION

Establishment of the maintainability data collection, analysis, and corrective action system is one of the twelve essential tasks required in the conduct of a well organized Maintainability Engineering Program Plan (Ref. 1). Army Regulations (Ref. 2) require suitable planning of maintainability data collection to assure effective performance of the maintainability design function during the phases of the materiel life cycle. Therefore, it is stressed that an effective and economical maintainability data system should involve only those data essential to timely and intelligent assessment of the maintainability design function and logistic support costs.

Since the principal function of the maintainability engineering discipline is to design materiel so as to improve availability for mission performance and to reduce logistic support costs by reducing maintenance downtime, this function can be stated in simplified form as: "Minimize maintenance downtime" and "Minimize logistic support costs".

Thus, effectiveness of the maintenance functions is based upon the ability of maintainability engineering to design equipment and develop maintenance plans which will optimize the materiel uptime while minimizing logistic support costs. The starting point for both the theory and practice of the maintainability engineering discipline is to gather, validate, and evaluate existing maintainability data from similar equipment and associated maintenance concepts. This allows one to: (1) effectively develop the maintainability design characteristics required to fulfill mission objectives, and (2) develop and evaluate the life-cycle support requirements. Maintainability data collection must be planned, maintained, controlled, validated, and evaluated in order to assure continued appropriateness of the maintainability designs throughout the life cycle and proper performance of the maintenance function.

Both the Military Standards (Ref. 1) and the Army Regulations (Ref. 2) state that existing data sources will be used to the maximum extent possible, basing the data system requirements only on essential data and specific plans. One must therefore define exactly the data to be collected and the controls to be exercised in order to assure that all data relevant to the equipment under consideration are appropriately routed, analyzed, and acted upon in an effective and expeditious manner. Military data banks, e.g., TAMMS, make the data available to any Government or industrial agency on a "need-to-know" basis and subject to security limitations (Refs. 1, 2).

There are numerous maintainability data sources in the military agencies, NASA, Bureau of Standards, Energy Research and Development Administration, and industrial agencies (Ref. 3). However, the data available from these sources are not always directly applicable to a particular materiel. Ideally, the necessary information would be obtained by collecting and analyzing all sources of data and observing the performance of such materiel in current use. Since this is not feasible because of costs and time involved, the current Army policy is to plan and control limited data collection (Refs. 4, 5). From the various sources of data available, maintainability engineers should construct the necessary maintainability data tables to formulate baseline historical information applicable to the specific materiel of concern. This is a current practice in the aircraft and electronic industries (including military aviation and commercial airlines) where historical data concerning existing materiel constitute the basis of the maintainability predictions and maintainability design requirements for new materiel (Ref. 6). A maintainability data system therefore entails the following essential functions:

1. Observe-collect-record-store maintainability data

2. Analyze-validate-evaluate maintainability data

3. Retrieve-utilize maintainability data (for immediate product compliance and corrective action, as well as for long range improvement of product design).

Maintainability data collection, evaluation, and tabulation should be a continuous process during all phases of the equipment life cycle. There are no handbook sources of maintainability data tabulations available that are pertinent to a particular equipment application, such as is available to the reliability engineer in MIL-HDBK-217 (Ref. 7) and FARADA data. Therefore the maintainability engineer is required to evaluate available data on the basis of similarity to the equipment of interest. Fortunately over the past two decades, some contractor maintainability engineering activities have been accumulating and evaluating data pertaining to their sphere of interest (not publishable because of proprietary interests), and thereby are able to justify their particular usage in new equipment designs. In a few instances, such as data furnished in MIL-HDBK-472 (Ref. 8), data have been cited based on particular special applications as examples in explaining the use of prediction techniques but such data are not useful as a reliable data source for new equipment applications. Thus, in the planning for collecting and evaluating data for use at the various equipment life-cycle phases, the maintainability engineer must define the need for and the type of data required in relation to the cost value benefits and constraints.

Maintainability design requirements and maintenance concepts are developed early in the system life cycle by feasibility studies. Such feasibility studies consider maintainability design and logistic support problems in order to arrive at realistic assessments of their potential effect on the support costs and availability of the materiel. These considerations, in turn, are derived from historical maintainability and logistic support information from similar systems. The data are validated and evaluated in order to select the best technical approaches and to identify areas of high technical risks. Using operational analysis, system analysis, and cost effectiveness techniques, the maintainability and support requirements appropriate to the particular system of concern are selected.

Feasibility studies must include as a minimum the following considerations:

1. Cost (both initial procurement and life-cycle support)

2. Failure rates (actual rates versus MIL-HDBK generic rates, plus projected development of state-of-the-art)

3. Scheduled maintenance replacement rates (allowable and required preventive replacement actions; scheduled servicing; etc.)

4. Replacement and repair times (mean time, median time, maximum time, man-hours per operating hour, standard deviations, probabilities of repair accomplishment)

5. Maintenance concepts (organizational, intermediate, and depot levels at which maintenance is performed; effects of levels of maintenance on the equipment mission availability; maintenance skill level requirements)

6. Maintainability characteristics of the equipment (ease of maintenance, fault isolation and location methods, access, item remove and replace, item serviceability checkout, throw-away versus repair, preventive maintenance modifications, and overhaul)

7. Equipment failure modes and their effects on system mission performance

8. Item criticality and high risk (new or modified development).

The key factor in collecting and analyzing data during concept development is to obtain only those data essential to the feasibility study. Therefore, one has to plan a procedure for data Collection which will reveal the critical trends. Factors to be included must represent realistic and measurable requirements. For example, a study was recently made for a military support aircraft in which analysis of historical maintainability data for the existing system (which the projected aircraft was to replace) revealed that four items were found to cause 40% of the aircraft maintenance downtime (landing gear, wheels/brakes, windshield/cockpit items, and flight controls rigging). Design of each item was studied in depth, and the resulting approaches or modifications reduced predicted downtime by a factor of four at a lower cost of procurement. The feasibility study covered only those elements which caused excessive downtimes, failure rates and causes, and outlined improvements needed, risks involved, and anticipated benefits. In this manner, the maintenance concepts were defined, their reliability-mission effectiveness and related costs determined, and preferred alternatives established (Ref. 6).

In summary, collection, validation, and evaluation of data are essential in establishing firmly and effectively the need and justification for the design and procurement of maintainable equipment for military use, and in the proper application of the techniques of maintainability engineering. One must realize as a precaution in planning and evaluating data that a data system is only as effective and valid as the individual involved is capable of observing and recording the data. Constraints of data collection, skills, time, environment, and cost are ever-present.

Contractors or industrial agencies should make use of the available military maintainability dgta systems as a basis for evaluation of product improvement and for maintainability engineering analyses, predictions, and allocations. These data systems also provide information concerning current use of the equipment. Contractor or industrial agencies have access to standard military data collecting systems, provided their "need to know" is within proprietary and security limitations. Some overall summaries include too much general data and the cost of extracting the essential data is excessive. Therefore, essential data should be specified so that it can be retrieved from a general computerized data system at reasonable cost, to fit the need of the user. Thus, the data analyst who establishes the need for data is a key link in the proper planning evaluation, use, and dissemination of information. He is the interpreter between the "user" and the designer.

Subsequent paragraphs in this chapter furnish the details concerning the development, acquisition, retrieval, and processing of data as well as the types of data and available data banks.

9-2 DATA TYPES

There exist many types of data in the life cycle of an equipment. These data are collected in such formats as engineering design log books, equipment log books, analysis and evaluation studies and reports, test logs and reports, handbooks, time and motion tables, production log reports, quality assurance reports, design review reports, military in-service maintenance data formats-reports-summaries, special data observations and reports, performance reports, and scientific engineering data banks (Ref. 3).

Therefore, in order for the maintainability engineer to take advantage of the voluminous data sources available, the data program must be organized and controlled on a basis tailored to solve the problems involved at the activity level of concern (Ref. **4).** All maintainability data are concerned with the following fundamental factors:

1. Time required for and frequency of a maintenance action (task)

2. Cost of life-cycle support (skills, spares, equipment, facilities and procedures).

All aspects of the types of data sought and used should relate directly to these factors, and data collection categories should be constrained accordingly (Refs. 9, 10). When specifying types of data collection, the key objective is to assure that the rationale of the "need to know" is clearly explained so that personnel involved in recording the data understand its importance and are motivated to record accurately. Two methods of obtaining data are useful:

1. Standardized source document formats, containing basic data required by existing maintenance management data collection systems, require maintenance technicians to provide the desired information as directed. The raw data are collected by a central processing unit and distributed on a "need-to-know" basis or upon special request for specific data available on standard formats (tape, decks of cards, microfilm, or other printouts). This process has the advantages of lower cost and the use of fewer personnel. This source is of considerable value because the standard formats, printouts, and tapes can be used with existing industrial and military computer facilities to code and retrieve specific details needed from the data available in the standard formats. Also, computer programs can print out chart displays of past trends and future predictions depending upon the statistical evaluation methods used. Fig. 9-1 is a typical computer-printout display chart showing the reliability (frequency of maintenance) achievement of an item. This type of display can reveal for management the operational failure time results.

A disadvantage of the standardized form is that existing systems invariably have questionable accuracy and are incomplete in some of the specific details concerning cause and frequency of downtimes. In such cases, use of the method described in the next paragraph may be necessary. In any case, both methods should be used to develop a file of special background data and usable tables as a baseline for the practice of maintainability engineering.

2. Special observations and data collection are more costly and yet more accurate. This approach, in which technical personnel concerned with the specific evaluation of equipment performance observe and record certain factors concerning sample equipment, has the following advantages:

a. Personnel have a more thorough understanding of the specific objectives and the "need to know", and thus a higher interest in the study and the results is maintained at the source of the data.

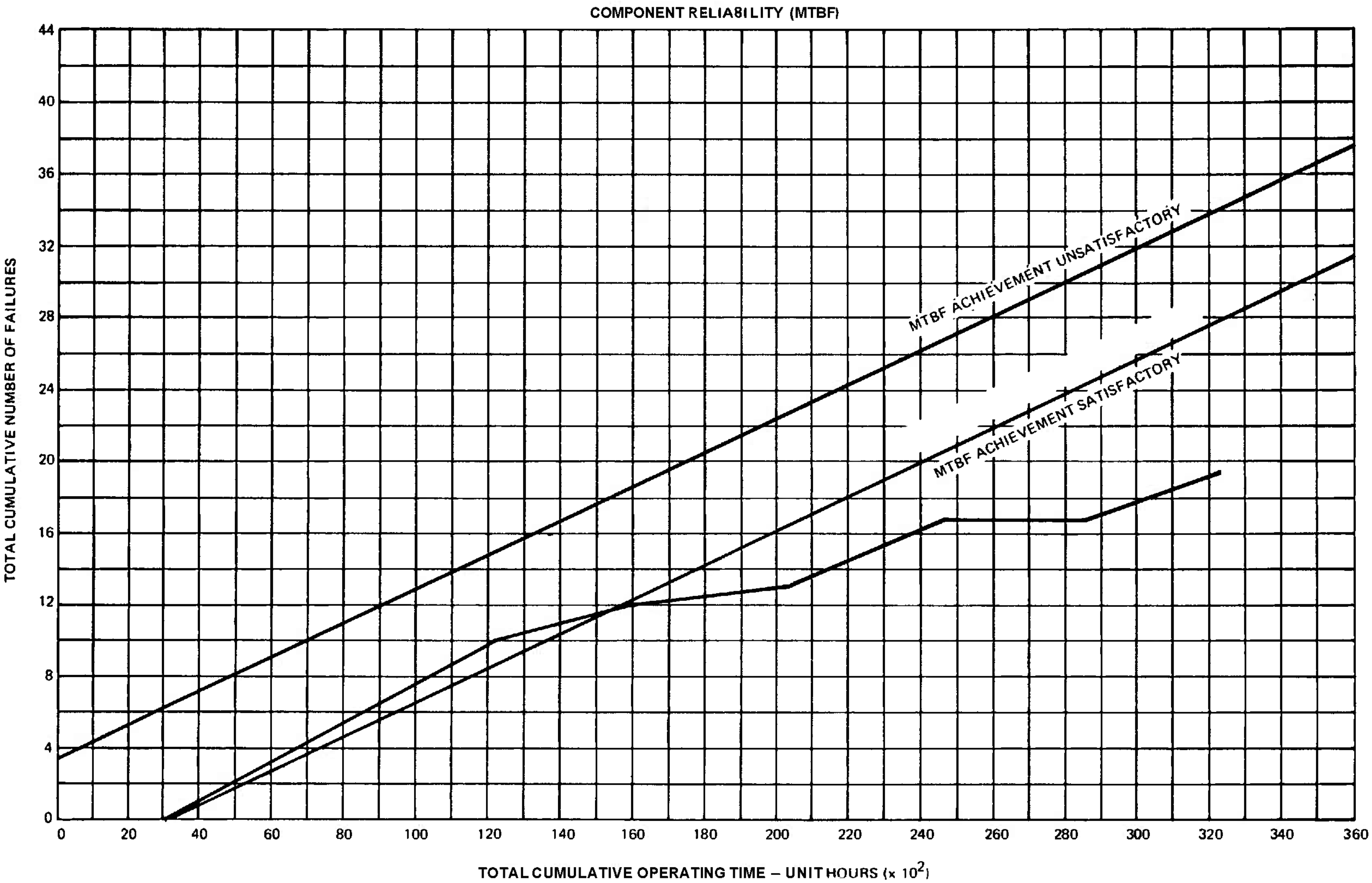


Figure 9-1. Computer Print-out of the Operational *MTBF* Achievement (simulated)

b. Close daily contact with the system user personnel and the evaluators keep the study on course.

c. Data which are supplied under close monitoring and recheck conditions require less interpretation before final processing.

d. Attention can be given to developing the required qualitative and quantitative details, trends, and human factor principles that are pertinent to the evaluation.

e. Inconsistencies and errors in the data can be detected through cursory checks and analysis with corrections applied at, or close to, the time of occurrence.

f. Cause of malfunctions can be analyzed in more detail at time of occurrence.

g. Evaluations can be made with fewer samples.

The disadvantage is that such a method is relatively expensive. However, considering the small sample used, the results are cost-effective for decisions regarding specific product improvement. Furthermore, logistic support and deployment interruptions are minimized. Therefore, in applying this method, the cost-effectiveness of an organized and controlled data collection plan must be considered (Refs. 4, **5**, 9).

9-2.1 HISTORICAL DATA

Historical maintainability data are used to establish the maintainability requirements for newly proposed equipment. These data should include consideratidns of both qualitative and quantitative factors. Without the knowledge of the maintenance downtimes, frequencies of their occurrence, anthropometric data, and operational and logistic factors of existing systems in an operational environment, any requirements set forth for improvement of proposed systems may result in unrealistic constraints, high risks, and costly life-cycle support.

The history of many new systems developed during the past two decades has revealed decreased reliability and increased cost of maintenance and logistic Support due to the complexity of the systems, even though a large portion of the systems used existing equipment items. This was partially due to improper use and validation of available data.

During the 1960's there were a number of contracts to study and evaluate historical maintainability data for various types of equipment in order to arrive at a basis for maintainability predictions for new equipment as delineated in MIL-HDBK-472 (Ref. 8). Similarly, data studied during the late 1950's resulted in many human factor handbook specifications and publications of tables. In addition, the military has banks of data which report the maintenance and logistic support expended over many decades for deployed equipment. The keeping of military use records started with individual equipment log books and has extended to sophisticated computer processing complexes. This has been supplemented by various industrial and scientific agencies, both under military contract and in-house efforts.

Despite the fact that data are available, adequate use has not been made of them, and evaluation has been neglected, invalid, or incomplete. Generally, the data that have been published are too generic to be applicable to a new "state-of-the-art" equipment. As an example, Tables 2 and 3, Procedure 11, MIL-HDBK-472, show vacuum tubes but not microcircuits, transistors, and diodes.

Several of the large prime contractors and subcontractors during the past decade have established data collection-analysis-evaluation systems and are using them for product improvement and for newly proposed systems. In almost every instance, the evaluation is proprietary and unavailable for general use in maintainability engineering. In a few cases, national symposia disclose benefits in a fragmentary way, such as Ref. 6, or data sources are disclosed by the military, such as those listed in Ref. 11.

The list of historical maintainability data regarding existing parts and components that may be incorporated in new systems is too voluminous and is of questionable reliability to be cited in this text. These types of historical data generally apply to electric, electronic, mechanical, and electromechanical items. It is true, most available data apply to functional electronic systems used in various environments, while little has been pinpointed to mechanical systems. However, the kind of data concerning fundamental maintainability characteristics remains the same.

To illustrate the use of historical data, previous data for an existing fighter aircraft landing gear brake and wheel assembly indicated a 10 to 15 hr downtime to replace and repair. By analysis of the maintenance tasks involved, a new design concept evolved whereby a simple, strut jacking system allowed one nut to be removed and an entire wheel-drum assembly to be replaced, reducing the downtime to 1 to 2 hr. By designing the wheel drum assembly as a one-part replacement instead of 15, the cost of the design of the new assembly procurement was lower, yet performance remained equal or better.

Another case was an airborne computer for a 1970 design concept. High density chips were mounted on one printed circuitboard of throw-away design in lieu of 1965-vintagesmall chips using 10 repairable printed circuit cards. The downtime and frequency of replacement decreased one order of magnitude and procurement costs remained the same. Not only were the quantitative factors of maintenance downtime and frequency considered, but the associated qualitative factors of skills required and ease of maintenance were also included.

Both of these cases resulted from product improvement and life-cycle cost feasibility studies of the historical data collected and analyzed from previous systems. Both serve to emphasize the principle that when the facts are known, the evaluation of alternatives to accomplish the same functions with equal or better performance at less cost is possible.

Another example of the use of historical data is illustrated by Fig. 9-1 and Fig. 9-2. These compare several user installations and an individual user's results. The collected data were programmed to allow retrieval of the essential data to reveal details as well as to extrapolate trends. These are also examples of the use of data collected from many sources to show the manufacturer the maintenance status of his equipment in an operating environment.

9-2.2 ANTHROPOMETRIC DATA

Anthropometric data are concerned with the geometry of the human body and clearances necessary for it to function properly while performing a given task. These types of data are essential qualitative factors in determining the ability of the technician to perform the maintenance tasks. The old cliche "if you can't get at it, you can't fix it" is a primary element in the ease of a maintenance function. The scoring criteria of Checklist "C" Design Dictates, ***Maintenance Skills Prediction Method 3,*** MIL-HDBK-472 (Ref. 8), furnish the definitions of 11 anthropometric characteristics with respect to capabilities of an average technician. Although these definitions appear complete, one has to refer to the standard human engineering handbooks to obtain the details of the span of reach, height, length, and lifting power for the minimum/maximum profiles for various human percentiles. It is interesting to note that the 11 characteristics are directly related to the maintenance task times determinations.

Three of the many convenient references for maintenance technician anthropometric measurements are given in:

1. ***Guide to Design of Electronic Equipment for Maintainability,*** WADC Technical Report 56-218, 1965

2. ***Anthropometry of One-Handed Maintenance Actions,*** Technical Report NAVTRADEVCEN 330-1-3, US Navy Training Device Center, 1960

3. "Maintainability Engineering" Chap. 6, Department of Army Pamphlet 705-1, June 1966.

Anthropometric data are acquired by observation and measurement of actual human factors relating to tasks performed by maintenance technicians. During the early development of equipment design, special observations are made by the maintainability engineer to verify and evaluate the qualitative design features built into the equipment. During the design process, the maintainability engineer in his Maintainability Design Guide stipulates the human body constraints that are both peculiar and general to the items involved in the design process. The end-item specifications should designate these peculiar and general constraints. In the review and monitoring of the drawing board designs, sometimes paper, cardboard, or wooden mock-ups are built for visual recognition of the constraints, especially for handling and accessibility factors. Mock-ups are the least costly method of insuring that designs meet these human factor constraints. In the Army, the tear-down maintenance phase of system development is the focal point for assessment of the human factors constraints. Whenever special maintainability data are obtained for a sample of the deployed equipment, human anthropometric data are observed and reported in order to evaluate the achievements or needs for product improvement. In studying, validating and evaluating historical data and trends for new or proposed equipment, it may be necessary to make special observations or observe mock-up of the equipment to determine the causes of excessive maintenance downtime and associated life-cycle support problems.

In all anthropometric observations, evaluations take into account the technician skill and rating requirements that represent the population of the technicians available to the military services. Statistical methods are used to determine the validity of the evaluations in terms of the various percentiles of maintenance types available for the tasks under observation. Any data without a suitable range of observations are suspect. Within the range of observations, one must evaluate each of the various maintenance environments.

9-2.3 OPERATIONAL REQUIREMENTS

The operational requirement for equipment, ability to perform its primary mission whenever necessary, is the reason for the need to build maintainability into the equipment. The primary measure is maintenance

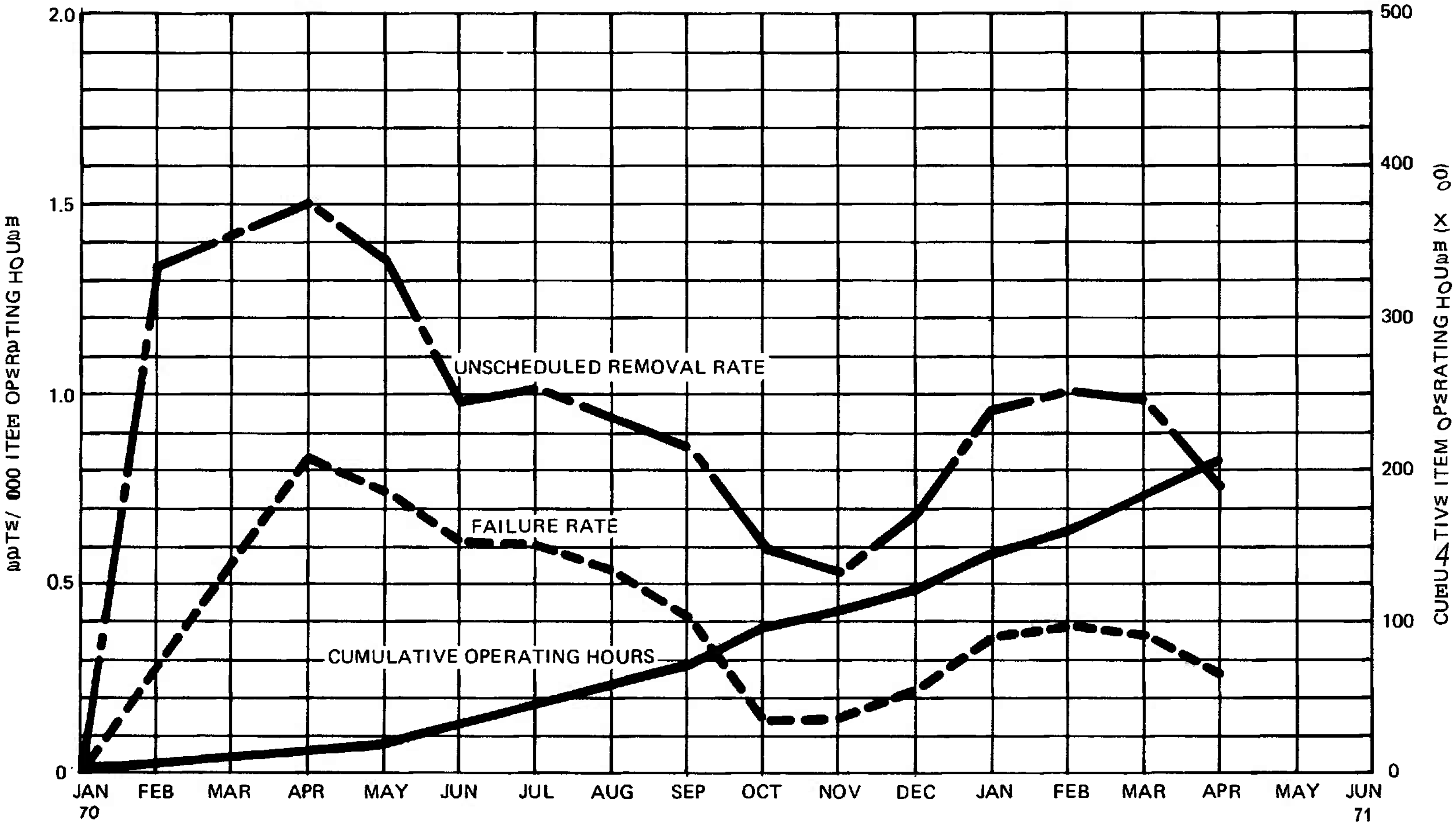


Figure 9-2. Graphic Display of Historical Data

downtime resulting from equipment failure. The fundamental sources of data that define this measure are:

1. The Failure Modes, Effects and Criticality Analysis of the equipment

2. The historical and/or selected data obtained from similar equipment (maintenance time frequency, levels of maintenance, associated qualitative features, and logistic support).

Also included in the operational requirements analysis are the secondary or ancillary missions of the equipment and the respective availability priorities. In considering operational performance, the effects of a malfunction that can be deferred and the related mission degradations are established. To complete the data analysis for maintainability, the effects on the costs of the life-cycle logistic resources are evaluated, and the maintenance concepts are established.

The following are specific operational data elements that are defined, analyzed, and evaluated:

1. Equipment operational availability

2. Mission goal (primary and secondary missions for which equipment is designed)

3. Mission duration parameters (maximum and average under war-time and peace-time operational environments)

4. Failure modes and effects of equipment failure

5. Equipment turn-around time (maximum, mean, and minimum)

6. Maintenance frequency in the operational environments (failure rates, scheduled maintenance rates)

7. Preventive maintenance time (scheduled periods, allowed-average, and maximum)

8. Corrective maintenance downtime (minimum, mean, and maximum at specified percentiles)

9. Standard deviations and confidence levels of maintenance downtime in the operational environment

10. Maintenance concept (operational levels of maintenance, repair levels, etc.)

11. Maintenance float requirements (peace-time/war-time end items authorized for storage at installations for replacement of unserviceable items when repair cannot be accomplished within a specified period of time by the intermediate level of maintenance)

12. Integrated logistic support for the maintenance plan (life-cycle costs, skills, training, technical procedures, support equipment, facilities, and technical services)

13. Maintenance, logistics, standby readiness, and other nonoperational downtimes.

The chief parameter, as stated previously, is availability for mission operation. This is the ratio of uptime to uptime plus downtime, with downtime for maintenance support as the principal parameter investigated for establishing the maintainability design principles to be built into the system. The downtime impact on availability is measured in terms of frequency of failure (need for corrective maintenance), frequency of a scheduled maintenance (need for preventive maintenance), and the time to perform the maintenance action (repair, replace, adjust, service, inspect, logistics delay, etc.) In evaluating the availability factor, it is essential that data from existing and/or prior equipment be considered in order to arrive at realistic requirements. A functional flow diagram will facilitate consideration of the specific operational data elements. The functional flow block diagram uses a noun-verb description, and shows the system mission (both primary and secondary) and the maintenance operational concepts as well as the effects on the specific operational elements. The diagram thus puts consideration of all of the elements in their proper perspective. When trade-off decisions become necessary, the definitions and the respective elements of cost and mission effectiveness are properly weighed to encourage effective and realistic decisions. The functional flow block diagram is discussed in par. 4-2.1.1.

In some complex equipment, the need arises for computerized mathematical models in order to arrive at optimum decisions. These analyses depend upon consideration of all of the relevant data elements in order to show trends, assumptions, and proposed constraints upon which trade-offs and other decisions are made. A ***system effectiveness operational analysis*** depends upon the maintainability engineering function as the data source for the required maintenance and logistic support information. Maintainability engineering, in turn, derives its data from data bank sources which are evaluated and validated according to the pertinent types of data.

Of special importance in the determination of operational requirements are data affecting various design parameters that influence downtime, i.e., the effects of detecting failures by associated built-in test, auxiliary test, and check-out equipment. This feature has a direct bearing on the effectiveness of the operational requirements as well as a direct influence on downtime and operational availability of the equipment. Validation of the failure modes and effect data should include evaluation of the method of fault detection and adequacy of the equipment operational check-out in determining the time parameters. Historically, fault detection time has been one of the largest contributors to maintenance downtime.

9-2.4 FAILURE RATES

Failure rates for corrective maintenance, and frequency of preventive maintenance are essential parameters in the determination of equipment maintenance downtimes. Time required to perform a maintenance task is not an arithmetic average of task times, but rather is a sum of the individual task times multiplied by their frequencies and divided by the sum of the frequencies of all the related tasks.

Frequency data are needed to assess all phases of maintainability engineering analysis, and perform predictions, and allocation. It is important to recognize and understand the types of failure rate data and scheduled maintenance rate data in use, e.g., generic rates, generic usage rates under various operational environments, and historical data based on actual use of similar equipment. Generic rates may be the only data available for preliminary predictions and allocations. In some cases, it may be necessary to use multiplication factors to increase the generic rates to operational use rates. In others, overall use factors have been applied, based on the types of equipment. In still other cases, generic use factors have been applied to individual components and their related operational environment. The most useful and reliable failure frequencies are based on historical operational data of similar equipment actually operating in the military combat environment.

9-2.4.1 Manufacturer's Data

Failure rates, used to show compliance with *MTBF* requirements, are usually insufficient to establish failure data of an item, system, or equipment to be used for maintainability analysis. Newly designed parts must undergo life tests of sufficient sample size to be valid. When tested under laboratory and simulated environments, the values obtained are generic rates. When related to field environment, operational degradation factors must be used. Whenever manufacturers' test data are used, the test conditions must be known, and the rates extrapolated accordingly. Refs. 3 and 12 through 17 list the current manufacturing sources of data and data types available, as well as the various governing agencies collecting failure rate data.

9-2.4.2 Operational Use Data

Operational use data are obtained either by general maintenance management data collection systems of the entire deployed population, as reported by the maintenance activity on standard source documents and summarized through the use of computer techniques; or specific controlled data collection systems which observe only a limited sample of the deployed population (Refs. 4, 5, 9).

In the first method, the validity depends upon the technicians' opinion as to cause of the failure and also upon the relative capabilities of the available technician. The overall failure rate under an operational environment gives a realistic use rate and therefore is a valid rate for assessing the real life factors. In some deployed units where the rate is higher than in others, one is able to pinpoint unusual environments or management controls, and determine where and how the use rates can be improved, or define areas for further special observation.

The second method of special observation and controlled data collection for a limited deployed sample of the population is the most desirable for validating equipment failure rates and making pinpoint determinations of the causes of failures. The selective method does provide detailed data showing the factors that affect the failure rates and the downtimes, and the knowledge of which is necessary to obtain design decisions that will ease maintenance and reduce the failure rates (Refs. 5, 18).

Both methods establish historical data upon which to base operational degrading factors over laboratory testing conclusions. This is the important factor that must be considered by the maintainability engineer in order to develop designs or product improvements in terms of the real life downtimes that affect availability. Analysis and evaluations of the various military, governmental, and industrial data bank information must include the source of data, validity of the data, causes of failure, and use times in order to establish confidence in the failure rate data (Refs. 3, 12, 15).

9-3 EQUIPMENT VALIDATION DATA

9-3.1 GENERAL

During equipment validation, data are collected, analyzed, and evaluated to determine the quantitative maintainability features and the associated qualitative design features (Ref. 9). Early in the life-cycle of the equipment, decisions are made concerning the specific maintenance level at which item replacement due to malfunction will be made. Also in this phase, decisions are made regarding preventive maintenance which will meet the requirements for equipment performance.

The first type of data generated consists of detailed maintainability engineering analyses of the maintenance tasks for each maintenance level function (operator maintenance, organizational maintenance, direct support maintenance, general support maintenance, and depot maintenance). These data consist of time estimates and predictions based on engineering judgment, historical data on similar equipment, time-and-motion studies, preliminary design breadboard and mock-up tests, environmental tests, user-service test reports, and manufacturer parts list and specification data.

The second type of data generated are demonstration and test data observations and evaluations resulting from prototype production items used for maintainability demonstrations, engineering prototype tests at proving grounds, maintenance evaluation, and other qualification testing.

In both data types, design reviews (in-house and customer) are conducted using the results of tests, design trade-offs, analyses, evaluations, and drawings to demonstrate that maintainability requirements established for the end-item specificationshave been achieved. The data so generated are used in the engineering analysis of the detailed maintenance tasks in order to define the integrated logistic support resources for the equipment life cycle (manpower skills and training, spares and maintenance float, support equipment, technical publications and procedures, facilities, and contract support services).

9-3.2 OPERATIONAL EQUIPMENT DATA

As previously stated, historical data from deployed operational equipment consist of two types:

1. General operational maintenance data collected by existing military maintenance data management policies

2. Special data from a selected sample of the deployed equipment. Included in the latter are specific manufacturers' data for equipment used in the military operational environment; caution must be taken to insure that the data are field operational data (as opposed to laboratory test data).

The essential maintainability data measurements are:

1. Mean time between maintenance (hours, cycles, miles, rounds, etc.)

2. Man hours (per operating hour or operating cycle)

3. Mean, median, and maximum maintenance action times

4. Standard deviation of maintenance task time distribution

5. Skill levels and ease of maintenance factors

6. Operational usage delay times (administrative, logistics, etc.)

7. Equipment availability

8. Failure modes and effects (fault detection and location elements).

Usually these types of operational data are limited to organizational and intermediate levels of maintenance.

Depot level measurements are generally reported not as operating data but in terms of summaries affecting the logistic parameters for reconditioning of equipment.

The key principle in the analysis of operational data is the evaluation of the equipment parameters that affect mission availability. The equipment data generated and analyzed must reflect the total equipment operation and degradation effects.

In many cases the operational data generated from the general types of data are insufficient to use with confidence, but trend projections may show the need for the respective maintainability parameters to be de signed into the equipment. For critical areas affecting availability the data must be obtained and analyzed using specific observation of a selected sample of deployed equipment.

9-3.3 DESIGN DATA

Design data are obtained by study, review, and evaluation of equipment drawings, specifications, and parts lists. Once the drawing has been approved, the designed-in maintainability features are fixed. Any subsequent changes are costly and in some cases prohibited. Therefore, a day-by-day, week-by-week review of designer progress is made so that the design data maintainability parameters are evaluated together with other design parameters. The maintainability qualitative design data are obtained from review of the drawings (ease of maintenance). Measurements of the time to perform a maintenance task are then analyzed. For instance, the feature of accessibility defines the time to get at an item for replacement; the attachment feature defines the time to remove and replace; the test point availability, identification and/or interface with built-in test equipment defines the fault location and detec tion time. The failure rate data or frequency of mainte nance is used to analyze the mean downtime for maintenance. The parts lists and the associated failure rates or preventive maintenance frequency rates are used to define the need for maintenance (generic failure rate is increased to determine the operational use rate). The specification data are used to obtain the qualitative and quantitative constraints of maintainability imposed on the design.

9-3.4 PROPOSED PARTS AND COMPONENTS TEST DATA

Generation of test data for parts and components provides the failure rate data required to determine how frequently maintenance will be necessary which information in turn is used in the computation and evaluation of the mean and maximum maintenance times. Such data are used in the failure modes and effects analysis that also defines the maintenance needs.

The data are generated in terms of failure rates, *MTBF,* degradation factors, operational use factors, and laboratory environmental factors: The data are collected from manufacturers' parts test records and specifications which show the results of testing and in special environment. The key factor in the use of such data is to have a full understanding of the conditions of the test and the relationship of the test to the equipment under consideration. Generally detailed information on parts and components involves proprietary rights, and the relationship to the specific application must be negotiated during the proposal-to-buy transaction.

9-3.5 RECORDS OF TRADE-OFF DECISIONS

These types of data are limited to a specific problem concerning specific equipment being developed at a particular time. Usually the data thus collected and used are lost for historical purposes.

Maintainability trade-off decisions require the following types of data to be collected, analyzed, and evaluated (Ref. 19):

1. Performance Factors
   a. *MTBM* (frequency of corrective and preventive maintenance)
   b. Repair times (mean, median, maximum, and associated standard deviations)
   c. Equipment availability (uptime versus uptime + downtime ratios)
   d. Maintenance concepts (levels of maintenance)
   e. Utilization rates
   f. Failure modes, effects, and criticality analysis
2. Handling Features
   a. Weight
   b. Size
   c. Human factor capabilities and associated support equipment
3. Equipment Placement Features
   a. Racks
   b. Accessibility
   c. Cable and mechanical interconnections
   d. Equipment layouts
4. Packaging Features
   a. Accessibility
   b. Modularization
   c. Standardization
   d. Functional grouping
   e. Interchangeability
   f. Plug-in components
5. Diagnostic Aids Factors
   a. Built-in test
   b. Test points location, identification
   c. Fault detector displays
   d. Ancillary test support equipment
6. Adjustment Requirements
   a. Accessibility and quantity of adjustment points
   b. Interaction effects
7. Displays and Control Features
   a. Panel layouts and illumination
   b. Self-indicating fuses and spare fuses
   c. Meters
8. Labeling (external/internal) Adequacy
9. Safety Factors
   a. Personnel safety
   b. Equipment safety
10. Integrated Life-Cycle Support Requirements
    a. Tools/test equipment
    b. Technical data
    c. Personnel skills, quantity, training
    d. Facilities
    e. Spares
11. Schedules
    a. RDT&E
    b. Production
12. Cost (unit, system) Parameters
    a. Initial investment, development, and acquisition
    b. Life-cycle support costs
    c. Effects of previous 11 factors on cost parameters
    d. Throw-away versus repair costs

e. Functional and esthetic values.

Not all of these data items are considered in every trade-off study and some that are unlisted may arise in a study and need to be considered. Of paramount concern in evaluating item trade-off studies is the cost effect in determining the weighting factors assigned to each element. Of similar importance is the effect of change decisions on the scheduling parameters. The two requirements for equipment operation (availability to perform its mission, and the life cycle logistic support) are also related to each of the weighting factors established. In all maintainability trade-off studies, the effects on interrelated performance data parameters must be considered, and similarly in all system analysis and effectiveness trade-off studies, the maintainability data parameters must be evaluated.

9-3.6 ENGINEERING TEST AND SERVICE TEST DATA

As stated previously, Engineering Tests and Service Tests (ET/ST) are conducted to verify and evaluate equipment adherence to the performance characteristics and requirements. During these tests, controlled detailed data observations are collected. The evaluations are usually statistically oriented to determine the performance built into the designs, to determine trends, to base modification for product improvement, and to verify and establish the integrated life-cycle resource support requirements. In some instances, such as gun tests, life tests to destruction are made.

The following are essential maintainability data to be obtained:

1. Quantitative Performance Elements
   a. *MTBM* (frequency of corrective and preventive maintenance tasks)
   b. Maintenance man hours per operating hour, round, mile, cycle
   c. Downtime for maintenance (mean, median, and/or maximum at specified percentiles, the associated standard deviation and confidence levels)
   d. Equipment availability (ratio of uptime to uptime + downtime)
   e. Quantitative factors of life-cycle logistic support, such as repair parts and maintenance floats
2. Qualitative Elements Affecting the Quantitative Elements
   a. Maintenance tear-down features
   b. Ease-of-maintenance design features, such as accessibility, diagnostic aids, and skills required
   c. Qualitative features of integrated logistic support, such as need for repair parts, need for levels of maintenance, and facilities.

At present these types of data are recorded and available through Army data banks and can be used as historical data sources.

9-3.7 HISTORICAL DATA FROM COMPONENTS AND PARTS

These types of data are similar to the data cited in par. 9-3.2, Operational Equipment Data. In the paragraph cited the essential data measurements are for an equipment-level summary. If a breakdown on a detailed parts and components level is desired, a controlled data collection plan is devised to obtain such in-depth data from the data banks as pertains to the specific type of equipment of concern (Ref. 2).

These types of data include the following elements for end item, components, and piece parts:

1. Repair times by level of maintenance
2. Maintenance man hours per operating hour, mile, or round
3. Skill levels used
4. Distribution of maintenance tasks
5. Failure rates for corrective maintenance and preventive rates for scheduled maintenance
6. Transportation time between user and the various supporting shops and supply facilities
7. Administrative delay times
8. Support equipment data and operational facilities information
9. Test instrumentation and test methods.

Generally, obtaining data from the entire deployed population is not economically feasible and is too unwieldy to analyze and evaluate. Therefore obtaining such data by selected controlled data collection methods is preferred (Refs. 4, 5). For example, MIL-HDBK-472, which contains data from specific studies of special populations, may be used to make maintainability predictions (Ref. 8). In the past, military activities have had to depend upon contractor support for such detailed evaluations or rely on gross summaries of the deployed equipment population. Ref. 11 lists a typical example for Air Force ground electronic end items. There is a trend toward military activities obtaining such data on a specialized controlled basis (Ref. 5 is an example.)

SECTION II

DATA ACQUISITION

9-4 PRINCIPLES OF DATA ACQUISITION

As stated previously, a basic principle in data acquisition is that data are only as reliable as the observer's ability to record the data. Also of importance is the observer's understanding of the need for recording the data. A third principle is the need for the data to be of sufficient quantity and on a timely basis in order to be really useful.

9-4.1 METHODS OF ACQUIRING MAINTAINABILITY DATA

The two fundamental methods of accumulating data related to the maintainability characteristics of a product are:

1. Use of standardized reports of corrective and scheduled maintenance for the total deployed equipment population, where the observer is the technician or supervisor of the maintenance activity;
2. Use of controlled special reports of maintenance actions for selected portions of the deployed equipment on detailed formats, where the observer is a specially trained technician.

Deployed equipment population is equipment in field use. It can also be considered, prior to field use, as the population of equipment under development and/or production where Engineering and Service Tests are used to obtain performance data. Although no attempt is made in this text to present an exhaustive list of characteristics, the following are essential generic data required for maintainability evaluation:

1. Frequency and reason for maintenance action
2. Downtime required to perform the maintenance action (minutes, hours, man-hours).
3. Manpower and skills required
4. Spare parts
5. Identification of equipment/systems/components being maintained
6. Logistic support required, and logistic delays
7. Cause of malfunctioning
8. Time between failures.

The following objectives are essential to assembling data (Refs. 9, 10, 19):

1. All terms, codes, and data elements shall be clearly defined.
2. The acquisition system shall be based on a complete analysis of the equipment performance requirements.
3. The individual technician, operator, supervisor, or observer must understand the need for and the type of data to be acquired.
4. Acquisition of the required data must be timely in order to be effective.
5. Determination of data required must be in terms of the needs of management, engineering, and logistics, and should be met in an economical manner. Visual charts must show the effects, distribution, and trends of the data for management decisions.
6. Quality assurance procedures will be incorporated to make certain that data collected is adequate for the intended purposes.
7. All data must be kept current to the extent required by the system.
8. Duplication of data should be prevented. Experience of industrial agencies has shown that many existing governmental computerized data acquisition systems have adequate data available. With proper evaluation of the needs, it has been found that existing tab runs can be obtained and coded to fit individual tab run needs for evaluation (Ref. 6).

No attempt is made here to publish maintenance formats used to collect and acquire data. The references cited in this section show the multiplicity of formats used (general and specific). All of the formats contain certain data as basic elements plus the individual elements considered essential to the particular system used.

9-4.2 METHODS OF PROCURING DESIGN PACKAGEDATA

The design package data are obtained by the procuring agency by negotiation at the time of the contract award. This design package contains data which were used to arrive at the design decisions made for the equipment and includes maintainability predictions, allocations, and task analyses. In some cases, these data involve proprietary rights which need to be firmly delineated in the procurement contracts. As is seen from the types of data listed in par. **9-4.1,** much of the needed information is available from design organizations (both industrial and military laboratory). Therefore it is essential at contract negotiation that the controlled plan to obtain such data be developed in detail, defining the need, availability, processing, and timely reporting. The objectives stated in par. **9-4.1** must be established and adhered to. Failure to properly define the problem has resulted in many contract cost escalations.

9-4.3 DATA FROM MANUFACTURER'S SPECIFICATIONS AND BROCHURES

Many of these types of data are of a generic summary nature for specific end items and are not in sufficient detail to be useful except for end-item purposes. Details of the end item are proprietary, and the contractor reserves the right to contract for overhaul and/or warranty repair. If the item procured is off-the-shelf under Government contract, repair data are usually available from military overhaul and reconditioning facilities. The Government standard supply catalog is also useful. In addition, military historical data are available from repair reports for similar items. In some cases, more in-depth data are available from proposal justifications for items to be procured. In all cases, the maintainability engineer (both military and industrial) must investigate and evaluate the validity of the data, especially the qualitative maintainability characteristics upon which the quantitative parameters are based.

9-4.4 INCIDENT STATISTICS REPORTING FAILURES

These types of data are obtained from breadboard, mock-ups, laboratory tests, environmental testing, prototype testing, and field sources—from both industrial and Governmental agencies. The basis of incident statistics must be evaluated to determine if the data are derived in terms of generic failure rates, or degradation failure rates resulting from factors of environment, operational degrading, or field deployment. Since frequency of repair is derived from failure rate data, the maintainability engineer must weigh the condition under which the data are derived so as to validate the use of the data.

9-4.5 GOVERNMENT DATA COLLECTION SYSTEMS

The primary military sources for maintainability data are the general maintenance management systems, such as Air Force AFM **66-1** data, Navy **3M** data, and the Army maintenance data collection systems. In addition there are various reports from the Army, Navy, and Air Force covering special and controlled data for statistically sampled equipment. This handbook does not describe the details and the reporting forms, since such details are shown both in the references cited and in the respective maintenance management regulations that define the procedures (Refs. 1, 2, **4, 5**).

9-4.6 LATEST THINKING ON CONTROLLED DATA COLLECTION

The collection, storage, retrieval, analysis, and dissemination of data is a major expenditure for system program management. Many agencies throughout industry and Government (especially military) collect and process raw data, but in almost all cases this has been done to satisfy their particular requirements.

The present state of the art in data acquisition, analysis, and processing does permit measurement of the effectiveness of competitive weapon systems. However, if this is to be accomplished, pertinent data must be made available so that the experience can be applied to the development of new weapon systems.

Ref. **20** describes the Air Force's System Experience Correlation and Analysis Program (SECAP) and the Reliability Analysis Central (RAC) System used to illustrate the latest thinking on integrating standard data systems within the Air Force. Ref. **21** describes the National Bureau of Standards Reference Data System. The Introduction in Ref. 9 cites the need for and the interest in establishing an integrated data system. Refs. **12** through **14** and **15** through **17** cite the Tri-Service Integration of Reliability and Failure Rate Data (FARADA and GIDEP).

9-5 CENTRAL DATA BANK

Improvement in the effectiveness of the national system for scientific and technical information is a matter of great popular concern (Refs. 20, **21**). "Experimental results in measurements are the backbone of physics. No theory is acceptable unless it agrees with experimental data. Conversely, a systematic study of experimental results can suggest new theoretical approaches. Tables and graphs of numerical data therefore play an important role in the progress of science . . . Thus it is obvious that data compilations are of great importance and should have the full cooperation of those producing data. It is also clear that modern computer techniques can handle such data more efficiently than old tabulations could, especially since their number and variety are growing more rapidly." These words of Samuel Goudsmidt ***(Physics Today,*** 19(9), **52)** define the problem and emphasize the importance of the availability of data. Such compilations of data have not been available to the maintainability engineer due to proprietary factors, and he has had to rely on a small amount of data obtained from limited sources.

Compilations of data are the basic tools of scientists and engineers. Handbooks containing useful information on the properties of substances and systems are available and of invaluable assistance, such as the ***Standard Handbook for Mechanical Engineers,*** by Baumeister and Marks. Yet, for the maintainability engineer, no such maintainability data handbook has been published. Maintainability practices and principles for application to product improvement of entire systems have evolved from small, specially selected samples of limited validity.

The problem of a Central Data Bank is recognized by the Government. The Nationwide National Standard Reference Data System (NSRDS) is a starting point. NSRDS is operated by the Bureau of Standards and coordinated by the Committee on Scientific and Technical Information (COSATI), which consist of representatives frum all Government departments and independent agencies having technical information programs. As recognized in Ref. **15,** NSRDS should then be regarded as the primary source of data measurements by the technical community of the United States. SECAP and RAC-Central are examples of the feasibility and usefulness of such central data banks. When all data from the NSRDS are available in the central data bank, it will have become a useful tool. Until the central data bank can be used, maintainability engineering does have available existing military central data bank agencies. Experience has shown that whenever these data banks have been used, the basic data thus made available to the maintainability engineering discipline has resulted in decided product/system/equipment improvement (Ref. **6).** Essential factors in the efficient use of central data bank information are:

1. A two-way flow of information between contractors, designers, and maintainability engineers

2. Standardization of methods of collection, analysis, and evaluation

3. Standardization of coding, processing, and dissemination of data.

The benefits to be derived when these factors are implemented are:

1. Reduction of man-power requirements
2. Reduction of training
3. Speed of obtaining and retrieving data
4. Accuracy of data
5. Data handling economy.

Standardization of the present formats will lower cost of data reduction. There will always be unique items of vital concern to the individual agencies, but these need not affect the standardization principle, as can be noted by review of the various formats and procedures. Ref. **14** shows the integration of the Air Force **66-1** systems and SECAP-RAC Central System to illustrate the standardization principle.

The technology of data communications in the United States has been widely developed and used in commercial applications. One of the problems holding up implementation of this technology in the military and other Government agencies is the cost of installation of communication terminals. Recent technological improvements involving lower costs will allow a data acquisition operator to use telex or similar machines to retrieve required data files and associated printouts and visual displays from central data banks. Time-sharing systems will make it economically feasible to extend computer utility throughout the system complex. The emphasis in information storage and retrieval has changed from design of document handling systems to design of handling systems, as used for the Air Force Reliability Analysis Central facility.

Until retrieval by telex or similar techniques is available, microfilm or tape can be used to collect data furnished from central data banks. Tape is the preferred method where large amounts of data are necessary. Such tape can be coded, and local computer programs can be developed for the print-out of the desired data and data displays. Microfilm is used where the amount of data is small, and hand analysis and evaluation can be performed. The decision as to which to use is based on cost-effectiveness.

9-6 DATA RETRIEVAL AND PROCESSING

Good data bank management is prerequisite to efficient data retrieval and processing. There are two basic principles: the product is only as good as the accuracy of the data input, and only that data for which there is a use should be stored. Therefore, the first management step is the establishment of adequate formats and coding of the data bank files to serve as the data base of the desired output (Refs. 20, 21).

The maintainability data elements to be retrieved and processed are those elements related to maintenance action (what, where, when, how, and why an item was maintained, together with maintenance time and the man-hours and skills expended), and those items needed to validate and evaluate the logistic support requirements.

REFERENCES

1. MIL-STD-470, ***Maintainability Program Requirements,*** 26 March 1966, Par. 5.10.
2. AR 702-3, ***Army Materiel Reliability, Availability, Maintainability.*** and 4.
3. ***Reliability and Maintainability Data Source Guide,*** US Naval Applied Science Laboratory, Brooklyn, N.Y., January 1967.
4. AMCR 750-43, ***Special Operational Data Collection,*** 22 April 1970.
5. H. H. Dobbs and J. Thompson, ***Field Experiment Controlled Data Collection 155MM S.P. Howitzer,*** Report CWS 70-002, US Army Weapons Command, January 1971.
6. ***Maintainability Seminar,*** Los Angeles Maintainability Association, 1 May 1972, Sections IV and VII, AD-723 227.
7. MIL-HDBK-217, ***Reliability Stress and Failure Rate Data for Electronic Equipment.***
8. MIL-HDBK-472, ***Maintainability Prediction.***
9. W. H. Van Alven, Ed., ***Reliability Engineering,*** Prentice-Hall, Inc., Englewood Cliffs, N.J., 1964, Chapter 4.
10. AMCP 705-1, ***Maintainability Engineering,*** June 1966, Chapter 12.
11. AFSC DH-1-9, ***Maintainability Handbook (for Ground Electronic Systems),*** 20 December 1970, Design Notes 2A2, 2A4, and 2A5.
12. ***FARADA Failure Rate Data Program,*** Naval Fleet Missile System Analysis and Evaluation Group, Corona, California, April 1969.
13. ***FARADA Program Participants,*** Naval Fleet Missile System Analysis and Evaluation Group, Corona, Calif., March 1971.
14. ***Army-Navy-Air Force-NASA FARADA Program,*** Naval Fleet Missile System Analysis and Evaluation Group, Corona, Calif., April 1970.
15. GIDEP (Government-Industry Data Exchange Program), ***Information Systems Handbook,*** Naval Fleet Missile Systems Analysis and Evaluation Group, Corona, Calif., 6th Annual Work Shop Conference, Los Angeles, Calif., May 1-3, 1968.
16. GIDEP (Government-Industry Data Exchange Program), ***Exchange Program Joint Agreement Letter, Dept. of Navy, Hq. Naval Material Command, Army Materiel Command, Chief Naval Material Command, Air Force Logistic Command, Air Force Systems Command;*** MAT 03 RE, Naval Fleet Missile Systems Analysis and Evaluation Group, Corona, Calif., 21 Dec. 1970.
17. GI-DEP (Ggvernment-Industry Data Exchange Program), ***Roster GI-DEP-III Representatives,*** Naval Fleet Missile Systems Analysis and Evaluation Group, Corona, Calif., April 1971.
18. D. W. Sharp, ***Data Collection and Evaluation,*** Proceedings of the 5th National Symposium on Reliability and Quality Control, 1959, pages 146-160.
19. B. S. Blanchard and E. E. Lowery, ***Maintainability Principles and Practices,*** McGraw-Hill Book Co., N.Y., 1969, Chapter 7.
20. M. Haus, et al., ***Air Force System Effectiveness and Reliability Data Systems,*** Reliability and Maintainability Conference, 6th Proceedings, Paper 670, July 1967, pages 116-139.
21. E. L. Brady and M. B. Wallenstein, "The National Standard Reference Data System", Science, 12 May 1967, Volume 156, No. 3776, pages 754-762.

CHAPTER 10

CONTRACTING RESPONSIBILITIES

SECTION I

INTRODUCTION

10-1 GENERAL

The user-producer dialogue is discussed in par. **1-4.1.** The Army as user represented to industry by AMC, and industry, as producer, must enter into a contractual agreement with respect to the system to be developed and produced. Both parties, therefore, have certain responsibilities for mutually establishing the contractual requirements for maintainability. Chapter 3 describes in detail the organization and management requirements for maintainability and how these requirements are carried out throughout the system life cycle.

To properly plan and implement a maintainability program, both the Government and the contractor must carry out certain tasks at certain times during the life cycle. These include:

1. Development of a maintenance support plan
2. Performance of maintenance engineering analyses
3. Establishment of maintainability design requirements reflecting the maintenance concept
4. Preparation of the maintainability program plan appropriate for the specific program
5. Performance of maintainability analyses, predictions, allocations, and demonstrations
6. Participation in design reviews for maintainability and in change control for maintainability
7. Participation in the performance of maintainability trade-offs and trade-offs of maintainability with other design disciplines
8. Preparation of necessary technical documentation as it affects maintainability.

Some of these tasks are the responsibility of the Government (Army) and others are the responsibility of the contractor. AR 702-3 (Ref. 1) establishes concepts, objectives, responsibilities, and general policies for the Army reliability and maintainability programs. This regulation establishes reliability and maintainability characteristics which must be specified for the design of materiel and must be considered and assessed concurrently throughout the life cycle. AR 750-1 (Ref. 2) defines the materiel maintenance component of the Army Logistic System, including maintenance support planning as an integral part of the acquisition process for Army systems or equipments. It also assigns functional responsibilities of the Government and the contractor, and prescribes policies for the maintenance function.

10-2 GOVERNMENT RESPONSIBILITIES

Maintainability recognized and acknowledged as one of the major system engineering disciplines and explicitly treated as part of the system/equipment contractual requirements. The Government has a number of responsibilities, such as:

1. Establishing maintenance support and maintainability requirements
2. Assuring that these requirements are adequately specified in requirements documents and in system/equipment specifications
3. Assuring that maintainability program requirements are covered in contractual documents
4. Monitoring, reviewing, testing, and evaluating the contractor's design effort for maintainability.

These responsibilities reside at different levels and among different organizations in the Army. For example, the Deputy Chief of Staff for Research, Development, and Acquisition has primary Army General Staff responsibility for the overall reliability and maintainability program pertaining to materiel and equipment.

He formulates, issues, and maintains current Army policies on, and exercises Army General Staff supervision over, reliability and maintainability programs (AR 702-3).

The Deputy Chief of Staff for Logistics has primary Army General Staff responsibility for maintenance policy and support planning. He develops and coordinates the Army General Staff position for maintenance concepts, conducts periodic reviews concerning status of maintenance support planning, and participates in support development and readiness planning (AR 750-1).

Using agencies have responsibility for establishing and maintaining controls at appropriate levels to assure effective coordination of reliability and maintainability program functions and execution of the policies set forth by the Army General Staff. They also have responsibility for delineation of new materiel requirements to include realistic reliability, maintainability, and availability characteristics consistent with the goal of integrating equipment into the Army at realistic costs in resources (AR 702-3).

The Commanding General, TRADOC, as representative of user agencies, is responsible for assuring that the maintenance concept—based on prescribed policy provided in the requirements documents—is realistic and sufficiently definitive to furnish essential data required by, and reflects appropriate inputs from, the developing and other participating agencies, such as the Deputy Chief of Staff for Personnel, and the Commanding General, US Army Forces Command (AR 750-1).

The Commanding General, US Army Materiel Command, has responsibility for

1. Establishing and maintaining integrated controls at appropriate levels to assure achievement of reliability, maintainability, and availability requirements and execution of the policies set forth by the Army General Staff

2. Validating the reliability, maintainability, or availability requirements proposed for items that are under his development responsibility

3. Preparing, coordinating, and distributing to all participating agencies a maintenance support plan for each end item

4. Assuring that the development acceptance test of an item includes the testing of all elements of maintenance support as well as the demonstration and evaluation of reliability, maintainability, and availability

5. Preparing and quantifying descriptions of the elements of maintenance support for incorporation in the system description applicable to each development project, including the maintenance concept

6. Implementing a maintenance engineering analysis system in accordance with TM 38-703-3

7. Assuring participation of appropriate personnel during the equipment development cycle as required at bidder's conferences, contract negotiations, source selection, and reviews made of contractor's plans, designs, or engineering change proposals (AR 702-3 and AR 750-1).

Specific responsibilities with respect to the life-cycle management of US Army materiel are delineated in detail in DA Pamphlet 11-25 (Ref. 3). AMC, as the primary Army agency interface with contractors, must see to it that the various requirements are reflected in system development specifications and contractual documents.

Army contracting responsibilities include specification of the following, as required for the item under contract:

1. Operational requirements and constraints

2. Logistic support, maintenance policy, and maintenance concept requirements or constraints

3. Reliability, maintainability, and availability requirements

4. Requirements for a formal maintainability program in accordance with MIL-STD-470

5. Maintainability prediction requirements and method to be used per MIL-HDBK-472 or otherwise

6. Maintainability demonstration requirements and method(s) to be used per MIL-STD-471 or otherwise

7. Maintenance engineering analysis requirements per TM 38-703-3

8. Related human factors, safety, standardization, and similar requirements

9. Maintainability design policies, criteria, and constraints to the extent required by the particular development program

10. Other maintainability-related requirements as applicable to the specific program, such as GFE, interface with other equipments or contractors which affect the specific program but which are not a part of it

11. Contract data requirements per DD 1664, Data Item Descriptions, and specified on DD 1423, Contract Data Requirements List.

10-3 CONTRACTOR RESPONSIBILITIES

The contractor's responsibilities with regard to maintainability vary with the size of the program and its contractual requirements. As a minimum, when maintainability requirements are not specifically called out, he must provide those design features and characteristics in his equipment which will promote ease of maintenance.

At the opposite extreme, the contractor must demonstrate his capability to perform a full-scale maintainability program effort in accordance with MIL-STD-470. He may have to demonstrate in his proposal, prepared in response to an RFP, that he has an established maintainability organization capable of preparing the maintainability program plan and conducting the approved maintainability program. In addition, he may have to supply information with regard to his management policies and practices as they affect maintainability. This information might include:

1. A chart that shows the maintainability organization and its relationship with other groups within the company organization
2. Responsibility and authority of the maintainability organization, specifically the organizational element responsible for implementing the maintainability program and internal effort planned in connection with the maintainability program
3. Operational methods and procedures
4. Identification of key personnel
5. Methods for reporting maintainability status to program management
6. Methods for program control of maintainability
7. Methods for effecting coordination and liaison, including design reviews and maintainability demonstration tests
8. Policy for monitoring and controlling subcontractor/vendor maintainability programs.

The contractor must demonstrate that he has a working knowledge of existing maintainability standards, specifications, analytical methods, and the published literature.

Among the contractor's responsibilities, as specified in contract task statements, are to

1. Prepare the maintainability program plan
2. Perform maintenance engineering analysis
3. Establish maintainability design criteria
4. Perform maintainability analyses
5. Perform maintainability predictions and allocations
6. Perform design trade-offs including interfaces with related disciplines
7. Perform maintainability demonstration tests
8. Participate in design reviews
9. Establish data collection, analysis, and corrective action system, coordinated with other disciplines
10. Prepare maintainability status reports
11. Establish a maintainability assurance program
12. Incorporate and enforce maintainability requirements in subcontractor and vendor contract/purchase specifications.

SECTION II

THE CONTRACTING CYCLE AND MAINTAINABILITY INPUTS TO CONTRACTING

10-4 GENERAL

Maintainability inputs to contracting are reflected in the system/equipment specification and in maintainability program requirements which are made part of the contract. It is essential that the maintainability requirements be related to and derived both from operational requirements (Plan for Use) and the maintenance concept. This latter, in turn, is derived from the Integrated Logistic Support concept (Plan for Support) and the maintenance support plan, as described in Section I of Chapter 5, and as illustrated in Fig. 5-1. It is only through such an orderly procedure that effective and thorough managerial direction, planning, programming, and resource allocation will be provided throughout the life cycle, and cost-effective, design-oriented maintainability requirements will be clearly specified so as to enable the objectives of the reliability and maintainability programs to be achieved for the system.

This will help contract negotiations with respect to maintainability since it will allow maintainability requirements to be stated in mission-oriented terms whose definition and meaning will be mutually understood and agreed upon by the Government and the contractor. It will also minimize design and contract changes resulting from ambiguity or improper planning and inadequate specification. One of the problems associated with contracting for and demonstration of maintainability is whether to specify an inherent or operational availability (par. 1-7.3). While it is desirable to be able to demonstrate that the system will meet an operational availability, the definition of operational availability includes supply and administrative delay times and other operational factors over which the designer has no control. This opens the door for later disputes and involves the Government as part of the demonstration and guarantee of the achievement of contractual requirements. This is why inherent availability, which is design controllable, is almost always used as the basis of maintainability design and demonstration.

10-5 MAINTAINABILITY INPUTS TO CONTRACT1NG

The establishment of an effective maintainability program requires that the life cycle approach described in Chapter 3 be applied to the contracting cycle, tailored as required by the nature, size, and complexity of the program. The logical development from operational (mission) requirements to the ILS concept to maintenance support planning to maintenance engineering analysis to maintainability design prediction and demonstration should be either an explicit or an implicit part of every maintainability program. Omitting or telescoping parts of these in order to meet unrealistic time deadlines only serves to introduce serious problems in meeting program objectives, resulting in costly changes and compromising the program. It is now recognized that program progress should be tied to measurable event milestones rather than schedule deadlines.

A number of regulations, policies, specifications, standards, and handbooks exist on the subject of maintainability which should be intelligently applied; tailored, as appropriate, to the needs of each program; and referenced in specifications, RFP's, contractors' responses to RFP's, and contractual documents. These have been discussed in some detail in the preceding chapters of this handbook. Both technical and contracting personnel of the developing agency should become familiar with the requirements of these documents insofar as they affect their respective areas of concern and responsibility. It is not sufficient just to incorporate these documents in specification and contractual documents by reference.

For example, MIL-HDBK-472 contains four maintainability prediction methods. The Government developer must determine whether any of these methods can be used or modified in his program, or whether some other prediction method would be more suitable, as specified by the Government or recommended by the contractor in his proposal or contract.

Similarly, MIL-STD-471 contains six maintainability demonstration methods. The developing agency must determine which of these should be used, or which combination if more than one is to be called out—for example, to demonstrate both corrective maintenance and preventive maintenance downtimes. A modified method or one not included in the MIL-STD are still other possibilities.

Of significant importance also is the decision as to which data items are to be furnished by the contractor as part of the contractual requirement, when these are to be available for review or delivered, and if and when they are to be updated. These requirements must be spelled out as part of the Contract Data Requirements List (DD 1423) and supported by Data Item Descriptions (DD 1664) for each data element required, all of which must be included as part of the contract. A sample specification of a maintainability program for a weapon system is given in Appendix A.

Availability, reliability, and maintainability requirements must be stated in terms appropriate to the item considering its intended purpose, its complexity, and the quantity expected to be produced. They must be clear, quantitative where possible, and capable of being measured, tested for, or otherwise verified. Statistical confidence levels and risks associated with demonstrating achievement of these requirements must be stated in documents describing test and evaluation requirements. Contracts for materiel must include detailed reliability and maintainability requirements as specific requirements for demonstrating achievement to the satisfaction of the Army.

Specification of quantitative availability, reliability, and maintainability requirements is discussed in detail in preceding chapters of this handbook. Specifically, it has been pointed out that, wherever possible, it is preferable to specify an availability with a minimum acceptable reliability (*MTBF*) and maximum acceptable maintainability (MTTR) to give the designer as much design freedom as possible, rather than specifying a specific *MTBF* and/or MTTR. Specification of system effectiveness requirements in mission-oriented terms also is discussed.

When the specification of specific maintainability parameters is desired by the nature of the equipment or program, the following considerations should be taken into account (Ref. 1):

1. Maintainability
   a. Probability of completing diagnosis, repair, and verification (of successful correction) within a specified time with specified personnel resources.
   b. Mean-time-to-repair (MTTR). This characteristic will apply to the prime item only and will not include repairs of components.
   c. Maximum-time-to-repair or maximum corrective maintenance downtime.
   d. Maintenance man-hours by skill level/specific maintenance action.
   e. The conditional probability that maintenance at a level higher than organizational maintenance will be required before the system or equipment can be restored to service, given that a malfunction occurs.
   f. Requirements for ease of maintenance, such as:
      (1) The location of high mortality parts to provide ready access when maintenance is required
      (2) The use of quick-release fasteners, wing nuts, and other features that will minimize requirements for special tools.

2. Frequency of Scheduled Maintenance Actions. Minimum allowable time between scheduled maintenance actions for each applicable maintenance level will be specified.

3. Test and Checkout Methods. The amount of built-in test capability, and the degree of failure diagnosis and fault location required without auxiliary equipment will be specified. Compatibility with multipurpose, automatic diagnostic test equipment, available in the field during the same time frame, will normally be required. As early in the life cycle as possible, any requirements for special, multipurpose, or automatic test equipment will be outlined, including requirements for test equipment access points or connections, and built-in sensors or measuring devices. Elimination of the need for checkout after issue (wooden round concept) will be considered.

4. Time Between Overhauls (TBO). This will be specified for items requiring periodic overhaul or rebuild, e.g., engines, vehicles, and missile systems.

5. Maintenance Man-hours/Operating Hour. This is a comprehensive measure that depends on several performance and maintenance characteristics such as MTBF, MTTR, and frequency of scheduled maintenance actions. It is sometimes referred to as the Maintenance Support Index.

It must be emphasized that requirements which are specified for system effectiveness, including reliability and maintainability, must be capable of being measured or demonstrated within a realistic time frame. This requires careful consideration by the developing agency before such requirements are put into specification and contractual documents. For example, an *MTBF* of 5000 hr may be desirable, but may not be demonstrable because of lack of sufficient test models and/or test time, yet it should be stated if it is a practical goal. Special testing techniques may be used to verify the results or the Government may have to rely on predictions or incentive/penalty clauses based on future operation (see Chapter 6).

10-6 GOVERNMENT FURNISHED EQUIPMENT

When Government furnished equipment (GFE) is part of the system and called out in the contract, the Government assumes certain responsibilities to the contractor in meeting system and contractual requirements. Among these are assuring that the contractor receives certain required information and data about the GFE, that it is complete and accurate, and that the contractor receives it in a timely manner.

With respect to maintainability, the following types of information about the GFE are required:

1. Performance specifications
2. Operational data
3. Environmental capabilities
4. Test and evaluation data
5. Maintenance concept and the maintenance support plan
6. Maintainability design features
7. Maintenance engineering analysis data
8. Reliability data, preferably operational from the field, such as MTBF, utilization rates, and reliability block diagrams
9. Maintainability data, such as *MTTR* and *MAXTR*.

Of particular importance is the effect of GFE reliability and maintainability data on system maintainability prediction and allocation. If the contractor must meet some specified availability or maintainability requirement, then, as discussed in Chapter 4, he must make maintainability allocations and predictions based upon the MTBF and MTTR data he has accumulated or estimated. If GFE is part of the system, then the contractor must be furnished such data on the GFE. If the contractor's prediction and allocation, based upon the furnished data, do not meet specification requirements for the system, he has one of several options. One is to change his allocations of reliability and maintainability goals to improve the reliability and/or maintainability of those parts of the system under his design control.

Another is to use techniques such as redundancy to achieve specified values. A third option is to turn the problem back to the Government by indicating that improvements must be made in the GFE or a relaxation of the system specification made. Which of these alternatives is the desired one depends upon cost-effectiveness trade-offs which should be made.

By specifying the use of GFE, the Government assumes part of the responsiblity for meeting system effectiveness requirements. To help avoid later disputes, the Government should furnish reliability and maintainability data on GFE to prospective contractors as part of the proposed package when the RFP is distributed. The contractors should be asked, as part of their proposals, to comment on the given data or to make a preliminary maintainability prediction. One of the dangers in this, however, is that poor data may be furnished and contractors will have a tendency to accept such data and not take exceptions, or will indicate that specified system effectiveness requirements cannot be met, on the basis that this can be postponed until after award of a contract and made the basis of a contract change.

10-7 DESIGN REVIEWS

Design review refers to informal and formal reviews of the development of the system/equipment design conducted at various points throughout the planning and acquisition phases of a program. Design review is intended to ensure that all facets of design (including maintainability) are being incorporated or adequately considered. The subject of design review encompasses a wide scope of interests and disciplines. It is basically agreed that holding separate formal reviews for maintainability, reliability, human factors, etc., would result in confusion, waste of man-hours, extra expense, schedule delays, and most likely would result in little or no accomplishment. Thus, the intent is to formally and logically review the proposed design from the total-system standpoint in the most effective and economical manner through a combined integrated review effort (Ref. 5).

As pointed out by Blanchard and Lowery (Ref. 4). "The formal design review serves a number of purposes.

1. It provides a formalized audit (check) of proposed system/equipment design with respect to contractual and specification require-

ments. Major problem areas are discussed and corrective action is taken.
2. It provides a common baseline for all project personnel. The designer is provided the opportunity to explain and justify his design approach, and the various technical disciplines (e.g., maintainability) are provided the opportunity to hear the design engineer's problems. This tends to facilitate a better understanding among design and support personnel.
3. It provides a means of solving interface problems, and promotes the assurance that the system elements will be compatible. For example, major interface problems between engineering and manufacturing relative to the producibility of the system being designed are often undetected until design data are released and production commences. The result of major problems discovered at that time are often quite costly. Another example includes the common problem of compatibility, or lack thereof between prime-equipment and support-equipment design. Such interface problems are often undetected at an earlier point in time due to a wide variance in organizational interests and activity. A design review prevents these major problems.
4. It provides a formalized record of what design decisions were made and the reasons for making them. Trade-off study reports are noted and are available to support design decisions. Compromises to maintainability are documented and included in the trade-off study reports.
5. It promotes a greater probability of mature design as well as the incorporation of the latest techniques. Group review may lead to new ideas, possibly resulting in simplified processes and subsequent cost savings."

There are a number of points in time along the system life cycle at which review of the status of development of a system/equipment design is not only desirable, but necessary to the achievement of a cost-effective system. The number of design reviews are a function of the system complexity and the state-of-the-art uncertainty. DA Pam 11-25 (Ref. 3) promulgates the Life Cycle System Management Model (LCSMM) for Army materiel systems. It details the process by which Army materiel systems are initiated, validated, developed, deployed, supported, and modified; and indicates where reviews are necessary.

For a complex system which is subject to the complete life cycle development process, starting with concept formulation and proceeding through acceptance of the production design for production and deployment, many design points are required. In fact, there should be a program and/or design review at each major milestone prior to proceeding to the next phase, and at the end of each design state discussed in Chapter 3.

Blanchard and Lowery (Ref. 4) describe conceptual design, system design, preliminary design, and critical design review activities which correspond to some of these design review points.

10-8 DESIGN CHANGES

One of the major concerns to both the system user and the system developer is the control of design changes. This is particularly true once the design has been approved and committed to production and deployment. It is also true, however, during the development phase of the life cycle. Unless closely controlled, changes can and do have serious consequences. If design changes are necessary during development, the potential impact of each change must be ascertained before it is approved. This impact also depends upon the stage in the development cycle in which it is proposed. The later in the design and development cycle in which a change occurs, the greater will be its impact on cost, schedule, and redesign of associated items. For example, changes proposed and made during the detail or production design stages will have much greater impact than those made during preliminary design or engineering development. Changes made during production and deployment may have even more serious implications. The watchword of the program manager, therefore, should be to resist desultory change and to institute adequate control over change proposals and their evaluation and approval so that indiscriminate changes are not permitted.

The control of system/equipment design, and changes to it, is generally called configuration management (Ref. 6). Configuration management has many facets, depending upon the phase of the life cycle to which it is applied. It can be overdone if imposed too early in the life cycle and can be abused and lead to lack of control if not rigorously enforced during later stages.

For example, there is no real need for imposing detailed control procedures in the preliminary design and engineering development stages during which various design alternatives are being explored. Change control, at these points, should be limited to those changes resulting from study and design efforts which indicate a need for changing or waiving specification or other contractual requirements and which require formalization. As the design develops in greater detail during the detailed design, test and evaluation, and production design stages, certain elements of the design will become frozen. For example, repaiddiscard, module size, standardization, test philosophy, and other design decisions will fix or constrain the design alternatives. Deviations from these design decisions must be controlled so that the design maintains its integrity. Such changes must be proposed and documented so that the trade-offs and reasons for the change can be evaluated and the appropriate decisions made. The interface impact of each proposed design change on the other design disciplines must be evaluated as well as the impact on performance, cost, schedule, production, and the possibility of having a number of different versions of the system/equipment fielded.

Ideally, formal configuration management procedures, as spelled out in configuration management manuals, should not be implemented until the complete design has been tested and approved and a first article configuration audit performed. While this is feasible for some individual equipments, the design is often approved in parts for complex systems and equipments. In these cases, formal change control must be initiated as soon as design releases are made.

The mechanism for making and controlling design changes is the Engineering Change Proposal (ECP). The control of change is a mutual responsibility of both the contractor and the Government. Just as for the other design disciplines, the review and approval of ECP's with respect to maintainability is a responsibility of the maintainability engineer.

REFERENCES

1. AR 702-3, *Army Materiel Reliability, Availability, Maintainability.*
2. AR 750-1, *Maintenance of Supplies and Equipment.*
3. DA Pam 11-25, *Life Cycle Management Models for Army Systems.*
4. B. S. Blanchard and E. E. Lowery, *Maintainability: Principles and Practices,* McGraw-Hill Book Company, New York, 1969.
5. AR 70-37, *Configuration Management.*

APPENDIX A

SPECIMEN FOR MAINTAINABILITY PROGRAM REQUIREMENTS

MAINTAINABILITY PROGRAM REQUIREMENTS FOR THE XYZ WEAPON SYSTEM

1.0 *SCOPE.*

1.1 ***COVERAGE.*** This specification provides the requirements for the establishment of a maintainability program by the contractor for the **XYZ** Weapon System.

1.2 ***APPLICATION.*** The requirements of this specification as applicable to the XYZ system throughout its design, development, and production phases. Its applicability commences with Contract Definition, wherein the contractor prepares a preliminary plan, sufficiently in detail to permit evaluation for continuing development of the system. This plan shall describe the efforts to be expended during development and production to assure adequate maintainability of the **XYZ** system, and which meets the requirements of this specification. The preliminary proposal shall be updated and approved by the Government prior to entering engineering development. After approval the plan shall be contractually binding upon the Government and contractor. The contractor's response to this specification shall have headings in the same sequence and title as those of section 5 "Detailed Requirements" below.

2.0 REFERENCED DOCUMENTS The issues of the following documents in effect on the data of request for proposal, form a part of this specification to the extent specified herein:

Specifications, Military

MIL-V-38352 = Value Engineering Program Requirements

MIL-H-46855 = Human Engineering Requirements for Military Systems, Equipment, and Facilities

Standards

MIL-STD-280 = Definitions of Terms for Equipment Divisions

MIL-STD-471 = Maintainability Demonstration

MIL-STD-72 1 = Definition of Effectiveness Terms for Reliability, Maintainability, Human Factors and Safety

MIL-STD-785 = Reliability Program for Systems and Equipments (Development and Production)

MIL-STD-882 = System Safety Program for Systems and Associated Equipment; Requirements for

MIL-STD-1472 = Human Engineering Design Criteria for Military Systems, Equipments and Facilities

Handbooks

MIL-HDBK-472 = Maintainability Prediction

3.0 ***DEFINITIONS***

3.1 The definitions included in MIL-STD-280, MIL-STD-47 1, and MIL-STD-72 1 apply. In addition, the following definition applies:

Contract Definition. That phase of development during which preliminary design and engineering are verified and accomplished, and the contract and management planning are performed.

4.0 ***GENERAL REQUIREMENTS***

4.1 ***Maintainability Program.*** The contractor shall establish and maintain an active and effective maintainability program. This program shall commence as a preliminary program during contract definition, updated and approved by the Government, prior to implementation, after which it becomes a contractual document at the beginning of engineering development of the **XYZ** system. The program shall include as a minimum, but not be limited to the following tasks:

a. Prepare a maintainability program plan
b. Perform maintainability analysis
c. Prepare inputs to the detailed maintenance concept and detailed maintenance plan
d. Establish maintainability design criteria
e. Perform design trade-offs

f. Predict maintainability parameter values
g. Incorporate and enforce maintainability requirements in subcontractor, vendor, and suppliers' contract specifications
h. Integrate other items
i. Participate in design reviews
j. Establish data collection, analysis, and corrective action system
k. Demonstrate achievement of maintainability requirements
l. Prepare maintainability status report.

5.0 DETAILED REQUIREMENTS

5.1 ***Maintainability Plan.*** The contractor shall prepare a preliminary plan for the development and conduct of a maintainability program in his response to the RFP for conducting Contract Definition on the **XYZ** system. This plan shall outline how he intends to develop and conduct the program to meet the requirements of pars. 5.1.1 through 5.12 below and shall be in sufficient detail to permit evaluation of the proposal. The program plan shall be continually updated and finalized during Contract Definition. The finalized plan shall be submitted for review and approval by the Government along with the contractor's proposal for engineering development of the **XYZ** System. Upon approval of the plan and contract, the plan shall be contractually binding upon the contractor and the Government.

5.1.1

Maintainability Organization. The contractor shall submit in his initial maintainability program plan, and update as required, an organizational chart showing the various management functions with specific names of individuals assigned each responsibility. This chart must commence with top management and extend down to a management level at which the maintainability function for the program will be administered. The relation of the maintainability function to reliability, logistic support, safety, system engineering and various other disciplines shall be clearly delineated on this organization chart. The contractor's maintainability shall not be under the control of engineering, manufacturing, marketing, or quality control functions and shall have direct access to the program manager. The contractor's organization chart shall make clear only direct or indirect lines of reporting, be they administrative or otherwise, which apply to the organization responsible for administering the maintainability program effort. The contract shall submit as a part of the maintainability program plan, a breakdown of the maintainability organization identified as being in charge of the **XYZ** program. This breakdown shall show the organization of this activity with the various functions and responsibilities of each organizational group, including the maintainability manager. Personnel shall be listed by name in conjunction with the assigned responsibility.

5.1.2

Authority and Responsibility. The contractor shall give a charter to his maintainability organization. This charter shall be from top management. A copy of this charter shall be included in the maintainability program plan. The charter shall clearly give necessary authority to the maintainability organization and its management to enforce its policies and actions. The contractor shall identify specifically, and by organizational title, the individual who shall serve as the single-point contact for the Government in the area of maintainability. This contact must be a member of the contractor's maintainability organization.

5.1.3

Management Tasks. The Contractor's Maintainability Program Plan shall identify and define the essential tasks specified in pars. 5.2 thru 5.12 below, which shall include:

a. The work to be accomplished for each task, including inputs and outputs
b. The time phasing of each task
c. The organizational element responsible for implementing the task
d. Appropriate milestones for program review by the Government and contractor
e. Specific technique(s) for allocating quantitative requirements to lower level functional elements such as sub-system, assembly, subassembly, accessory, part, or component
f. Specific technique(s) for maintainability predictions of quantitative requirements at lower level functional elements, such as subsystem, assembly, subassembly, part, or component
g. Method by which maintainability requirements are disseminated to designers and associated personnel
h. Provisions for internal training and indoctrination in connection with this project.

5.1.4

Maintainability Interface Compatibility. The maintainability program shall be coordinated with other interfacing efforts such as those cited below to insure an integrated and effective contractual effort.

a. Logistic support and Inputs to the Detailed Maintainability Plan
   1. Maintenance requirements analysis
   2. Maintenance task analyses

3. Tool and test equipment determinations to include calibration equipment and calibration requirements
4. Manpower training and skill requirements determination
5. Maintenance information systems: i.e., technical data, training manuals, etc.
6. Support equipment/facilities determination.

b. Reliability Program (MIL-STD-785)
c. Human resources (personnel subsystems) including human engineering (MIL-H-46855 and MIL-STD-1472)
d. System life cycle cost estimates and cost-effectiveness studies
e. System engineering and system effectiveness analysis activities
f. Design Engineering
g. Value Engineering (MIL-V-38352)
h. Safety Engineering (MIL-STD-882)

5.2 ***Maintainability Analyses.*** The contractor shall perform a maintainability analysis of the **XYZ** system as an integral part of the overall **XYZ** system analysis. Primary input to the analysis shall be data obtained from his studies and engineering reports and the following additional data from the Government System Specifications or Description:

a. Operational and support concepts and requirements, including environmental conditions
b. Overall quantitative maintainability requirements
c. Personnel subsystem constraints
d. Projected facility, training program, skills, equipment, and tool availability
e. Cost constraints
f. Studies and engineering reports for the **XYZ** system
g. Lists of standard tools and equipment.

As a major task of the analysis, the contractor shall allocate quantitative maintainability requirements to all significant functional levels of the **XYZ** system. The analysis shall document trade-offs and the quantitative and qualitative requirements, which then become design criteria and are incorporated into specifications. The maintainability analysis shall also be used during design development and test to evaluate the degree of achievement of the maintainability design requirements.

5.3 ***Inputs to the Detailed Maintenance Concept and Detailed Maintenance Plan.*** The contractor shall, as a result of the maintainability analysis (par. 5.2 above) prepare inputs to a detailed maintenance concept for the **XYZ** system. The concept shall be based upon the operational and support concepts and requirements established in the System Specification. A detailed maintenance plan shall be prepared from the concept, based on the planned operational environment of the **XYZ** system as described in the system specification. The plan shall include but not be limited to:

a. Depth and frequency of maintenance requirements at each level
b. Facilities required
c. Support equipment and tools required
d. Skill levels and number of people required.

5.4 ***Maintainability Design Criteria.*** The contractor shall establish and periodically update detailed maintainability design criteria, determined from the **XYZ** system Maintainability analysis. These criteria shall be implemented by maintainability guidelines, techniques, and procedures previously developed and incorporated into Military/industrial handbooks. Appropriate consideration of maintainability design criteria by the contractor shall be reflected in design concept reviews, item selection, design reviews, and design trade-offs. Maintainability design criteria shall include but not be restricted to:

a. Reduction of maintenance complexity
b. Reduction of need and frequency of design-dictated maintenance activities
c. Reduction of maintenance downtime
d. Reduction of design dictated maintenance support costs
e. Limitation of maintenance personnel requirements
f. Reduction of potential for maintenance error.

5.5 ***Design Trade-offs.*** The contractor's maintainability plan shall indicate how maintainability is considered in all design trade-off analyses. This shall include the effects of maintainability on reliability, safety, and other disciplines and constraints. All analyses shall be documented and submitted as part of the maintainability status report, paragraph 5.12 below.

5.6 ***Maintainability Parameter Values.*** The contractor shall predict Maintainability values for the **XYZ** system. The prediction technique shall be based on Method 2 of MIL-HDBK-472 or a contractor formulated Method which can provide the same values. In all cases the technique used shall estimate quantitatively the **XYZ** system parameter values for the planned design configuration. The quantitative estimates shall be used to evaluate the adequacy of the proposed design to meet the maintainability quantitative requirements and identify design features requiring corrective action.

5.7 ***Maintainability Requirements in Subcontractor and Vendor Specifications.*** The contractor shall include in his maintainability program plan, appropriate quantitative maintainability requirements in all specifications for items which are procured from subcontractors and vendors or suppliers. The requirements shall be stipulated in terms which can be demonstrated in accordance with MIL-STD-471. The method to be selected shall take into consideration risk, cost, time, and validity of underlying assumptions. The method to be used for demonstration shall be selected by the contractor and approved by the Government prior to adoption. The specifications shall include, but not be restricted to the following:

a. System/equipment constraints and requirements
b. Maintenance concepts and support requirements
c. Standardization and interchangeability requirements
d. Maintainability demonstration requirements and procedures.

The program plan shall also provide means for evaluating the subcontractor's and vendor's maintainability program and methods to assure compliance by the subcontractors and/or vendor of all specific maintainability requirements, including corrective action as required.

5.8 ***Integration of Other Items.*** The contractor shall obtain maintainability parameters for all subsystems, assemblies, parts, and components furnished by the Government and all subcontractors and/or suppliers. Parameter values for Government Furnished Equipment shall be obtained from the Government, and values for other items shall be furnished by the respective subcontractors and/or suppliers. If the maintainability values are unavailable or unknown, the contractor shall estimate them. If the estimated or furnished values are incompatible with the **XYZ** system, or if analysis indicates that the **XYZ** system will not satisfy the operational or maintainability requirements when these values are used, the contractor shall identify problem areas, propose alternate courses of action or revised statements of requirements, and estimate values which shall allow the maintainability or operational requirements to be met. The contractor shall notify the Government if the maintainability values for GFE are incompatible with maintainability requirements of the system. Values approved by the Government shall be used to determine quantitative requirements to be entered into contract specifications, and as the basis for determining the contractor's compliance with quantitative maintainability requirements during the maintainability demonstration.

5.9 ***Design Reviews.*** The contractor shall conduct maintainability design reviews to assure the accomplishment of maintainability requirements of the **XYZ** system. Reviews shall be of two types: formal and informal, and shall be keyed to significant design milestones. The formal design reviews shall be indicated as milestone events on the maintainability program schedule submitted to the Government. The formal reviews shall be conducted as an integral part of the contractor's system engineering review and evaluation, and shall be documented with copies furnished to the Government. Notice of the formal review shall be given to the Government not less than ten (10) days in advance to permit attendance at these reviews. The informal reviews may consist of maintainability and engineering personnel, although more than likely they will include other personnel also. All design changes shall be reviewed and their effects on achievement of quantitative maintainability requirements evaluated. A formal review shall be held prior to release of drawings/specifications for production.

5.10 ***Data Collection, Analysis, and Corrective Action System.*** The contractor shall establish a maintainability data collection system for prediction during design and for evaluation of demonstration results. The system shall be a closed loop system which is integrated with other activities such as reliability. The contractor's existing data system shall be utilized with minimum changes necessary to meet the requirements that follow. The data collection system for maintainability prediction shall be defined as early as possible but not later than contract definition, and used during design. The data collection system for demonstration shall receive preliminary planning during contract definition and shall be finalized in the maintainability demonstration prior to testing. The contractor shall evaluate the data against maintainability quantitative and qualitative requirements, identify problems, recommend solutions, document corrective actions, and include data collected which proves the effectiveness of corrective action. The data collection formats used shall permit determination of maintainability during early design and maintainability demonstration. The system shall be compatible with the Government Maintenance Data Collection System so that as the vehicle enters the operational phase, transition to in-service reporting can be accomplished with minimum disturbance and maximum continuity of effort.

5.11 ***Achievement of Maintainability Requirements.*** The contractor shall demonstrate the achieved maintainability of the **XYZ** system. The demonstration shall be based on a maintainability demonstration plan prepared in accordance with MIL-STD-471, and submitted to the Government for approval prior to implementation. A report shall be issued upon completion of the formal demonstration. The demonstration plan shall be responsive to the program plan established by the requirements of this specification. The demonstration effort shall be integrated to the maximum extent possible with other system test requirements such as proof of design, breadboard, prototype, environmental production, and acceptance. The contractor shall update the maintainability parameter values obtained from maintainability analyses and predictions with the values obtained during the demonstration. All demonstrations held for contract compliance shall be conducted in the operational environment spelled out in the system specification.

5.12 ***Maintainability Status Reports.*** The contractor shall provide to the Government a Monthly Maintainability Status Report. This report shall provide a current accounting of required, allocated, predicted, and observed values for the **XYZ** system and its components, subassemblies, parts, and accessories. The report shall include a narrative and graphical treatment of trends, problem areas, and actions taken or proposed and effective date, both estimated and actual. The reports may be combined with other reports, provided the maintainability information is summarized in a separate section and all supporting information is cross-referenced.

6.0 *NOTES.*

6.1 ***Data Requirements.*** The selected data which shall be furnished by the contractor for each phase of the **XYZ** System Development Program are outlined in the Contractor Data Requirement List Form DD 1423 for each phase. The requirements are attached to the RFP for this system and form an integral part thereof.

INDEX

INDEX (Cont'd)

INDEX (Cont'd)

INDEX (Cont'd)

INDEX (Cont'd)

INDEX (Cont'd)

INDEX (Cont'd)

INDEX (Cont'd)

INDEX (Cont'd)

W

(AMCRD-TT)

FOR THE COMMANDER:

OFFICIAL:

S/ Harold

G. J. HAROLD
LTC, GS
Adjutant General

ROBERT L. KIRWAN
Brigadier General, USA
Chief of Staff

DISTRIBUTION:
Special

☆U.S. GOVERNMENT PRINTING OFFICE: 1976—0-211-372/5859A

ENGINEERING DESIGN HANDBOOKS

Available to DA activities from Letterkenny Army Depot, ATTN: AMXLE-ATD, Chambersburg, PA 17201. All other requestors--DOD, Navy, Air Force, Marine Corps, nonmilitary Government agencies, contractors, private industry, individuals, universities, and others—must purchase Handbooks from National Technical Information Service, Department of Commerce, Springfield, VA 22151. See Preface for further details and AMC policy regarding requisitioning of classified documents.

No. AMCP 706-	Title
100	Design Guidance for Producibility
104	Value Engineering
106	Elements of Armament Engineering, Part One, Sources of Energy
107	Elements of Armament Engineering, Part Two, Ballistics
108	Elements of Armament Engineering, Part Three, Weapon Systems and Components
109	Tables of the Cumulative Binomial Probabilities
110	Experimental Statistics, Section 1, Basic Concepts and Analysis of Measurement Data
111	Experimental Statistics, Section 2, Analysis of Enumerative and Classificatory Data
112	Experimental Statistics, Section 1, Planning and Analysis of Comparative Experiments
113	Experimental Statistics, Section 4, Special Topics
114	Experimental Statistics, Section 5, Tables
115	Environmental Series, Part One, Basic Environmental Concepts
116	Environmental Series, Part Two, Natural Environmental Factors
117	Environmental Series, Part Three, Induced Environmental Factors
118	Environmental Series, Part Four, Life Cycle Environments
119	Environmental Series, Part Five, Glossary of Environmental Terms
120	Criteria far Environmental Control of Mobile Systems
171	Packaging and Pack Engineering
123	Hydraulic Fluids
124	Reliable Military Electronics
125	Electrical Wire and Cable
127	Infrared Military Systems, Part One
128(S)	Infrared Military Systems, Part Two (U)
130	Design for Air Transport and Airdrop of Materiel
132	Maintenance Engineering Techniques (MET)
133	Maintainability Engineering Theory and Practice (METAP)
134	Maintainability Guide for Design
135	+Inventions, Patents, and Related Matters
13h	**Servomechanisms, Section 1, Theory
137	**Servomechanisms, Section 2, Measurement and Signal Converters
138	**Servomechanisms, Section 3, Amplification
139	**Servomechanisms, Section 4, Power Elements and System Design
140	Trajectories, Differential Effects, and Data for Projectiles
150	Interior Ballistics of Guns
158	Fundamentals of Ballistic Impact Dynamics, Part One
159(S)	Fundamentals of Ballistic Impact Dynamics, Part Two (U)
160(C)	Elements of Terminal Ballistics, Part One, Kill Mechanisms and Vulnerability (U)
161(C)	Elements of Terminal Ballistics, Part Two, Collection and Analysis of Data Concerning Targets (U)
162(SRD)	Elements of Terminal Ballistics, Part Three, Application to Missile and Space Targets (U)
163(S)	*Basic Target Vulnerability (U)
165	Liquid-Filled Projectile Design
170(S)	Armor and Its Applications (U)
175	Solid Propellants, Part One
176	+Solid Propellants, Part Two
177	Properties of Explosives of Military Interest
17R	+Properties of Explosives of Military Interest, Section 2 (Replaced by -177)
179	Explosive Trains
180	Principles of Explosive Behavior
181	Explosions in Air, Part One
182(SRD)	Explosions in Air, Part Two (U)
185	Military Pyrotechnics, Part One, Theory and Application
186	Military Pyrotechnics, Part Two, Safety, Procedures and Glossary
1H7	Military Pyrotechnics, Part Three, Properties of Materials Used in Pyrotechnic Compositions
188	Military Pyrotechnics, Part Four, Design of Ammunition for Pyrotechnic Effects
1H9	Military Pyrotechnics, Part Five, Bibliography
190	*Army Weapon System Analysis
191	System Analysis and Cost-Effectiveness
192	Computer Aided Design of Mechanical Systems, Part One
193	Computer Aided Design of Mechanical Systems, Part Two
195	*Development Guide for Reliability, Part One, Introduction, Background, and Planning for Army Materiel Requirements
196	*Development Guide for Reliability, Part Two, Design far Reliability
197	*Development Guide for Reliability, Part Three, Reliability Prediction
198	*Development Guide for Reliability, Part Four, Reliability Measurement
199	*Development Guide for Reliability, Part Five, Contracting for Reliability
2nn	*Development Guide for Reliability, Part Six, Mathematical Appendix and Glossary
201	Helicopter Engineering, Part One, Preliminary Design
202	*Helicopter Engineering, Part Two, Detail Design
203	Helicopter Engineering, Part Three, Qualification Assurance
204	Helicopter Performance Testing
205	Timing Systems and Components
210	Fuzes
211(C)	Fuzes, Proximity, Electrical, Part One (U)
212(S)	Fuzes, Proximity, Electrical, Part Two (U)
213(S)	Fuzes, Proximity, Electrical, Part Three (U)
214(S)	Fuzes, Proximity, Electrical, Part Four (U)
215(C)	Fuzes, Proximity, Electrical, Part Five (U)
z35	Hardening Weapon Systems Against RF Energy
237	*Mortar Weapon Systems
238	*Recoilless Rifle Weapon Systems
239	*Small Arms Weapon Systems
240(C)	Grenades (U)
242	Design for Control of Projectile Flight Characteristics (Replaces -246)
744	Ammunition, Section 1, Artillery Ammunition--General, with Table of Contents, Glossary, and Index for Series
245(C)	Ammunition, Section 2, Design for Terminal Effects (U)
246	+Ammunition, Section 3, Design for Control of Flight Characteristics (Replaced by -242)
247	Ammunition, Section 4, Design for Projection
24R	+Ammunition, Section 5, Inspection Aspects of Artillery Ammunition Design
249	Ammunition, Section h, Manufacture of Metallic Components of Artillery Ammunition
25'1	Guns--General
251	Muzzle Devices
252	**Gun Tubes
253	*Breech Mechanism Design
255	Spectral Characteristics of Muzzle Flash
260	Automatic Weapons
270	Propellant Actuated Devices
280	Design of Aerodynamically Stabilized Free Rockets
281(SRD)	Weapon System Effectiveness (U)
282	+Propulsion and Propellants (Replaced by -285)
2Rh	Structures
290(C)	Warheads--General (U)
291	+Surface-to-Air Missiles, Part One, System Integration
292	+Surface-to-Air Missiles, Part Two, Weapon Control
293	+Surface-to-Air Missiles, Part Three, Computers
294(C)	+Surface-to-Air Missiles, Part Four, Missile Armament (U)
295(C)	+Surface-to-Air Missiles, Part Five, Countermeasures (U)
29h	+Surface-to-Air Missiles, Part Six, Structures and Power Sources
297(C)	+Surface-to-Air Missiles, Part Seven, Sample Problem (U)
300	Fabric Design
312	Rotational Molding of Plastic Powders
313	Short Fiber Plastic Base Composites
327	Fire Control Systems--General
329	Fire Control Computing Systems
331	Compensating Elements
335(SRD)	*Design Engineers' Nuclear Effects Manual (DENEM), Volume I, Munitions and Weapon Systems (U)
336(SRD)	*Design Engineers' Nuclear Effects Manual (DENEM), Volume II, Electronic Systems and Logistical Systems (U)
337(SRD)	*Design Engineers' Nuclear Effects Manual (DENEM), Volume III, Nuclear Environment (U)
338(SRD)	*Design Engineers' Nuclear Effects Manual (DENEM), Volume IV, Nuclear Effects (U)
340	Carriages and Mounts — General
341	Cradles
342	Recoil Systems
343	[illegible]
344	Bottom Carriages
145	Equilibrators
346	Elevating Mechanisms
347	Traversing Mechanisms
350	Wheeled Amphibians
355	The Automotive Assembly
356	Automotive Suspensions
357	Automotive Bodies and Hulls
360	Military Vehicle Electrical Systems
361	Military Vehicle Power Plant Cooling
410	*Electromagnetic Compatibility (EMC)
411(S)	*Vulnerability of Communication-Electronic and Electro-Optical Systems (Except Guided Missiles) to Electronic Warfare, Part One, Introduction and General Approach to Electronic Warfare Vulnerability (U)
412(C)	*Vulnerability of Communication-Electronic and Electro-Optical Systems (Except Guided Missiles) to Electronic Warfare, Part Two, Electronic Warfare Vulnerability of Tactical Communications (U)
413(S)	*Vulnerability of Communication-Electronic and Electro-Optical Systems (Except Guided Missiles) to Electronic Warfare, Part Three, Electronic Warfare Vulnerability of Ground-Based and Airborne Surveillance and Target Acquisition Radars (U)
414(S)	*Vulnerability of Communication-Electronic and Electro-Optical Systems (Except Guided Missiles) to Electronic Warfare, Part Four, Electronic Warfare Vulnerability of Avionics (U)
415(S)	*Vulnerability of Communication-Electronic and Electro-Optical Systems (Except Guided Missiles) to Electronic Warfare, Part Five, Optical/Electronic Warfare Vulnerability of Electro-Optic Systems (U)
416(S)	*Vulnerability of Communication-Electronic and Electro-Optical Systems (Except Guided Missiles) to Electronic Warfare, Part Six, Electronic Warfare Vulnerability of Satellite Communications (U)
445	Sabot Technology Engineering
470	*Metric Conversion Guide for Military Applications

*UNDER REVISION--not available
**REVISION UNDER PREPARATION
+OBSOLETE--out of stock