UNCLASSIFIED

AD NUMBER

AD875609

NEW LIMITATION CHANGE

TO

Approved for public release, distribution unlimited

FROM

Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; Jul 1970. Other requests shall be referred to US Army Research Office, Durham, NC 27706.

AUTHORITY

US Army Research Office ltr, 15 Mar 1971

THIS PAGE IS UNCLASSIFIED

AD875609

ARO-D Report 70-2

PROCEEDINGS OF THE FIFTEENTH CONFERENCE ON THE DESIGN OF EXPERIMENTS IN ARMY RESEARCH DEVELOPMENT AND TESTING

AD No. ____
DDC FILE COPY



DDC
OCT 12 1970

This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of the U. S. Army Research Office—Durham, Durham, North Carolina.

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

Sponsored by
The Army Mathematics Steering Committee
on Behalf of
THE OFFICE OF THE CHIEF OF RESEARCH AND DEVELOPMENT

U. S. ARMY RESEARCH OFFICE-DURHAM

Report No. 70-2
July 1970

PROCEEDINGS OF THE FIFTEENTH CONFERENCE

ON THE DESIGN OF EXPERIMENTS

Sponsored by the Army Mathematics Steering Committee

Host

U. S. Army Missile Command
Redstone Arsenal, Alabama

22-24 October 1969

This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of the U. S. Army Research Office-Durham, Durham, North Carolina. 27706

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

U. S. Army Research Office-Durham
Box CM, Duke Station
Durham, North Carolina

REPRODUCTION QUALITY NOTICE

This document is the best quality available. The copy furnished to DTIC contained pages that may have the following quality problems:

- Pages smaller or larger than normal.
- Pages with background color or light colored printing.
- Pages with small type or poor printing; and or
- Pages with continuous tone material or color photographs.

Due to various output media available these conditions may or may not cause poor legibility in the microfiche or hardcopy output you receive.

☑ If this block is checked, the copy furnished to DTIC contained pages with color printing, that when reproduced in Black and White, may change detail of the original copy.

FOREWORD

In a letter under the date of 2 November 1967, Dr. John L. McDaniel, Technical Director of the Research and Engineering Directorate at the U. S. Army Missile Command (MICOM), offered to hold the Fourteenth Conference on the Design of Experiments in Army Research, Development and Testing at his installation. Since arrangements were already underway to hold this conference in the Washington area, this invitation had to be declined by the Army Mathematics Steering Committee (AMSC), the sponsor of this series of conferences. Dr. McDaniel, when made aware of this situation, was willing for the Committee to treat his request to hold the conference as a standing invitation. Members of the AMSC were very pleased to hear this and then discussed with him the possibility of holding the Fifteenth Conference at Redstone Arsenal. These negotiations were brought to a successful conclusion; and, on 29 November 1968, Major General Charles W. Eifler issued a formal invitation to host this conference at his command on 22-24 October 1969. He appointed Dr. Siegfried Lehnigk to serve as Chairman on Local Arrangements and Mr. Raymond V. Knox to handle administrative requirements.

MICOM had already served as the host to the Ninth Conference in this series. It is interesting to note that Dr. Lehnigk, as well as Henry A. Dihm, and W. H. Ewart served as members of the Local Arrangements Committee for the Ninth Conference, as well as the Fifteenth Conference. Those in attendance at this 22-24 October meeting are much in debt to these gentlemen, as well as to many others at Redstone Arsenal, for the excellent handling of the many details connected with a meeting of this size.

Among the many highlights of the Fifteenth Conference on the Design of Experiments was the banquet talk given by Professor Oskar Morgenstern of Princeton University and the following invited speakers:

Reliability Applied to Space Flight
Dr. John E. Condon, National Aeronautics and Space Administration

Systems Reliability
Dr. Nancy R. Mann, Rocketdyne

A Probability Approach to Catastrophic Threat
Dr. Clifford J. Maloney, National Institutes of Health

The Empirical Bayes Approach to the Design and Analysis of Experiments
Professor Richard G. Krutchkoff, Virginia Polytechnic Institute

On Confidence Limits for the Performance of a System When Few Failures are Encountered
Dr. S. C. Saunders, Boeing Scientific Research Laboratories

Everyone had the opportunity to hear the above-mentioned talks, as they were given in general sessions. Unfortunately, one was not privileged to hear all of the thirty-two contributed papers. These covered a wide range of interesting statistical problems and had to be scheduled so that three talks were conducted simultaneously. Following the banquet, it was my privilege to award the Fifth Samuel S. Wilks Memorial Medal, sponsored by the American Statistical Association and the Army, to Dr. W. J. Youden. Details of this presentation are included in these Proceedings.

This conference was attended by 156 scientists; and 52 organizations were represented. Speakers and panelists came from: Boeing Scientific Research Laboratories; Cornell University; Honeywell, Inc.; Litton Systems, Inc.; National Aeronautics and Space Administration; National Institutes of Health; North Carolina State University; Princeton University; Rocketdyne; University of Alabama; University of Georgia; University of Michigan; University of Wisconsin; Vanderbilt University; Virginia Polytechnic Institute; and, 12 Army facilities.

Members of the AMSC would like to express their thanks to the many speakers, chairmen and panelists for all their efforts in behalf of this important scientific meeting. Most of the papers presented at the conference are being made available to the public through these Proceedings. The AMSC asked that copies of this manual receive wide distribution among Army laboratories and Technical Libraries.

At this time, let me express my appreciation to all members of the Program Committee (Clifford Cohen, Jr., Henry Dihm, Francis Dressel, Walter Foster, Fred Frishman, Bernard Harris, Boyd Harshbarger, Raymond Knox, Siegfried Lehnigk, H. L. Lucas, Clifford Maloney, and Herbert Solomon) for their many suggestions and advice on the selection of the speakers and the organization of the whole conference.

Frank E. Grubbs
Conference Chairman

TABLE OF CONTENTS

Title	Page
Foreword	iii
Table of Contents	v
Program	ix
Reliability Applied to Space Flight John E. Condon	1
Computer-Aided Solution of Prior Distributions for Generating Monte Carlo Confidence Bounds on System Reliability Nancy R. Mann	7
Comparative Analysis of the LCSS-ETG-3 Performance Compability Using Statistical Probabilities Andrew H. Jenkins	27
Estimation of Vehicle Parameters for the Given Model: $Y = \theta_1 e^{\theta 2t} \sin(\theta_3 t)$ John Howerton and D. Ray Campbell	75
A Method of Improving the Estimation of Variance John Gurland and J. S. Mehta	97
A Classification of Bivariate Variance Comparisons Clifford J. Maloney	113
Computerized Quality Control as Applied to Upper Atmospheric Data Oskar M. Essenwanger	123
A Statistical Model for the Analysis of Simultaneous Two-Station Ionospheric Soundings Erwin Biser and Richard D'Accardi	147
Position Location Via Multiple Triangulation G. A. Stoops and E. L. Spitznagel, Jr.	189
A General Computational Algorithm for Bayesian Confidence Bounds R. W. Clarke	197
Exact Lower Confidence Limits on Normal and Lognormal Reliability R. W. Soanes	205

Real-Time Simulation Technique for Evaluating a Gyro-Seeker Assembly
Elwood D. Bias 215

An Empirical Approach to Analysis of the Interaction Characteristics of a Six-Component Rocket Engine Test Stand
Aubrey W. Presson. 243

Interlaboratory Study of a Method for Measuring Ammonium Perchlorate Particle Size
Bernard J. Alley 265

New Analyses and Methods Leading to Improved Target Acquisition Requirements Involving Systems, Geodetic and Re-entry Errors, and Increased Weapons Effectiveness for Conventional Weapons
Hans Baussus von Luetzow 279

An Air Defense Comparative Model
R. E. Shannon, J. P. Ignizio and J. L. Stimach 293

Probabilistic Manpower Planning for the Research and Development Organization
Larry H. Johnson 315

Analysis of Factorial Arrangement in Non-connected Block Design
Sadrig Kurkjian and R. C. Woodall. 337

Design of Field Test Programs and Statistical Techniques for Analysis of the Performance of Navigation and Positioning Systems
Emil H. Jebe and Ralph A. King 349

A Unified Procedure for Selecting Alternate Experimental Designs
Edwin M. Bartee. 373

A Problem in Continuous Sampling Verification
Mary E. Blome. 403

Toward a Stochastic Model of Terrain
R. H. Peterson and W. C. Taylor. 423

A Suggested Procedure for Analyzing Missile Performance by a Least Squares Fit to a Generalized Linear Statistical Model and a Quick Check for Normality of the Data
Nancy R. Rich. 437

A Use of Reliability Techniques in Army Experiments
D. R. Barr and T. Jayachandran 495

Optimizing a Four-Part Assay Procedure
Walter D. Foster 509

An Application of Linear Programming to Experimental Design
J. Richard Moore 515

Transmission of Infrasonic Waves Generated by Large Missile Launches
Raymond E. Lacy and C. E. Sharp 523

Determining the Flight Reliability of an Antitank Missile with Side Jets
Robert G. Conard and Nancy R. Rich 535

Empirical Bayes and the Design and Analysis of Experiments
Richard G. Krutchkoff. 579

Youden Awarded the Samuel S. Wilks Memorial Medal
Frank E. Grubbs. 599

The Use of a Hybrid Computer to Evaluate Man-Machine Performances of Complex Vehicle Control Systems
Myrna L. Toivanen, Bernard S. Gurman and Erwin Biser 607

Experimental Design Considerations in Validating a Method of Modeling a Man-Organized System
B. B. Lukens and R. A. Brown 653

An Investigation of the Effect of Some Prior Distributions on Bayesian Confidence Intervals for Attribute Data
Alan W. Benton 663

Some Techniques for Constructing Mutually Orthogonal Latin Squares
W. T. Federer, et al. 673

On Confidence Limits for the Performance of a System When Few Failures are Encountered
S. C. Saunders 797

A Probability Model for the Assessment of Human Incapacitation from Penetrating Missile Wounds
William P. Johnson and William J. Bruchey, Jr. 835

List of Attendees 855

FIFTEENTH CONFERENCE ON THE DESIGN OF EXPERIMENTS

IN ARMY RESEARCH, DEVELOPMENT AND TESTING

22-24 October 1969

Wednesday, 22 October

0800-0830 REGISTRATION - Lobby of Rocket Auditorium, Building 7120

0830-0845 OPENING OF THE CONFERENCE - Rocket Auditorium

Dr. Siegfried Lehnigk, Chairman on Local Arrangements

WELCOME

0845-1130 GENERAL SESSION I - Rocket Auditorium

Chairman: Dr. Walter D. Foster, Biomathematics Division, Biological Laboratories, Ft. Detrick, Frederick, Maryland

RELIABILITY APPLIED TO SPACE FLIGHT

Dr. John E. Condon, Reliability and Quality Assurance, National Aeronautics and Space Administration, Washington, D. C.

SYSTEMS RELIABILITY

Dr. Nancy R. Mann, Rocketdyne, Canago Park, California

1130-1330 LUNCH - Officer's Club

1330-1515 TECHNICAL SESSION 1 - Conference Room, Building 7101

Chairman: John S. Hagen, U. S. Army Development and Proof Services, Aberdeen Proving Ground, Maryland

DEVELOPMENT OF TESTING PROGRAMS TO MINIMIZE OVERALL PROJECT COST OR FAILURE PROBABILITY

Roger L. Lapp, Corps of Engineers, Huntsville, Alabama

COMPARATIVE ANALYSIS OF THE LCSS-ETG-3 PERFORMANCE CAPABILITY USING STATISTICAL PROBABILITIES

Andrew H. Jenkins, U. S. Army Missile Command, Redstone Arsenal, Alabama

Preceding page blank

1300-1515 TECHNICAL SESSION 2 - Rocket Auditorium

Chairman: Joseph Weinstein, Electronics Computing Laboratory, U. S. Army Electronics Command, Fort Monmouth, New Jersey

ESTIMATION OF VEHICLE PARAMETERS FOR THE GIVEN MODEL:

$Y = \theta_1 e^{\theta_2 t} \sin(\theta_3 t)$

John Howerton and D. Ray Campbell, Systems Evaluation Branch, Advanced Systems Laboratory, Research and Engineering Directorate, Redstone Arsenal, Alabama

A METHOD OF IMPROVING THE ESTIMATION OF VARIANCE

John Gurland, Department of Statistics, University of Wisconsin. Sponsored by the Mathematics Research Center, U. S. Army, University of Wisconsin, Madison, Wisconsin

A CLASSIFICATION OF BIVARIATE VARIANCE COMPARISONS

Clifford J. Maloney, Biometrics Section, Department of Health, Education and Welfare, National Institutes of Health, Bethesda, Maryland

1330-1515 TECHNICAL SESSION 3 - Conference Room, Building 7101

Chairman: Eugene F. Dutoit, Quality Assurance Directorate, U. S. Army Munitions Command, Picatinny Arsenal, Dover, N. J.

COMPUTERIZED QUALITY CONTROL AS APPLIED TO UPPER ATMOSPHERIC DATA

Oskar M. Essenwanger, Aerophysics Branch, Physical Sciences Laboratory, Research and Engineering Directorate, Redstone Arsenal, Alabama

A STATISTICAL MODEL FOR THE ANALYSIS OF SIMULTANEOUS TWO STATION IONOSPHERIC SOUNDINGS

Erwin Biser and Richard D'Accardi, Avionics Laboratory, U. S. Army Electronics Command, Fort Monmouth, New Jersey

POSITION LOCATION VIA MULTIPLE TRIANGULATION

G. A. Stoops and E. L. Spitznagel, Jr., Math Sciences Section, Litton Scientific Support Laboratory, Litton Systems, Inc., Fort Ord, California

1515-1545 BREAK

1545-1700 CLINICAL SESSION A - Rocket Auditorium

Chairman: Henry Dihm, Advanced Systems Laboratory, Directorate of Research and Development, U. S. Army Missile Command, Redstone Arsenal, Alabama

Panelists:
Robert Bechhofer, Cornell University
O. P. Bruno, U. S. Army Ballistics Research & Development Center
A. C. Cohen, University of Georgia
Bernard Harris, Mathematics Research Center, U. S. Army
Boyd Harshbarger, Virginia Polytechnic Institute
H. L. Lucas, North Carolina State University

AN EMPIRICAL APPROACH TO ANALYSIS OF THE INTERACTION CHARACTERISTICS OF A SIX-COMPONENT ROCKET ENGINE TEST STAND

Aubrey W. Presson, Test Research and Analysis Branch, Test and Research and Engineering Directorate, U. S. Army Missile Command, Redstone Arsenal, Alabama

INTERLABORATORY STUDY OF A METHOD FOR MEASURING AMMONIUM PERCHLORATE PARTICLE SIZE

Bernard J. Alley, U. S. Army Missile Command, Redstone Arsenal, Alabama

1545-1700 TECHNICAL SESSION 4 - Conference Room

Chairman: William McIntosh, U. S. Army Test and Evaluation Command, Aberdeen Proving Ground, Maryland

A GENERAL COMPUTATIONAL ALGORITHM FOR BAYESIAN CONFIDENCE BOUNDS

R. W. Clarke, U. S. Army Watervliet Arsenal, Watervliet, New York

CONFIDENCED NORMAL AND LOGNORMAL RELIABILITY FOR ANY SAMPLE SIZE

R. W. Soanes, U. S. Army Watervliet Arsenal, Watervliet, New York

1545-1700 TECHNICAL SESSION 5 - Control Room

Chairman: Raymond B. Schnell, U. S. Army Advanced Materiel Concepts Agency, Washington, D. C.

REAL-TIME SIMULATION TECHNIQUE FOR EVALUATING A GYRO-SEEKER ASSEMBLY

Elwood D. Baas, White Sands Missile Range, White Sands, New Mexico

BIOCHEMICAL ASPECTS OF FEEDBACK EFFECTS IN BIOCELLULAR SYSTEMS

George I. Lavin, Terminal Ballistic Laboratory, U. S. Army Ballistic Research and Development Center, Aberdeen Proving Ground, Maryland

Thursday, 23 October

0830-1015 CLINICAL SESSION B - Rocket Auditorium

Chairman: David Howes, Strategy and Tactics Analysis Group, Bethesda, Maryland

Panelists:
Robert Bechhofer, Cornell University
O. P. Bruno, U. S. Army Ballistics Research & Development Center
A. C. Cohen, University of Georgia
Bernard Harris, Mathematics Research Center, U. S. Army
Boyd Harshbarger, Virginia Polytechnic Institute
H. L. Lucas, North Carolina State University

A PROBLEM IN CONTINUOUS SAMPLING VERIFICATION

Mary E. Blome, U. S. Army Ammunition Procurement and Supply Agency, Joliet, Illinois

TOWARD A STOCHASTIC MODEL OF TERRAIN

R. H. Peterson and W. C. Taylor, Army Materiel Systems Analysis Agency, Aberdeen Research and Development Center, Aberdeen Proving Ground, Maryland

0830-1015 TECHNICAL SESSION 6 - Conference Room

Chairman: Robert G. Stimson, Air Defense Systems Group, Office of the Chief of Staff, Washington, D. C.

NEW ANALYSES AND METHODS LEADING TO IMPROVED TARGET ACQUISITION REQUIREMENTS INVOLVING SYSTEMS, GEODETIC AND RE-ENTRY ERRORS AND INCREASED WEAPONS EFFECTIVENESS FOR CONVENTIONAL WEAPONS, PART I

Hans Baussus von Luetzow, U. S. Army Topographic Laboratories, Fort Belvoir, Virginia

AIR DEFENSE SYSTEMS COMPARATIVE MODEL

R. E. Shannon, J. P. Ignizio, and J. L. Stimach, Research Institute, University of Alabama, Huntsville, Alabama. Sponsored by the U. S. Army Missile Command, Redstone Arsenal, Alabama

PROBABILISTIC MANPOWER PLANNING FOR THE RESEARCH AND DEVELOPMENT ORGANIZATION

Larry H. Johnson, Research and Engineering Directorate, U. S. Army Missile Command, Redstone Arsenal, Alabama

0830-1015 TECHNICAL SESSION 7 - Control Room

Chairman: Bruce C. Gray, Biomathematics Division, U. S. Army Biological Laboratories, Fort Detrick, Frederick, Maryland

ANALYSIS OF FACTORIAL ARRANGEMENT IN DISCONNECTED BLOCK DESIGNS

Badrig Kurkjian and R. C. Woodall, U. S. Army Materiel Command, Harry Diamond Laboratories, Washington, D. C.

DESIGN OF FIELD TEST PROGRAMS AND STATISTICAL TECHNIQUES FOR ANALYSIS OF THE PERFORMANCE OF NAVIGATION AND POSITIONING SYSTEMS

Emil H. Jebe, Institute of Science and Technology, The University of Michigan, Ann Arbor, Michigan; and,

Ralph A. King, Department of Industrial Engineering, University of Wisconsin, Madison, Wisconsin

0830-1015 TECHNICAL SESSION 7 (Continued)

A UNIFIED PROCEDURE FOR SELECTING ALTERNATE EXPERIMENTAL DESIGNS

Edwin M. Bartee, Center for Engineering Management Studies, Vanderbilt University, Nashville, Tennessee. Sponsored by the U. S. Army Missile Command, Redstone Arsenal, Alabama

1015-1045 BREAK

1045-1200 CLINICAL SESSION C - Rocket Auditorium

Chairman: Frank E. Grubbs, U. S. Army Aberdeen Research and Development Center, Aberdeen Proving Ground, Maryland

Panelists:
Robert Bechhofer, Cornell University
O. P. Bruno, U. S. Army Ballistics Research & Development Center
A. C. Cohen, University of Georgia
Bernard Harris, Mathematics Research Center, U. S. Army
Boyd Harshbarger, Virginia Polytechnic Institute
H. L. Lucas, North Carolina State University

DETERMINING THE RELIABILITY OF AN ANTITANK MISSILE WITH SIDE THRUSTERS

R. G. Conard and N. R. Rich, Systems Evaluation Branch, Advanced Systems Laboratory, Research and Engineering Directorate, Redstone Arsenal, Alabama

TRANSMISSION OF INFRASONIC WAVES GENERATED BY LARGE MISSILE LAUNCHES

Raymond E. Lacy and C. E. Sharp, Acoustic/Seismic Communications Research Area, Institute for Exploratory Research, U. S. Army Electronics Command, Fort Monmouth, N. J.

1045-1200 TECHNICAL SESSION 8 - Conference Room

Chairman: Carol D. Rose, Design of Experiments Branch, U. S. Army Tank-Automotive Command, Warren, Michigan

A SUGGESTED PROCEDURE FOR ANALYZING MISSILE PERFORMANCE BY A LEAST SQUARES FIT TO A GENERALIZED LINEAR STATISTICAL MODEL AND A QUICK CHECK FOR NORMALITY OF THE DATA

Nancy R. Rich, Systems Evaluation Branch, Advanced Systems Laboratory, Research and Engineering Directorate, Redstone Arsenal, Alabama

1045-1200 TECHNICAL SESSION 8 (Continued)

A USE OF RELIABILITY TECHNIQUES IN ARMY EXPERIMENTS

D. R. Barr and T. Jayachandran, Mathematical Sciences Section, Litton Scientific Support Laboratory, Litton Systems, Inc., Fort Ord, California

1045-1200 TECHNICAL SESSION 9 - Control Room

Chairman: Gerhard J. Isaac, Statistics Branch, U. S. Army Medical Research and Nutrition Laboratory, Fitzsimons General Hospital, Denver, Colorado

OPTIMIZING A FOUR-PART ASSAY PROCEDURE

Walter D. Foster, Biomathematics Division, U. S. Army Biological Laboratories, Fort Detrick, Maryland

AN APPLICATION OF LINEAR PROGRAMMING TO EXPERIMENTAL DESIGN

J. Richard Moore, U. S. Army Aberdeen Research and Development Center, Aberdeen Proving Ground, Maryland

1200-1300 LUNCH - Cafeteria - Building 7101

1300-1500 GENERAL SESSION II - Rocket Auditorium

Chairman: Professor Robert M. Thrall, Department of Mathematical Sciences, Rice University, Houston, Texas

A PROBABILITY APPROACH TO CATASTROPHIC THREAT

Dr. Clifford J. Maloney, Biometric Section, Department of Health, Education and Welfare, National Institutes of Health, Bethesda, Maryland

THE EMPIRICAL BAYES APPROACH TO THE DESIGN AND ANALYSIS OF EXPERIMENTS [Part 1]

Professor Richard G. Krutchkoff, Department of Statistics, Virginia Polytechnic Institute, Blacksburg, Virginia

1500-1530 BREAK

THE EMPIRICAL BAYES APPROACH TO THE DESIGN AND ANALYSIS OF EXPERIMENTS [Part 2]

Professor Richard G. Krutchkoff

1830-1930 SOCIAL HOUR - Officer's Club

1930- BANQUET

Master of Ceremonies: Dr. John L. McDaniel, U. S. Army Missile Command, Redstone Arsenal, Alabama

Presentation of the Samuel S. Wilks Memorial Award by Dr. Frank E. Grubbs, U. S. Army Aberdeen Research and Development Center, Aberdeen Proving Ground, Maryland

Banquet Speaker:

Professor Oskar Morgenstern, Princeton University, Princeton, New Jersey

Friday, 24 October

0830-0900 TECHNICAL SESSION 10 - Rocket Auditorium

Chairman: Gideon A. Culpepper, Quality Control Division, White Sands Missile Range, New Mexico

THE USE OF A HYBRID COMPUTER TO EVALUATE MAN-MACHINE PERFORMANCE OF COMPLEX VEHICLE CONTROL SYSTEMS

M. L. Toivanen, Honeywell, Inc.; Bernard S. Gurman and Erwin Biser, U. S. Army Electronics Command, Fort Monmouth, New Jersey

0830-0900 TECHNICAL SESSION 11 - Conference Room

Chairman: Badrig Kurkjian, U. S. Army Materiel Command, Harry Diamond Laboratories, Washington, D. C.

EXPERIMENTAL DESIGN CONSIDERATIONS IN VALIDATING A METHOD OF MODELING A MAN-ORGANIZED SYSTEM

B. B. Lukens and R. A. Brown, Research Institute, University of Alabama, Huntsville, Alabama. Sponsored by the U. S. Army Missile Command, Redstone Arsenal, Alabama

0830-0900 TECHNICAL SESSION 12 - Control Room

Chairman: Alan S. Galbraith, Mathematics Division, Army Research Office-Durham, Durham, North Carolina

AN INVESTIGATION OF THE EFFECT OF SOME PRIOR DISTRIBUTIONS ON BAYESIAN CONFIDENCE INTERVALS FOR ATTRIBUTE DATA

Alan W. Benton, Surveillance and Reliability Division, Aberdeen Research and Development Center, Aberdeen Proving Ground, Maryland

0905-0930 Rocket Auditorium

OPEN MEETING OF THE AMSC SUBCOMMITTEE ON PROBABILITY AND STATISTICS

Chairman: Dr. Walter D. Foster, Biomathematics Division, U. S. Army Biological Laboratories, Fort Detrick, Maryland

0930-1000 BREAK

1000-1200 GENERAL SESSION III - Rocket Auditorium

Chairman: Dr. Joseph Bluhm, U. S. Army Mechanics and Materiel Research Center, Watertown, Massachusetts

TECHNIQUES FOR CONSTRUCTING MUTUALLY ORTHOGONAL LATIN SQUARES

Professor W. T. Federer, Department of Plant Breeding and Biometry, Cornell University, Ithaca, New York. Presently at the Mathematics Research Center, U. S. Army, University of Wisconsin, Madison, Wisconsin

ON CONFIDENCE LIMITS FOR THE PERFORMANCE OF A SYSTEM WHEN FEW FAILURES ARE ENCOUNTERED

Dr. S. C. Saunders, Boeing Scientific Research Laboratories, Seattle, Washington

CLOSING OF THE CONFERENCE

Dr. Frank E. Grubbs, Chairman of the Conference

1200-1300 LUNCH - Cafeteria - Building 7101

1300-1500 TOUR

****

ALTERNATE PAPER

The following paper will be presented if there is a cancellation in the program:

A PROBABILITY MODEL FOR THE ASSESSMENT OF HUMAN INCAPACITATION FROM PENETRATING MISSILE WOUNDS

William P. Johnson and William J. Bruchey, Jr., Vulnerability Laboratory, Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland

****

PROGRAM COMMITTEE

Clifford Cohen, Jr.
Henry Dihm
Francis Dressel (Secretary)
Walter D. Foster
Fred Frishman
Boyd Harshbarger
Bernard Harris
Raymond Knox
Siegfried Lehnigk
H. L. Lucas
Clifford Maloney
Herbert Solomon
Frank E. Grubbs (Chairman)

RELIABILITY APPLIED TO SPACE FLIGHT

John E. Condon
NASA Headquarters
Washington, D. C.

This month, October, marks the eleventh anniversary of NASA. Reflecting on NASA's accomplishments during the past eleven years, I feel we can point with pride to an outstanding record of success. Our record of mission success during these eleven years is over 75%, topped by a manned flight record of outstanding success in the Mercury, Gemini and Apollo programs.

The superlatives have been exhausted in describing the success and significance of Apollo - particularly Apollo 11. I suspect that many of you are keenly interested in knowing how we have attained the level of reliability so vital to the success of the Apollo program. I have given a great deal of thought to this subject during the past three months and regretfully - though not unexpectedly - have not found a simple, concise answer to this question. There are many factors which have contributed to the reliability of Apollo and thus it is not possible to single out any one factor as being all encompassing. However, there are two areas which, in my view, are worthy of special attention:

1. major attention by top management to the reliability of Apollo hardware; and,

2. emphasis, through all phases of the program, on the engineering aspects of reliability.

I will devote my remarks to the latter of these two points following some brief comments on the former.

The effective attainment of reliable space hardware requires the attention of all members of program/project team coupled with strong management support. This has been a key factor in the success of Apollo as top management has actively participated in key milestone reviews which are so important to the successful performance of the system. To illustrate this point, the following are examples of key Apollo milestone reviews.

> Critical Design Review. The purpose of this review is to formally review the design of the Contract End Item when the design is essentially complete. The review is intended to precede the release of engineering for manufacture. Among other things, this review established the integrity of the design by review of analytical and test data, and reliability apportionment and analysis available at that particular point in time.

Certification of Flight Worthiness. The purpose of this milestone is to certify that each flight stage and module is a complete and qualified item of hardware prior to shipment and is accompanied by adequate and accurate supporting documentation. Through this review the Apollo Program Director is informed of any deficiencies prior to shipment of the stage or module. This review certifies, for example, that:

1. acceptance, qualification and reliability tests have been successfully completed and meet the specification requirements;

2. departures from specification and drawing requirements have been approved by Material Review Boards;

3. critical hardware failures have been analyzed and corrected.

Flight Readiness Review (FRR). This is a two part review scheduled for each mission by a joint letter signed by the Program Director and the Mission Director. The purpose of the Program Director's FRR is to determine that the space vehicle hardware and launch complex are ready to commence the mission period. This includes consideration of the checkout and qualification status of all hardware, the summary of failures and disposition thereof, with particular emphasis on failures that have occurred during the pre-launch and checkout phase, and all modifications, deviations and waivers. The purpose of the Mission Director's FRR is to make a judgment for initiating the mission period and committing the deployment of world-wide forces to support the mission. Upon satisfactory completion of the Flight Readiness Review the mission period will commence.

The active participation of top management in these reviews gives emphasis to their importance, helps ensure that all factors which influence the successful performance of the hardware have received proper attention, and results in a "team" approach to system reliability.

The nature of NASA systems - highly complex, small quantity, R&D systems - requires that we concentrate on the engineering aspects of reliability rather than the analytical aspects, particularly at the system and major subsystem levels. In this regard, I would like to discuss the following:

1. adequacy of design for mission requirements;

2. identification and control of failure modes;

3. testing; and,

4. identification and correction of all failures.

We place heavy emphasis on the design review function and require our contractors, as part of their reliability program, to have a design review program. Contractors are required to establish and conduct a formal program of planned, scheduled and documented design reviews at the system, subsystem and component levels. These reviews are comprehensive critical audits of all pertinent aspects of the design of the hardware and software and are conducted at major program milestones beginning in the feasibility stage. Participation in these design reviews should be inter-organizational including competent personnel from such areas as design, fabrication, test, reliability assurance, quality assurance, and parts applications. In this way, inter-disciplinary engineering competence is brought to bear on all aspects of hardware design so as to identify and eliminate potential problems. NASA personnel may participate in these design reviews as deemed necessary. Each design review must be documented and the contractor's reliability organization is responsible for follow-up action to ensure that all recommendations are satisfactorily completed. An effective design review program pays high dividends through the early identification and elimination of problems which would manifest themselves at a later time when correction may be more costly.

Also, as an integral part of the early design phase, we require the contractor to develop analyses to determine possible modes of failure and their effects on mission objectives and crew safety. These analyses are conducted at the system, subsystem and component levels. Each potential failure is considered in terms of its probability of occurrence and is categorized as to probable effect on mission success; e.g., loss of life of crew member, mission termination, launch scrub or delay, etc. These analyses, generally referred to as Failure Mode, Effect and Criticality Analyses (FMEA) have the following important applications:

1. determining the need for redundancy, fail-safe design and derating;

2. determining the need to select parts and components of higher reliability;

3. identifying single failure points and reducing such to acceptable levels of risk;

4. supporting reliability predictions and assessments;

5. supporting system safety and hazard analyses;

6. assuring that test programs are responsive to known and suspected potential failure modes;

7. establishing allowing operating times or cycles; and,

8. determining operational contingency plans.

Of particular importance in our manned flight program is the use of FMEA's to identify single failure points which could adversely effect crew safety and mission objectives.

NASA places strong emphasis on testing throughout all phases of hardware development and fabrication. We require the contractor to develop an integrated test program which will evaluate all aspects of system performance capability to the extent practical. In terms of reliability considerations we expect the testing program to be directed towards:

1. verifying the capability of the design;

2. evaluating the susceptibility of the design and hardware to failures;

3. identifying unexpected interactions among components and assemblies;

4. identifying failure modes which reflect defects in materials, workmanship and fabrication processes; and,

5. obtaining failure rate and other reliability data.

To the extent practical, tests are planned using statistical design-of-experiment techniques and are conducted under environmental conditions and for time periods commensurate with mission conditions.

The final area to be discussed is that of failure reporting and corrective action. We expect all failures and nonconformances to be identified, analyzed and effective correction action taken - we cannot tolerate unexplained failures or ineffective corrective action in our space programs. We specifically require our contractors to employ a controlled system for identification, reporting, analysis, correction and prevention of recurrence of all nonconformances and suspected non-conformance of a functional nature which occur throughout the contract period. Some of the requirements which the system must satisfy are as follows:

1. it shall cover hardware, certain software, the interfaces between hardware and software and the interfaces between hardware or software and test or operational personnel;

2. it shall cover all nonconformances or suspected nonconformances of a functional nature such as:

a. unusual condition occurring in test or handling which are suspected to have an effect on the hardware;

b. transient malfunctions and suspected malfunctions; and,

c. notable deviations from previous performance - parameter drift.

3. It shall provide for investigation of each reported failure by an engineering analyses, followed, where appropriate, by laboratory analysis of failed hardware. Such investigation shall be adequate to assess causes, mechanisms, and potential effects of the failure and serve as a basis for decisions on the most efficient remedial and preventive actions;

4. it shall provide for a review of the technical closeout decision on each reported failure by higher levels of technical management commensurate with the criticality category of the failure involved; and,

5. closeout action shall be considered complete when:

a. remedial actions have been accomplished;

b. necessary preventive design and software changes have been devised and accomplished;

c. necessary design or computer program changes have been verified in test;

d. effectivity of preventive actions have been established;

e. change has been made in existing identical items of hardware to which the change is pertinent; and,

f. closeout documentation has been signed by proper management authority.

Such a system may seem unnecessarily extensive but experience has shown that it is necessary and pays high dividends.

In conclusion, I would like to point out that a significant portion of our reliability problems are due to nonelectronic parts and components. Such items as valves, fittings, seals, actuators, etc., continue to receive major attention as we strive to attain the levels of reliability necessary for mission success.

As we look to the future we will be striving to decrease, significantly, our cost per pound of payload, the complexity of our systems will continue to increase and, thus, our need for strong emphasis on the engineering aspects of reliability will not abate.

COMPUTER-AIDED SELECTION OF PRIOR DISTRIBUTIONS FOR GENERATING MONTE CARLO CONFIDENCE BOUNDS ON SYSTEM RELIABILITY*

Nancy R. Mann
Rocketdyne
Canoga Park, California

ABSTRACT. A description is given of results of preliminary investigations (by a group at North American Rockwell Corporation) related to the Monte Carlo generation of lower confidence bounds on the reliability of a logically complex system. In calculating system confidence bounds by use of a Monte Carlo procedure, one must generate the distribution of each independent subsystem reliability, given the life-test failure data for that subsystem. Therefore, an assumption of a specified a priori distribution for each subsystem reliability is implicit in the procedure.

In order that clues may be obtained as to optimum prior assumptions to be used in calculating Monte Carlo bounds for a complex system, the model has been restricted to a series system wherein each independent subsystem has exponentially distributed failure time and prototypes of each subsystem are tested until a fixed (but not necessarily the same for each subsystem) number of failures occurs. For this model, optimum (uniformly most accurate unbiased) exact classical confidence bounds on the reliability $R(t_m)$ at a specified mission time t_m are available, although not easily calculated (Lentner, M. M and Buehler, R. J., 1963. J. Amer. Statist. Assoc. 58, 670-677 and El Mawaziny, A. H., 1965. Unpublished doctoral dissertation, Iowa State University). Computer programs for calculating the optimum classical bounds and the Bayesian Monte Carlo bounds were written, and a means of numerically comparing various forms of prior distributions against an optimum standard was thus provided. *One prior distribution widely used in obtaining Monte Carlo and general Bayesian exact lower confidence bounds on system reliability is thereby shown numerically to yield bounds which are conservative in the classical sense for this series-system model. Another suggested prior distribution is shown to give bounds which are usually conservative but under certain conditions are liberal, and hence not truly confidence bounds. Moreover, it is demonstrated by a combination of numerical and analytical results, that for a series system containing more than one independent subsystem*

*This research was sponsored by the Mathematics and Statistics Panel of the Aerospace and Systems Group (A&SG) of North American Rockwell Corporation and funded under the Internal Research and Development program of the Executive Offices of A&SG.

Preceding page blank

there exists no prior distribution for subsystem reliability which is independent of the data and which yields the optimum lower bounds. Other numerical results related to the selection of optimum methods for generating the bounds and evaluation of certain approximate methods are described.

BACKGROUND AND APPROACH

Review of Pertinent Literature. If it is possible to determine confidence bounds on system reliability solely from the testing of the subsystems of which the system is comprised, saving of expensive system testing can be effected. It may, in fact, sometimes be infeasible to test the system as a whole. Furthermore, this method of obtaining system confidence bounds can be used for exploratory system design.

The subject of confidence bounds for system reliability from subsystem testing is one about which much has been written, but not a great deal is known. Consider a series system in which the failure times of k independent subsystems are exponentially distributed; i.e., for T a random variable representing failure time, Prob (T>t) = R(t) = $\exp(-\lambda t)$, $t \geq 0$, $\lambda > 0$. Suppose n_j prototypes of the jth subsystem, are subjected to life test and the life test is terminated at the time of the r_jth ordered failure, j=1,2,...,k. For this special model, there exist optimum (uniformly most accurate unbiased)[1] exact[2] confidence bounds on the reliability $R(t_m)$ at time t_m, the probability that the system will survive at least until time t_m. [See Lentner and Buehler (27) and El Mawaziny (12)]. No such optimum bounds have been found for a model which is equivalent to this exponential-failure-time series-system model, except for the fact that total test time t_j rather than number of failures r_j is specified for the life test of jth subsystem, j=1,2,...,k, and number of failures is the observable random variable. For either the fixed-time or fixed-number of failures model, optimum exact confidence

---

[1]The definitions of uniformly most accurate and unbiased confidence bounds are as given by E. Lehman (26). They are as follows: A confidence bound $\underline{\theta}(X)$ satisfying $P_\theta\{\underline{\theta}(X) \leq \theta\} \geq 1-\alpha$ for all θ and for all $\theta' < \theta$, $P_\theta\{\underline{\theta}(X) \leq \theta'\} = \text{minimum}$ is a uniformly most accurate lower confidence bound for θ at level $1-\alpha$. A family of lower confidence bounds at level $1-\alpha$ is said to be unbiased if $P_\theta\{\underline{\theta}(X) \geq \theta'\} \leq 1-\alpha$ for all $\theta' < \theta$ for all θ.

[2]A lower confidence bound at level $1-\alpha$ is said to be exact if $P_\theta\{\underline{\theta}(X) \geq \theta\} = 1-\alpha$ for all θ.

bounds have not been derived for cases in which either failure time has other than an exponential distribution (or can be converted by a transformation of the data to an exponential distribution) or the system is other than an independent series system.

Another much used failure model, often called the "attribute" model, is one in which only pass-fail binomially distributed data are collected for each independent subsystem. For this model, optimum exact confidence bounds on reliability (or probability of successful operation) of a series system have been derived [see Buehler (6)], but the problem of actually constructing such optimum bounds has not been completely solved [see Lipow (31), Lipow (32), Lloyd and Lipow (33), Steck (48), and Schick (43)]. If a Poisson approximation to the binomial distribution is applicable, then results of Harris (22) provide optimal exact bounds on the reliability of an independent series system for the attribute, model if one randomizes appropriately in obtaining the bounds. One would expect the Poisson approximation to the binomial distribution to apply when the number of prototypes of each subsystem tested is large and the probability of failure for each subsystem is small. There appears to be some question, however, [see Garner (19)] as to whether the approximation loses its applicability as the number of subsystems increases.

Many approximate and non-optimal exact confidence bounds on system reliability have been derived. There have been several approximate confidence bounds on system reliability at time t_m derived for the exponential fixed-number-of-failures model wherein the independent subsystems for a series system. Some of the papers containing these derivations were written prior to the publication of the derivation of the optimum bounds [see Takenaga (49) and Kraemer(25)].

Other work has been directed at providing a more tractable method of calculating confidence bounds than that of El Mawaziny's generalization to k subsystems, $k>2$, of the Lentner-Buehler bounds which apply to 2 subsystems only [see El Mawaziny and Buehler (12), Sarkar (41) and Grubbs (21)]. The method suggested by El Mawaziny and Buehler depends upon large-sample theory and the others use the fact that a function of the estimator of subsystem mean-time-to-failure has a chi-square distribution. The method of Sarkar does not require that the subsystems be independent and is exact for equal numbers of failures for all subsystems.

Some rather limited numerical comparisons have been made of some of these non-optimal methods for obtaining confidence bounds by, for example, Sarkar (41) and Grubbs (21). Apparently none of these methods have, until this time, been subjected to a thorough comparison with the Lentner-Buehler-El Mawaziny bounds, which must be calculated iteratively from an expression which demands extremely complicated calculations when the number of subsystems is more than two or three. (Problems involving loss of precision and use of excessive amounts of computer time also arise in calculating the El Mawaziny bounds when the product of the number of independent subsystems and the number of failures for any given subsystem is more than about 50.)

Other work dealing with the derivation of confidence bounds for system reliability under the exponential-failure-time model include a Bayesian approach for a parallel system with a single failure for each subsystem by Springer and Thompson (47) and two reports by Allen, Carlson and Hubach (2) and Saunders (42), which discuss the fixed-test-time model for a series system.

For the case in which only pass-fail data are collected for each subsystem many methods involving large-or small-sample approximations or Bayesian techniques have been derived for obtaining confidence bounds on the probability of successful operation of an independent series system. Among the large-sample methods are those suggested by Madansky (34) (based on the asymptotic chi-square distribution of -2 log likelihood ratio), by Myhre and Saunders (37) (which gives a generalization of Madansky's method) and by Rosenblatt (40), DeCicco (11) and Thomas (50) (all three of which are based on the asymptotic normality of maximum-likelihood estimators). The methods of Rosenblatt and Madansky are discussed and compared by Myhre and Saunders (38), who demonstrate that the likelihood ratio method attains its asymptotic properties for smaller sample sizes than the method suggested by Rosenblatt and in practical situations appears to yield more accurate bounds. Madansky (34), however, points out that the Rosenblatt method has slightly higher asymptotic (Bahadur) efficiency. The methods of DeCicco and Thomas use Taylor-series approximations to the variance of the maximum likelihood estimator of the system reliability R and would be expected to have asymptotic properties like those of the Rosenblatt method.

Small-sample approximate confidence bounds on R for an independent series system and binomial data have been derived by Nishime (39), Garner and Vail (20), Connor and Wells (8), Abraham (1) and Lindstrom and Madden [see Lloyd and Lipow (33)]. The first three of these approaches use various methods of combining confidence bounds on subsystem reliability to obtain the desired bounds on system reliability. The others use binomial or Poisson approximations for certain statistics. Some of these methods are sensitive to inequality of sample sizes for subsystems. Lower confidence bounds obtained by most of these approximate methods have been compared by the use of three sets of data by Schick and Prior (44) with three different sets of "exact" bounds obtained using results of Lipow [see (31) and (32)], based on Buehler's theory (6) and Poisson approximations. The data apply to systems composed of two subsystems, and in each of the three cases the sample sizes are equal. Only the Lindstrom and Madden method compares favorably with what appear to be the best of the Lipow "exact" bounds. Since there is some question about the standard used to judge the quality of the approximate methods, however, and since only three sets of data, two subsystems and equal sample sizes have been used in the comparisons, it is very difficult to make useful general inferences concerning these results.

Another method investigated numerically by Schick and Prior (44) is the Bayesian approach wherein reliability for each subsystem is assumed to

have a prior distribution which is uniform over the unit interval. Confidence bounds which are exact in the Bayesian sense (under the assumed prior distribution) are derived, by Zimmer, Briepohl and Prairie (51) and by Springer and Thompson (45). The latter authors use a Mellin transform technique for obtaining in closed form the distribution of system reliability, given all the subsystem data. A Monte Carlo application of this Bayesian model is suggested by Mastran (35) for a system which is logically more complex than a series system. In the numerical comparisons given by Schick and Prior (44), there appears to be no particular agreement between the sets of Bayesian bounds calculated on the basis of the procedure prescribed by Springer and Thompson and by Zimmer, et al. (Which incidentally agree to two or three significant figures, as one might expect) and the three sets of "exact" bounds calculated. In particular, the Bayesian lower confidence intervals on R are all larger than those based on what is for these three sets of data the smallest of the "exact" intervals.

One would expect the Bayesian bounds to be exact in a classical sense if sample sizes for all subsystems were "sufficiently" large. This is so because a prior density of the assumed type will have less effect upon the confidence bound as the sample sizes for all subsystems increase. Whether or not the bound is exact in a classical sense has not been established. Furthermore, the accuracy of this bound (see footnote 1) has not been investigated for small sample sizes. It is interesting to note that an approximate method, described by Dalton (10) and attributed to TRW's Florida Operations, yields bounds which agree to within 3 in the third significant figure with the three examples calculated in (44) by means of this particular Bayesian approach. The TRW method has the distinction of being extremely amenable to hand calculation.

Among very recently derived approximate methods for obtaining confidence bounds on the probability of successful operation of a series system are (1) those derived by Woods and Borsting (51) (discussed by Lieberman (30)), which are shown by Monte Carlo investigations in their paper to be very nearly exact; (2) those derived by J. R. Johnson (23) based on the exact multi-variate binomial distribution of component test data, and (3) those arising from a Bayesian approach which formally uses subjective judgment concerning prior knowledge by J. Bram (5).

The Monte Carlo Confidence Bound Problem. We now examine the problem of obtaining lower confidence bounds on the reliability of a logically complex system when testing will be performed on the k independent subsystems only. We assume that an equation relating true subsystem reliabilities to true system reliability is available, say by means of computer programs which can provide such information [see Levy (28) and McKnight, Modiest and Schmidt (36)]. We now, in lieu of an appropriate analytical method of obtaining such bounds, consider the possibility of the use of Monte Carlo techniques as suggested by Burnett and Wales (7), Bosnikoff and Klion (4), Costello, Meisel and Letow (9), Levy and Moore (29) and Mastran (35).

At first glance the creation of a Monte Carlo computer program for obtaining the bounds seems to be a straightforward problem of simulating the distribution of system reliability for a given set of failure data in an efficient manner. It soon becomes apparent, however, that there are important Bayesian questions implicit in the problem. That is, in order to generate the distribution of system reliability for a given data set, one must generate for each subsystem what is essentially the posterior distribution of subsystem reliability, given the subsystem life-test failure data. Hence, some prior distribution or something equivalent to such a prior distribution for subsystem reliability must be implicitly or explicitly assumed. In other words, in carrying out the Monte Carlo approach outlined by the authors mentioned above, one uses the density of some appropriate function of the data and implicitly or otherwise combines this information with a prior density of subsystem reliability by means of Bayes' Theorem, $P(A_i|B) = P(B|A_i)P(A_i) / \sum_{\text{all } j} P(B|A_j)P(A_j)$, to obtain the posterior density function of subsystem reliability, given the data. In agreement with the classical analytical method derived in (48), the Monte Carlo procedures described in (4), (7), (9), (29), and (35), in some cases directly suggest and in others tacitly imply a prior distribution for subsystem reliability which is the appropriate prior leading to the classical optimum bounds when the system consists of one subsystem only. One may then inquire as to whether such an assumption is appropriate when the system consists of more than one subsystem.

Springer and Thompson (47) analytically derive their exact Bayesian confidence bounds on $R(t_m)$ for an exponential-failure-time model, wherein one failure is allowed for each independent subsystem of a parallel system, using an alternative a priori assumption. They assume a uniform prior distribution on subsystem reliability over the unit interval, which leads to the classical optimum bounds on successful system operation for the pass-fail model when the system consists of a single subsystem. Springer and Thompson reason that a flat prior for subsystem reliability is in keeping with the intent of Bayes' Theorem when no prior information is known. They point out that the prior density $p(R_j)$ for the jth subsystem reliability yielding the classical optimum bounds for a system containing a single subsystem and an exponential fixed-failures model, $p(R_j) = R_j^{-1}[\ln(1/R_j)]^{-1}$ or equivalently, $q(\lambda_j) = \lambda_j^{-1}$, where $R_j = R_j(t_m) = \exp(-\lambda_j t_m)$ and $0 \leq R_j \leq 1$, $j=1,2,\ldots,k$, is "improper" in that the area under the frequency curve cannot be made equal to unity. Mastran (35) suggests for pass-fail data that prior densities for subsystems which lead to a uniform prior density for system reliability might be appropriate. In other words all the suggested prior distributions are derived from the concept of optimality for one subsystem for some model, even though the model may have little relationship to the one of interest.

In the following, a description is given of results of a study (by members of a group at North American Rockwell Corporation) to determine optimum prior assumptions to be used in generating Monte Carlo confidence bounds on the reliability of a logically complex system. The investigation was conducted principally by K. W. Fertig of Rocketdyne Division and the present author. Mr. Fertig wrote all computer programs needed for the investigation, except for one routine linking the Monte Carlo program to the reliability equation for the complex system. He also provided [see (17)] the important analytical derivation of the necessary form for a special restricted model of an optimum prior density function independent of the data and proved that no such prior density exists. Jerome Spanier of the North American Rockwell Science Center provided consultation on problems related to the Monte Carlo computer program. Shirley Stoneberger of the Los Angeles Division wrote the subroutine which makes use of the reliability equation generated from engineering flow chart information by the SCOPE (28) or the ARMM (36) program for a logically complex system.

RESULTS OF INVESTIGATION

Computer Programs Written and Utilized. An optimum standard against which to judge suggested prior distributions provides a means of attacking the Monte Carlo problem. Therefore, the model was first restricted to a series system wherein the ith independent subsystem has exponentially distributed failure time T_j with Prob $[T_j > t_m] = R_j = R_j(t_m) = \exp(-\lambda_j t_m)$ and n_j prototypes of the jth subsystem, $j=1;2,\ldots,k$, are tested until r_j failures occur. If one can determine an appropriate prior distribution for this model, then it should also be possible to make useful inferences concerning the fixed-failure-time series-system model and to determine a method of using prior information for more complex systems.

A computer program was coded in Fortran H for the IBM S/360 system for calculating for this restricted model the optimum classical confidence bounds of Lentner, Buehler and El Mawaziny discussed in the introduction of this paper. The bounds are based on the conditional distribution of $W = Z_1$, given $Z_1 - Z_2 = u_2, \ldots, Z_1 - Z_k = u_k$, where $Z_j = \sum_{i=1}^{r_j} T_{i,j} + (n_j - r_j)T_{r_j,j}$, with $T_{i,j}$ an observable failure time of the ith prototype of the jth component, and where the subscript 1 is arbitrarily assigned. Then, when u_j is less than zero for $j=2,3\ldots,k$, the optimum classical $(1-\alpha)$-level lower confidence bound $R_B(\alpha)$ on $R(t_m) = \exp(-\phi t_m)$, (where $\phi = \sum_{j=1}^{k} \lambda_j$) is obtained by finding the solution $\phi_B(\alpha)$ of the following equation and then calculating $R_B = \exp[-\phi_B(\alpha)\, t_m]$, with $\phi_B > 0$,

$$H\,(w|\underline{u};\phi) = A^{-1}\,(\underline{u};\phi) \sum_{i_2} \sum_{i_3} \cdots \sum_{i_k} \prod_{j=2}^{k} \left[\binom{a_j}{i_j} (-u)^{a_j - i_j} \right] \tag{1}$$

$$\phi^{-a_1-\Sigma_j\, i_j-1}\, \Gamma_{\phi w}\,(a_1+\Sigma_j\, i_j+1) = 1-\alpha,$$

where

$$A(\underline{u};\phi) = \phi^{-1-\sum_1^k a_j} \sum_{i_2} \sum_{i_3} \cdots \sum_{i_k} (a_1 + \sum_{j=2}^{k} i_j) \prod_{j=2}^{k} \left[\binom{a_j}{i_j} (-\phi u_j)^{a_j - i_j} \right]$$

and where

$$\Gamma_{\phi w}(a_1 + \Sigma\, i_j+1) = \int_0^{\phi w} y^{a_1 + \sum_{j=1}^{k_i} i_j} e^{-\phi y}\, dy \quad .$$

A similar expression is used if any u_j, $j=2,3,\ldots,k$, is greater than zero, and the solution is obtained by joint application of Newton-Raphson iterative procedures, the method of false position and bisection techniques. Then a computer program for generating Monte Carlo confidence bounds was coded and combined with that for obtaining the Lentner-Buehler-El Mawaziny confidence bounds. A listing and flow chart of the combined computer program are available [see Fertig (16) and (18)].

The Monte Carlo program calculates the confidence bounds on the basis of a specified prior density for subsystem reliability which is a member of the "conjugate" family of prior densities. That is, the prior density yields a posterior density of subsystem reliability, given the subsystem data, of the same general form (belonging to the same family of density functions) as the prior density. The prior density function $p_j(R_j)$ used for the jth subsystem reliability was, therefore,

$$p(R_j) = \frac{(\beta_{oj}+1)^{r_{oj}+1}}{\Gamma(r_{oj}+1)} R_j^{\beta_{oj}} [\ln(1/R_j)]^{r_{oj}} \qquad (2)$$

$\beta_{oj}, r_{oj} > -1$, $j = 1,2,\ldots,k$, with β_{oj} and r_{oj} subjectively chosen. This yields a posterior density $f(R_j|\hat{\beta}_j;r_j)$ for R_j of the form,

$$f(R_j|\hat{\beta}_j;r_j) = \frac{(\hat{\beta}_j+\beta_{oj}+1)^{r_j+r_{oj}+1}}{\Gamma(r_j+r_{oj}+1)} R_j^{\hat{\beta}_j+\beta_{oj}} [\ln(1/R_j)]^{r_j+r_{oj}} \qquad (3)$$

where the random variable $Z_j/t_m = (\sum_{i=1}^{r_1} T_{i,j} + (n_j - r_j)T_{r_j,j})/t_m$ is equal to $z_j/t_m = \hat{\beta}_j$, $j=1,2,\ldots,k$, for the observed set of data.

If β_{oj} and r_{oj} each have the value -1, then the prior density for R_j corresponds to the "improper" prior which is used by (48), (14), (7), and (29) and which gives the optimal classical bounds for a system consisting of a single subsystem [see Epstein and Sobel (14)], that is, $p(R_j) = R_j^{-1} [\ln(1/R_j)]^{-1}$, $j=1,2,\ldots,k$. (It is true, therefore, that even though the prior density corresponding to $\beta_{oj} = r_{oj} = -1$ is "improper," the corresponding posterior density is proper.) If β_{oj} and r_{oj} are both equal to zero, $j=1,2,\ldots,k$, then each subsystem prior density function for subsystem reliability is uniform over the interval from 0 to 1, as suggested by Springer and Thompson (47) for their special case of a parallel-system model mentioned earlier.

For generating the posterior distribution of R_j using the expression (3), a given set of data and specified values for β_{oj} and r_{oj}, a random number ρ_j is generated for the value of the integral $\gamma_j = \gamma(R_j|\hat{\beta}_j;r_j)$ given by the expression (3) from $R_{\gamma,j}$ to 1. The integration is performed by an evaluation of the incomplete gamma function and the value of $R_{\gamma,j}$ determined iteratively. The Newton-Raphson method of iteration in conjunction with the method of false position is used. Because this procedure is quite expensive in terms of computer time, the computer

program was written to calculate for a given set of subsystem data a table of 100 values of $R_{\gamma,j}$ corresponding to equally spaced values of $\gamma(R_j|\beta_j;r_j)$. The computer then samples from and interpolates cubically in this table for $a_1<\beta_j<a_2$ where a_1 and a_2 are functions of the data. For $\beta_j<a_1$ and $\beta_j>a_2$, a different table is sampled. In generating values for this table, $\gamma(R_j|\hat{\beta}_j;r_j)$ is calculated from a specified value of $R_{\gamma,j}$, so that no iteration is necessary, but the values of γ_j are not equally spaced (making interpolation more difficult). The second table, which contains values of γ_j much closer together than the one used for non-extreme values of $R_{\gamma,j}$, is necessary because of the steepness of the curve relating γ_j and $R_{\gamma,j}$ for values of $R_{\gamma,j}$ close to 0 or 1.

The first investigation made by means of the computer was of the two familiar prior distributions corresponding to β_{oj} and r_{oj} both equal to zero and both equal to -1, j=1,2,...,k. The Bayesian approach corresponding to $\beta_{oj} = r_{oj} = -1$, incidentally, is sometimes called the fiducial model since the posterior distribution of R_j, j=1,2,...,k, can be thought of as obtainable from the distribution of a function of the data for the jth subsystem, as detailed in (25). The preliminary phases of this investigation made use of the Monte Carlo program, but the results given below were obtained using instead a computer program which utilizes a Mellin transform technique [see Springer and Thompson (46)] to calculate the posterior distribution of $R(t_m)$ from the posterior distributions of the R_j's. This Mellin transform program was originally written to calculate the variance of the Monte Carlo confidence bound and is applicable to a series system when the posterior density of R_j has the form given by the expression (3) with r_{oj} an integer. The Mellin transform computer program is faster than the Monte Carlo program and gives better precision, but in its present form cannot be used if r_{oj} is other than an integer.

Study of Suggested Prior Densities. For each combination of input, involving from three to twenty-five components having λ's in various proportions, numbers of failures ranging from 1 to 10 and three or four different values of α ranging from .05 to .50, data were generated and a comparison was made of the two Bayesian bounds with the optimum classical bound. In each case (of a total of 156 cases), the Bayesian bound based on $\beta_{oj} = r_{oj} = -1$ is smaller than the corresponding classical bound obtained. It, therefore, appears that though exact in a Bayesian sense (under the assumed prior distribution for R_j,j=1,2,...k), the bounds based on such a prior assumption are conservative in the classical sense.

When the optimum bound is standardized at .800 by adjusting the mission time t_m, the fiducial bound ranges from .538 to .793. When the optimum bound is equal to .368, the fiducial bound ranges from .062 to .354.

El Mawaziny and Buehler (13) show that their large-sample approximation of the optimal bound, a bound obtained by the Rosenblatt method (40) and the fiducial bound will approach the optimal bound as numbers of failures for all subsystems become large. For three samples having ten failures for each of three identical components, the fiducial bounds were of the order .787 and .341 for optimal bounds of .800 and .378, respectively, with deviation between any two corresponding fiducial bounds less than three in the third decimal place.

Analytical results described later indicate that the fiducial bounds for a fixed number of failures per subsystem will agree less well as the number of subsystems increases and the subsystems become more variable with respect to failure rate. Unfortunately, because of the computer-time factor and considerations of precision, it is impossible at present to compare bounds for systems containing as many as ten subsystems when as many as ten failures occur for more than one or two of these subsystems. In any case, the large-sample methods cannot be expected to give bounds agreeing well with the optimal bounds when some of the subsystems have been subjected to few tests. Furthermore, it is impossible on the basis of these results to say whether bounds based on this specified prior might be conservative, liberal, or exact for a particular logically complex system.

The uniform prior distribution for subsystem reliability gives bounds even lower than those based on the fiducial method except in 24 cases (out of 150) in which all three bounds have values fairly close to zero. In these 24 cases they are higher than the optimum classical bounds. It appears that the distribution of the bounds based on the uniform prior may be less disperse than those of the optimal bounds, but these bounds seem to be even more conservative than the "fiducial" bounds given by $\beta_{oj} = r_{oj} = -1$ for true reliabilities of a reasonable size and α's of interest. For systems with low reliabilities, bounds obtained using a uniform prior density for subsystem reliability should be liberal rather than conservative when the confidence level is sufficiently low, but not exact in general. This inconsistent behavior may be due to the fact that a uniform prior density for R_j implies a prior density for λ_j (the failure rate for the jth subsystem) of the form $q(\lambda_j) = t_m \exp(-\lambda_j t_m)$, $j=1,2,\ldots,k$, or, strangely, one which is a function of t_m, the specified mission time.

The result for $\beta_{oj} = r_{oj} = -1$ is keeping with the analysis and numerical results of Saunders (42), who studies a fixed-test-time exponential series-system model and the Bayesian approach suggested in (2). This Bayesian approach uses a prior density for the fixed-test-time

model equivalent to the so-called fiducial method. Saunders (42) points out that in using such a Bayesian model for an exponential series system (and his argument applies to any true Bayesian model, that is, one based on a prior assumption which does not involve information concerning the number of components in the system), one can obtain different confidence bounds depending upon what one chooses to call a subsystem. Saunders points out, too, that such inferences apply to more logically complex systems which are highly reliable, since one can approximate an extremely reliable coherent system [see Birnbaum, Esary and Saunders (3) for a definition of a coherent system] by a series-system model, as indicated by Esary, Proschan and Walkup (15).

The Search for Optimum Prior Assumptions. Initially, it has been planned that a trial and error procedure would be used in attempting to determine appropriate prior assumptions for our series-system model. Saunders' argument might lead one to consider trying prior assumptions which are not truly Bayesian in that they are dependent upon the configuration (or number of components in the system). At this point in the study, however, an analytical result was derived, modifying the subsequent approach. The details of the analysis are given by Fertig (17) and are summarized below.

First, the form of a prior density function, or generalization of such a function, for R_j, $j=1,2,\ldots,k$, corresponding to the optimum classical bounds for our system model was determined. This was accomplished by setting the Laplace transform of $H_{\phi}(w|\underline{u};\phi)$ equal to the Laplace transform of the posterior density of $\phi = \sum_{j=1}^{k} \lambda_j$, obtained under a general prior assumption (not restricted to conjugate priors) for the special case $u_2 = u_3 = \ldots = u_k = 0$. If the prior assumptions which yield the optimum bounds are independent of the data, true Bayesian priors, for example, then an assumption concerning the value of the u's will have no effect on the result. The fact that the optimum classical bounds are invariant under permutations of the subsystems was used to obtain the improper "prior density" for the jth subsystem yielding these bounds. It is

$$p(R_j) = R_j^{-1} [\ln(1/R_j)]^{-(2-1/k)} \quad , \; j=1,2,\ldots,k.$$

We note that this improper density depends upon k, the number of subsystems in the series system and for k=1 does yield the optimum classical bound.

The Monte Carlo program was then used to test whether this prior assumption (which we may think of as a weighting function since it does not correspond to a strict Bayesian prior density) would yield the optimum bounds for variations in the data. For the case where all the u's equal zero ($z_1 = z_2 = \ldots = z_k$), eight values of the Monte Carlo bounds based

on 5,000 replications agree with the optimum bounds to within three in the third decimal place. Data were randomly generated for five subsystems using the fact that $2Z_j/\lambda_j$ is distributed as chi-square with $2r_1$ degrees of freedom [see Epstein and Sobel (14)], where $r_1 = 3$, $r_2 = 4$, $r_3 = 4$, $r_4 = 2$, $r_5 = 2$, $\lambda_1 = 1/12$, $\lambda_2 = 1/13$, $\lambda_3 = 1/15$, $\lambda_4 = 1/10$, $\lambda_4 = 1/11$. The mission time was taken as 1.0. For ten such data sets, the Monte Carlo bounds are uniformly larger than the classical confidence bounds with deviations ranging from 1 to 6 in the second decimal place.

The confidence bound obtained from any set of data by El Mawaziny's formula given by Eq. (1) is the unique optimum (uniformly most accurate unbiased) confidence bound for this exponential-failure-number series-system model. This is proved in the Appendix of Fertig's paper. Thus, any optimum bound is equivalent to the bound defined by Eq. (1) and must give the same result for any given set of data. Since the error in the Monte Carlo procedure is very small compared with the deviations obtained for the u's not all equal to zero, the empirical evidence indicates that the prior assumptions which yield the optimum classical bounds do depend upon the data.

Fortunately, a means of proving this result analytically then presented itself [see Fertig (17) for details]. The Laplace transform of $H(w|\underline{u};\phi)$ is a horrendous expression which gives no apparent clue as to how it might be factored and assigned to the various subsystems. The problem was made tractable earlier by letting all the z's be equal.

Another method of simplifying the expression was found to be to assume $r_1 = r_2 = \ldots = r_k = 1$. If this is done, then it is possible to demonstrate that the "prior density" for the jth subsystem yielding the optimum bound for $r_1 = r_1 = \ldots = r_k = 1$ cannot have the form $p(R_j) = R_j^{-1}\,[\ln(1/R)]^{-(2-1/k)}$ unless $z_1 = z_2 = \ldots = z_k$, as assumed in the earlier case. Hence, as indicated by the numerical evidence, one must incorporate present data into the "prior assumptions" or more properly the weighting functions, for obtaining optimum confidence bounds for the exponential fixed-failure-number series- system model.

Now that we have established what is not fruitful for obtaining the optimal bounds, one may properly inquire as to the next step in the investigation with respect to confidence for a complex system. Two approaches present themselves. The first is to consider each series system which is a part of the complex system as a single subsystem of the total system. If the fiducial approach were to be used, one could obtain an estimate $R_{\rho,\ell}$ of true reliability, given the failure data, for the ℓth subsystem consisting of the k_ℓ independent subsystems making up

this series-system subsystem by substituting a random number p_ℓ for [illegible] in Eq. (1). Then the reliability equation could be used to obtain an estimate of total system reliability given the failure data. This method will probably not yield confidence bounds that are exact in the classical sense, though of course they are exact in a Bayesian sense. In lieu of the fiducial approach one could somehow modify the value obtained for $R_{0,\ell}$ in an attempt to obtain bounds exact in the classical sense. Clues as to how this might be accomplished may possibly be obtained by investigating a simple parallel system (again an exponential-failure-time model with fixed number of failures) and methods of obtaining confidence bounds exact in the classical sense for the simpler model.

A considerable amount of computer time will be required with this approach when the product of the number of subsystems in any series system and the number of failures for any subsystem in that series system becomes large. Hence, one might in such cases, abandon the idea of considering the series system as a single subsystem. Instead one might consider a method of approximating the optimum confidence bound for a series system by using series system data in the prior assumptions (or weighting functions) for the subsystems of the series system that yield the posterior distribution of series system reliability. For the case $r_1 = r_2 = \ldots r_k = 1$, one possible factoring of the Laplace transform gives

$$p_\ell(R_j) = R_j^{-1-\hat{\beta}_j+\hat{\beta}_{(1)}} \quad [\ln(1/R_j)]^{-(2-1/k_\ell)} \quad ,j=1,2,\ldots,k_\ell, \quad (4)$$

where $\hat{\beta}_{(1)}$ is the smallest of z_1/t_m, z_2/t_m,...,$3_{k_\ell}/t_m$. This function or some modification of this function involving the true values for r_1, r_2, ..., r_k could be tried. If the expression (4) is used without modification, then the posterior distribution of series system reliability, given the failure data, depends only upon $\hat{\beta}_{(1)}$. It can be shown that, in fact, the fiducial distribution of series system reliability for this model depends only upon $\hat{\beta}_{(1)}$ if and only if $r_1 = r_2 = \ldots = r_{k_\ell} = 1$.

The method of Kraemer (24) depends solely upon $\hat{\beta}_{(1)}$ and hence would be expected to give poor results for large numbers of failures per component. This is born out by the comparisons made by Sarkar (42) and Grubbs (21).

The bounds derived by Grubbs (21) are approximations to the fiducial bounds, and for the bounds compared during this study, the approximation appears to be excellent. The Grubbs method is based on the fact that $2r_j \lambda_j/\hat{\lambda}_j$ is distributed as chi-square with $2r_j$ degrees of freedom,

$j=1,2,\ldots,k$, so that $\phi = \sum_{i=1}^{k} \lambda_j$ can be thought of as distributed as a weighted sum of chi-square variates. The weights used by Grubbs, namely, $\hat{\lambda}_j/2r_j$, $j=1,2,\ldots,k$, are appropriate for obtaining the fiducial bounds. One could obtain, instead, an approximation to the optimal bounds by adjusting the weights appropriately, obtaining clues from the expression (4) above. In this way, one can test approximations to the optimum prior assumptions and avoid the time-consuming Monte Carlo calculations. If successful in closely approximating the optimum bounds by the proper modification of the Grubbs method, one can use this approximation in place of the Lentner-Buehler-El Mawaziny bounds in considering series systems within a complex system as single subsystems.

The investigation is being continued along these lines.

REFERENCES

1. Abraham, John K., 1962. A confidence interval for the reliability of multi-component systems. PROCEEDINGS SEVENTH CONFERENCE ON THE DESIGN OF EXPERIMENTS IN ARMY RESEARCH, DEVELOPMENT AND TESTING, ARODR 62-2, 483-518.

2. Allen, D. C., Carlson, C. H. and Hubach, C. E., 1967. Procedure for Reliability Assessment and Confidence. Federal Electric Corporation Document No. AS-A-86-67.

3. Birnbaum, Z. W., Esary, J. D. and Saunders, S. C., 1961. Multi-component systems and structures and their reliability, TECHNOMETRICS 3, 55-71.

4. Bosniköff, I. and Klion, J., 1962. Development of new prediction techniques. PROCEEDINGS EIGHTH NATIONAL SYMPOSIUM ON RELIABILITY AND QUALITY CONTROL, 382-387.

5. Bram, J., 1969. "Confidence Limits for System Reliability." OEG Research Contribution No. 79, Center for Naval Analyses, University of Rochester.

6. Buehler, R. J., 1957. Confidence intervals for the produce of two binomial parameters. J. AMER. STATIST. ASSOC. 52, 482-493.

7. Burnett, Thomas L. and Wales, Beverly A., 1961. System reliability confidence Limits. PROCEEDINGS SEVENTH NATIONAL SYMPOSIUM ON RELIABILITY AND QUALITY CONTROL, 118-128.

8. Conner, W. S. and Wells, W. T., 1962. Simulating tests of a system from tests of its components. PROCEEDINGS OF THE EIGHTH NATIONAL SYMPOSIUM ON RELIABILITY AND QUALITY CONTROL, 14-16.

9. Costello, D. L. Meisel, R. M. and Letow, A. M., 1962. Compliance demonstration of a multimode system using Monte Carlo analysis. PROCEEDINGS EIGHTH NATIONAL SYMPOSIUM ON RELIABILITY AND QUALITY CONTROL, 446-457.

10. Dalton, R. E., 1966. "An Evaluation of Methods for Construction of Confidence Limits for System Reliability." TRW Systems, Florida Operations, Contract AF04 (694)-806.

11. DeCicco, Henry, 1960. "The Error in Linearized Estimates of the Variance of Products." Technical Note No. 2, Office for Reliability Research and Effects (Reliability Branch - ORDSW-DR).

12. El Mawaziny, A. H., 1965. "Chi-Square Distribution Theory with Applications to Reliability Problems." Ph.D. Thesis, Iowa State University, Ames.

13. El Mawaziny, A. H. and Buehler, R. J., 1967. Confidence Limits for the reliability of a series system. J. AMER. STATIST. ASSOC. 62 1452-59.

14. Epstein, B. and Sobel, M., 1963. Life Testing. J. AMER. STATIST. ASSOC. 48, 486-502.

15. Esary, J. D., Proschran, Frank, and Walkup, D. W., 1966. "A Multivariate Notion of Association, with a Reliability Application," Boeing Document D1-82-0567.

16. Fertig, K. W., 1969. "Flow Chart for Monte Carlo Program for Confidence Bounds on System Reliability." Rocketdyne Report MSM 69-11.

17. Fertig, K. W., 1969. "A Result Concerning Bayesian Prior Distributions and Confidence Bounds on the Reliability of Serial Systems with Exponential Failure Times." Rocketdyne Research Report RR 69-6. (Submitted for publication.)

18. Fertig, K. W., 1969. "Monte Carlo Program for Confidence Bounds on System Reliability." Rocketdyne Report MSM 69-12.

19. Garner Norman R., 1969. "Estimation for Serially Connected Systems." Technical Report RCS59-4 Aerojet-General Corporation.

20. Garner, Norman R. and Vail, Richard W. J., 1961. Confidence Limits for system reliability MILITARY SYSTEMS DESIGN 7, No. 5.

21. Grubbs, Frank, 1969. "On Confidence Limits for the Reliability of a Series System for Which each Component has an Exponential Time-to-Fail Distribution." (Submitted for publication.)

22. Harris, Bernard, 1968. "Hypothesis Testing and Confidence Intervals for Products and Quotients of Poisson Parameters with Applications to Reliability." MRC Technical Summary Report No. 923. U. S. Army Mathematics Research Center, The University of Wisconsin, Madison.

23. Johnson, J. R., 1969. "Confidence Interval Estimation of the Reliability of Multicomponent Systems Using Component Test Data." Ph.D. Thesis, University of Delaware.

24. Kraemer, H. C., 1963. One-sided confidence intervals for the quality indices of a complex item. TECHNOMETRICS 5, 400-403.

25. Kendall, M. G. and Stuart, A., 1967. THE ADVANCED THEORY OF STATISTICS VOL. 2, Second Edition, Hafner.

26. Lehman, E. L., 1959. TESTING STATISTICAL HYPOTHESES. John Wiley.

27. Lentner, M. M. and Buehler, R. J., 1963. Some inferences about gamma parameters with an application to a reliability problem. J. AMER. STATIST. ASSOC. 58, 670-677.

28. Levy, Sherwin, 1969, "System to Compute Operational Probability Equation." Program XC0003. Space Division, North American Rockwell Corporation.

29. Levy, Louis L. and Moore, Albert H., 1967. A Monte Carlo technique for obtaining system reliability confidence limits from component test data. I.E.E.E. TRANSACTIONS ON RELIABILITY R-16 No. 2, 69-72.

30. Lieberman, Gerald J., 1969. The status and impact of reliability methodology. NAV. RES. LOG. QUART. 16, 17-35.

31. Lipow, M., 1958. "Measurement of Over-All Reliability Utilizing Results of Independent Subsystem Tests." GM-TR-0165-00506, Space Technology Laboratories.

32. Lipow, M., 1959. "Tables of Upper Confidence Bounds on Failure Probability of 1, 2, and 3 Component Serial Systems." TR-50-0000-00756, Space Technology Laboratories, 2 Volumes.

33. Lloyd, D. K. and Lipow, M., 1962. Reliability: Management, Methods, and Mathematics, Prentice-Hall.

34. Madansky, Albert, 1965. Approximate confidence limits for the reliability of series and parallel systems. TECHNOMETRICS 7, 495-503.

35. Mastran, David V., 1968. A Bayesian approach for assessing the reliability of Air Force re-entry systems, PROCEEDINGS OF THE ASME RELIABILITY AND MAINTAINABILITY SYMPOSIUM, 380-383.

36. McKnight, C. W., Modiest, L. J. and Schmidt, N. E., 1965. An Automatic reliability mathematical model. PROCEEDINGS ELEVENTH NATIONAL SYMPOSIUM ON RELIABILITY AND QUALITY CONTROL, 518-532.

37. Myrhe, Janet and Saunders, Sam C., 1965. On confidence limits for the reliability of systems. ANN. MATH. STATIST. 39, 1463-1472.

38. Myhre, J. M. and Saunders, Sam C., 1968. Comparison of two methods of obtaining approximate confidence intervals for system reliability. TECHNOMETRICS 10, 37-49.

39. Nishime, F., 1959. Techniques for ... the Establishment of Confidence Limits for the Estimated Reliability, Unpublished memorandum, Space Technology Laboratories.

40. Rosenblatt, Joan R., 1963. Confidence limits for the reliability of complex systems. STATISTICAL THEORY OF RELIABILITY, 115-137. Ed. Marvin Zelen, The University of Wisconsin Press, Madison.

41. Sarkar, Tapas K., 1969. "An Exact Lower Confidence Bound for the Reliability of a Series System Where Each Component Has an Exponential Time to Failure Distribution." Technical Report No. 117, Department of Operations Research and Department of Statistics, Stanford University.

42. Saunders, Sam C., 1969. "On Confidence Limits for the Performance of a System When Few Failures are Encountered." Boeing Document D1-82-0676, Revised.

43. Schick, G. J., 1959. "Reliabilities, Confidence Limits and Their Improvements as Applied to Missile Reliability." Technical Publication, Aerojet General Corporation.

44. Schick, G. J. and Prior, R. J., 1966. Reliability and confidence of serially connected systems. PROCEEDINGS OF THE THIRD SPACE CONGRESS, Cocoa Beach, Florida, 352-360.

45. Springer, M. D. and Thompson, W. E., 1966. Bayesian confidence limits for the product of n binomial parameters. BIOMETRIKA 53, 611-613.

46. Springer, M. D. and Thompson, W. E., 1969. Bayesian confidence limits for system reliability. PROCEEDINGS OF THE ASME RELIABILITY AND MAINTAINABILITY SYMPOSIUM, 515-523.

47. Springer, M. D. and Thompson, W. E., 1968. Bayesian confidence limits for reliability of redundant systems when tests are terminated at first failure. TECHNOMETRICS 10, 29-36.

48. Steck, G. P., 1957. "Upper Confidence Limits of the Failure Probability of Complex Networks." SC-433 (TR), Sandia Corporation.

49. Takenaga, R., 1962. "Predicting System Reliability with Associated Confidence Level from Component Test Data." Technical Memorandum 3024-43-MA-003, Autonetics, A Division of North American Aviation, Inc., Downey, California.

50. Thomas, Ralph E., 1960. "An Improved Formula for the Standard Deviation of the Reliability Product Rule." Technical Report No. 2, Battelle Memorial Institute.

51. Woods, W. Max and Borsting, J. R., 1968. "A Method for Computing Lower Confidence Limits on System Reliability Using Component Failure Data with Unequal Sample Sizes." NPS 55Wo/Gg 8061A. United States Naval Postgraduate School, Monterey, California.

52. Zimmer, W. J., Prairie, R. R., and Breipohl, A. M., 1965. A consideration of the Bayesian approach in reliability evaluation. I.E.E.E. TRANSACTIONS ON RELIABILITY R-14, No. 2, 107-113.

COMPARATIVE ANALYSIS OF THE LCSS-ETG-3 PERFORMANCE CAPABILITY USING STATISTICAL PROBABILITIES

Andrew H. Jenkins
U. S. Army Missile Command
Redstone Arsenal, Alabama

ABSTRACT

A test plan is formulated and executed to obtain a random sample of measurements on the U.S. Army Missile Command's Land Combat Support System Electronic Test Group equipment.

Analyses of the data establish sample estimates of bias, accuracy, and stimulus setting errors and the standard deviation of measurement on 14 combinations of parameters and scales (e.g., dc voltage, 10-volt scale). The analyses pose hypotheses about the statistics and test these hypotheses against appropriate frequency distributions. They include the principle of analysis of variance, which makes use of bias error, accuracy error, stimulus setting error, and sample variance. These four parameters are used as response variables to establish the effects of the main factors of test durations, time delays, and machines and combinations of the main factors (i.e., interactions) on the computed response statistics for each of the 14 parameters and scales considered.

The overall estimates of the precision (standard error of measurement) for each parameter and scale are related to actual weapon system tolerances to obtain probability estimates of the risk of passing a bad unit ("undetected defect") or holding a good unit ("false alarm") in a single test in a checkout procedure. Single checkout probabilities are related to multiple sequential checkout probabilities.

ACKNOWLEDGEMENTS

The author is grateful to the following persons and their organizations for their assistance: The LCSS Project Office of the U.S. Army Missile Command for running the test tapes; F. Seeley and W. Barron of the Metrology Laboratory, U.S. Army Missile Command, for compiling the test tapes; W. Wigham, W. Jones, and C. Wood of the Computation Center, U.S. Army Missile Command, for processing the data; and Dr. B. Harshbarger of Virginia Polytechnic Institute for guidance on the statistical aspects.

This article has been reproduced photographically from the author's manuscript.

Preceding page blank

Section I. INTRODUCTION

The purpose of the Land Combat Support System (LCSS) electronic test group (ETG) equipment is to provide maintenance support and check operational readiness of major modules, assemblies, and subassemblies of the Shillelagh, Lance, and TOW missile systems. The primary requirement is for direct and general support missions. A detailed description of the LCSS-ETG can be found in a previous report [1]. The ETG is designed to automate the testing of critical components of the missile systems to achieve:

a) Rapid evaluation of the operational status of the unit under test (UUT).

b) Rapid fault isolation of a defective UUT.

c) Automated decision making as to operational status by comparison of measured values with prescribed standards.

d) A standardized automated test capability for several weapon systems.

Automation of the ETG equipment requires the preparation of a programmed test sequence. The test program instructs the operator on the required manual operations for the checkout such as external connections to "make" and "break." The program includes all necessary tests for functional checkout of the UUT's as prescribed by the weapon system design engineers. Typical tests are to measure stimuli and responses of such parameters as ac and dc voltage, resistance, optical alignment, frequency, phase, and time, and to compare these measurements with prescribed values. The required values of such parameters along with acceptable tolerances (deviations) are prescribed in the test program. The test equipment makes the measurement and compares it with the specified value and decides on a "go/no-go" basis as to a fault determination.

In the testing of missile components on a go/no-go basis there are a combination of conditions which may exist. A unit may be good and check good resulting in a go decision. The unit may be good and check bad resulting in a no-go decision. On the other hand, a unit may be bad and check good resulting in a go decision, or it may be bad and check bad resulting in a no-go decision. There are certain probabilities associated with these combinations of actual component condition and checkout results. These are shown in Table I. The $p(\alpha)$ is the probability that a unit checks good when in fact it is bad. The $p(\beta)$ is the probability that a unit checks bad when in fact it is good. Some authors refer to these probabilities as an "undetected defect" and a "false alarm,"

TABLE I. CHECKOUT PROBABILITY VERSUS UNIT CONDITIONS AND TEST DECISIONS

Unit Condition and Checkout	Decision	
	Go	No-Go
Bad		
Checks good (undetected defect)	$p(\alpha)$	NA
Checks bad	NA	$[1 - p(\beta)]$
Good		
Checks good	$[1 - p(\alpha)]$	NA
Checks bad (false alarm)	NA	$p(\beta)$

respectively. It can be seen that the $p(\alpha)$ is related to a go decision and the $p(\beta)$ is related to a no-go decision based on the test results. A unit that is good and checks good will not result in a no-go decision. Similarly, a unit that is bad and checks bad will not result in a go decision. These combinations are not applicable and are shown as NA in Table I. In a go decision situation there is a probability that a bad unit has checked good; i.e., there is a $p(\alpha)$ chance that a defect exists and it is undetected by the test equipment which is an "undetected defect." In a no-go decision situation there is a probability that a good unit has checked bad; i.e., there is a $p(\beta)$ chance that the test equipment has falsely indicated a defect that does not exist which is a "false alarm." Therefore, it is seen that the $p(\alpha)$ is the probability of simultaneously getting a measured value within the specification limits (or the decision limits) and an actual value outside the specification limits. The $p(\beta)$ is the probability of simultaneously getting a measured value outside the specification limits (or decision limits) and an actual value inside the specification limits.

In other words, given a go decision, the probability that it is wrong (bad checks good) is $p(\alpha)$ and the probability that it is right (good checks good) is $[1 - p(\alpha)]$. Given a no-go decision, the probability that it is wrong is $p(\beta)$ and the probability that it is right is $[1 - p(\beta)]$.

The $p(\alpha)$ and $p(\beta)$ set for the test equipment should be realistically determined in light of the weapon mission and test equipment environment. If the probability $p(\alpha)$ is set too high, an excessive number of bad units going to the troops will result.

On the other hand, if $p(\beta)$ is too high, an excessive amount of time is spent in checking for a defect that does not exist. Therefore experience, knowledge of military tactics and good judgment should govern the compromises between logistics, field troop effectiveness, troop operational conditions, military objectives, etc., to determine the levels of $P(\alpha)$ and $P(\beta)$. (Note: $p(\alpha)$ and $p(\beta)$ refer to single test probability and $P(\alpha)$ and $P(\beta)$ refer to multiple test probability.) The determination of $P(\alpha)$ and $P(\beta)$ considering the above military factors is not germane to this effort. This effort is concerned with the analysis of probabilistic relations between error probabilities $p(\alpha)$ and $p(\beta)$, and standard deviation error of measurement instrument $\left(\sigma_m\right)$, standard deviation of test parameter $\left(\sigma_p\right)$, parameter tolerance for a given $y\sigma$ confidence level (θ), and decision limits (γ). Also included are the relations between single test probabilities $p(\alpha)$, $p(\beta)$ and multiple test probabilities $P(\alpha)$, and $P(\beta)$.

Section II. MATHEMATICAL MODELS

1. General

In the measurement of any one individual parameter by the ETG, there are three things to be considered. The first is the value specified by the weapon system for the UUT. This is called the nominal value of the parameter (N). The second is the actual value of the parameter (X). The third is the measured value of the parameter (M). It is assumed that the actual value (X) of the parameter is related to the measured value (M), according to the normal probability density function. This assumption is based on the fact that there is no inherent bias error as would be caused by coupling, feedback loops, and switching in the instrument and that all errors in measurement are completely random and normally distributed. It is also assumed that the actual value X is distributed normally about the nominal value N, according to a normal probability density function.

Since the go/no-go decision is made on the measured value of the parameter, the normal probability density distribution for the random measurement error is considered in the following way. The density function is considered to be centered at the measured value M of the parameter. The density function with standard deviation σ_m describes the distribution of the possible actual value X_i's that could have resulted in a given measured value M.

2. Single Check Probability

The normal probability density distribution for the measurement error, for a given value M, has the form

$$f(X) = \frac{1}{\sigma_m \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X - M}{\sigma_m}\right)^2} \qquad (1)$$

The actual parameter value, X, is also a random variable with a probability density function f(X). It is reasonable to assume that the actual value X is normally distributed about the nominal parameter value N with a standard deviation for the nominal parameter value of σ_p. With no bias error in the measurement device, the measured value M will also be normally distributed

about the nominal value N with a density function f(M). If the measurement standard deviation σ_m is an order of magnitude less than the parameter standard deviation σ_p for the nominal value N then

$$f(M) \approx f(X) \ . \tag{2}$$

Therefore

$$f(M) = \frac{1}{\sigma_p \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{M-N}{\sigma_p}\right)^2} \tag{3}$$

The nominal value, N, as specified by the weapon system, also has prescribed tolerances. These tolerances are usually assumed to be $\pm n\sigma_p$ where σ_p is the standard deviation from the acceptable mean value of the parameter. The $+ n\sigma_p$ value is the upper specification limit and the $- n\sigma_p$ value is the lower specification limit, S_u and S_l, respectively. The tolerances may be specified as the $n\sigma_p$ level, where n = 1, 2, 3, 5, etc. Whatever the specified $n\sigma$ level, it represents the allowable limits for the parameter values by the weapon system for proper operation of the unit [2].

In order to assure that the probability of an undetected defect does not exceed a specified maximum, the measured value, M, must fall between even tighter test limits, defined as upper and lower decision limits D_u and D_l, respectively. A go/no-go decision is then based on whether the measured value falls inside or outside the decision limits and not the weapon system specification limits. The decision limits may be set at the specification limits or some fraction of the specification limits. That is:

$$\left(D_u, D_l\right) = a\left(S_u, S_l\right) \tag{4}$$

where $0 < a \leq 1$.

This is shown graphically in Figure 1 [3].

In Figure 1, it can be seen that $\theta = \pm n\sigma_p$ and $\gamma = \pm a\left(n\sigma_p\right)$ represents the upper and lower specification limits and the upper and lower decision (test)

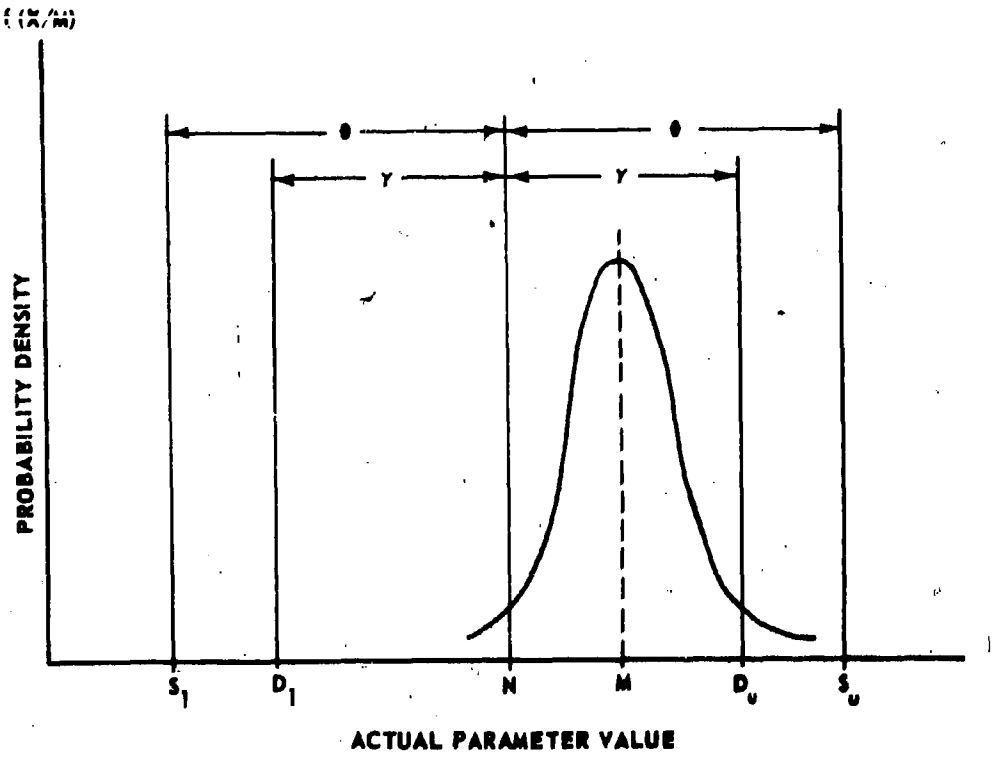


FIGURE 1. MEASUREMENT ERROR PROBABILITY DENSITY DISTRIBUTION

limits. The normal distribution curve around M is also shown to clarify the proposition of setting the decision limits less than the specification limits.

The probability of an undetected defect $p(\alpha)$ is the probability of simultaneously getting a measured value, M, within the decision limits and an actual value X outside the specification limits. The probability of getting a measured value, M, within the decision limits is $f(M)dM$. The probability that the measured value resulted from an actual value, X, outside the specification limits is

$$1 - \int f(X/M)dX \,. \qquad (5)$$

The simultaneous probability is the product of the two individual probabilities:

$$\left[f(M)dM\right]\left[1-\int_{S_l}^{S_u} f(X/M)dX\right]. \qquad (6)$$

The probability of an undetected defect is the summation of the above probability over all possible M's between the decision limits:

$$p(\alpha) = \int_{D_l}^{D_u} f(M)\left[1-\int_{S_l}^{S_u} f(X/M)dX\right]dM. \qquad (7)$$

The probability of a false alarm $p(\beta)$ is the probability of simultaneously getting a measured value, M, outside of the decision limits and an actual value, X, inside the specification limits. Similarly,

$$p(\beta) = \int_{-\infty}^{D_l} f(M)\left[\int_{S_l}^{S_u} f(X)dX\right]dM + \int_{D_u}^{+\infty} f(M)\left[\int_{S_l}^{S_u} f(X)dX\right]dM. \qquad (8)$$

In equations (7) and (8) the limits are expressed as follows:

$$D_l = N-\gamma$$

$$D_u = N+\gamma$$

$$S_l = N-\theta$$

$$S_u = N+\theta \qquad (9)$$

and f(X) and f(M) are as defined in equations (1) and (3) above.

Substituting f(X) and f(M) into equations (7) and (8) gives an integral equation which is, according to Duncan [4], in the noncumulative form and cannot be integrated in closed form. Numerical approximations have been obtained and set up in tabular form. However, in order to obtain reasonably close engineering estimates of $p(\alpha)$ and $p(\beta)$, an exponential of the form e^{X} and e^{-X}

is used and the integration performed with the limits of equation (9) substituted. In this manner, the $p(\alpha)$ and $p(\beta)$ equations are obtained in terms of σ_p, σ_m, θ, and γ (the desired parameters) and reduce to the following equations:

$$p(\alpha) = \left[\frac{\sigma_m}{2\,\sigma_p - \sigma_m}\right] e^{-1.15\left(\frac{\theta\sigma_p - \gamma\sigma_p + \gamma\sigma_m}{\sigma_p\sigma_m}\right)} - \left[\frac{\sigma_m}{2\left(\sigma_p + \sigma_m\right)}\right] e^{-1.15\left(\frac{\theta\sigma_p + \gamma\sigma_p + \gamma\sigma_m}{\sigma_p\sigma_m}\right)} - \left[\frac{\sigma_m^2}{\sigma_p^2 - \sigma_m^2}\right] e^{-1.15\left(\frac{\theta}{\sigma_m}\right)} \qquad (10)$$

$$p(\beta) = e^{-1.15\left(\frac{\gamma}{\sigma_p}\right)} - \left[\frac{\sigma_p^2 - 2\sigma_p\sigma_m - \sigma_m^2}{\sigma_p^2 - \sigma_m^2}\right] e^{-1.15\left(\frac{\theta}{\sigma_p}\right)} + \left[\frac{\sigma_m}{2\left(\sigma_p - \sigma_m\right)}\right] e^{-1.15\left(\frac{\theta\sigma_p - \gamma\sigma_p + \gamma\sigma_m}{\sigma_p\sigma_m}\right)} - \left[\frac{\sigma_m}{2\left(\sigma_p + \sigma_m\right)}\right] e^{-1.15\left(\frac{\theta\sigma_p + \gamma\sigma_p + \gamma\sigma_m}{\sigma_p\sigma_m}\right)} \qquad (11)$$

Equations (10) and (11) were solved parametrically assuming that the weapon systems specification limits fall at the 3-sigma points $\left(\text{i.e., } \theta = 3\sigma_p\right)$ for the actual parameter value distribution for a series of values of the following ratios:

$$\frac{\sigma_m}{\sigma_p} \equiv \frac{\text{measurement deviation}}{\text{parameter deviation}} = \text{accuracy ratio}$$

$$\frac{\gamma}{\theta} \equiv \frac{\text{measurement limit}}{\text{specification limit}} = \text{decision ratio.}$$

The computed values were plotted as functions of $p(\alpha)$, $p(\beta)$, σ_m/σ_p, and γ/θ. The plots are shown to different scales in Figures 2, 3, and 4.

3. Multiple Check Probability

The discussion up to this point has been concerned with individual measurement error probability. It is often necessary, in the checkout of a UUT, to make two or more sequential tests on the same unit. Under such conditions the overall error probability becomes a function of the number of sequential tests, m, and the individual test probabilities, $p(\alpha)$ and $p(\beta)$. The multiple check probabilities $P(\alpha)$ and $P(\beta)$ may be computed from the following equations:

$$P(\alpha) = 1 - \prod_{i=1}^{m} \left[1 - p(\alpha)_i \right] \tag{12}$$

$$P(\beta) = 1 - \prod_{i=1}^{m} \left[1 - p(\beta)_i \right], \tag{13}$$

where

m = number of tests
$p(\alpha)$ = individual test probability of undetected defect
$p(\beta)$ = individual test probability of false alarm.

Assuming that $p(\alpha)$ and $p(\beta)$ are the same for all m tests, then equations (12) and (13) reduce to

$$P(\alpha) = 1 - [1 - p(\alpha)]^m \tag{14}$$

$$P(\beta) = 1 - [1 - p(\beta)]^m . \tag{15}$$

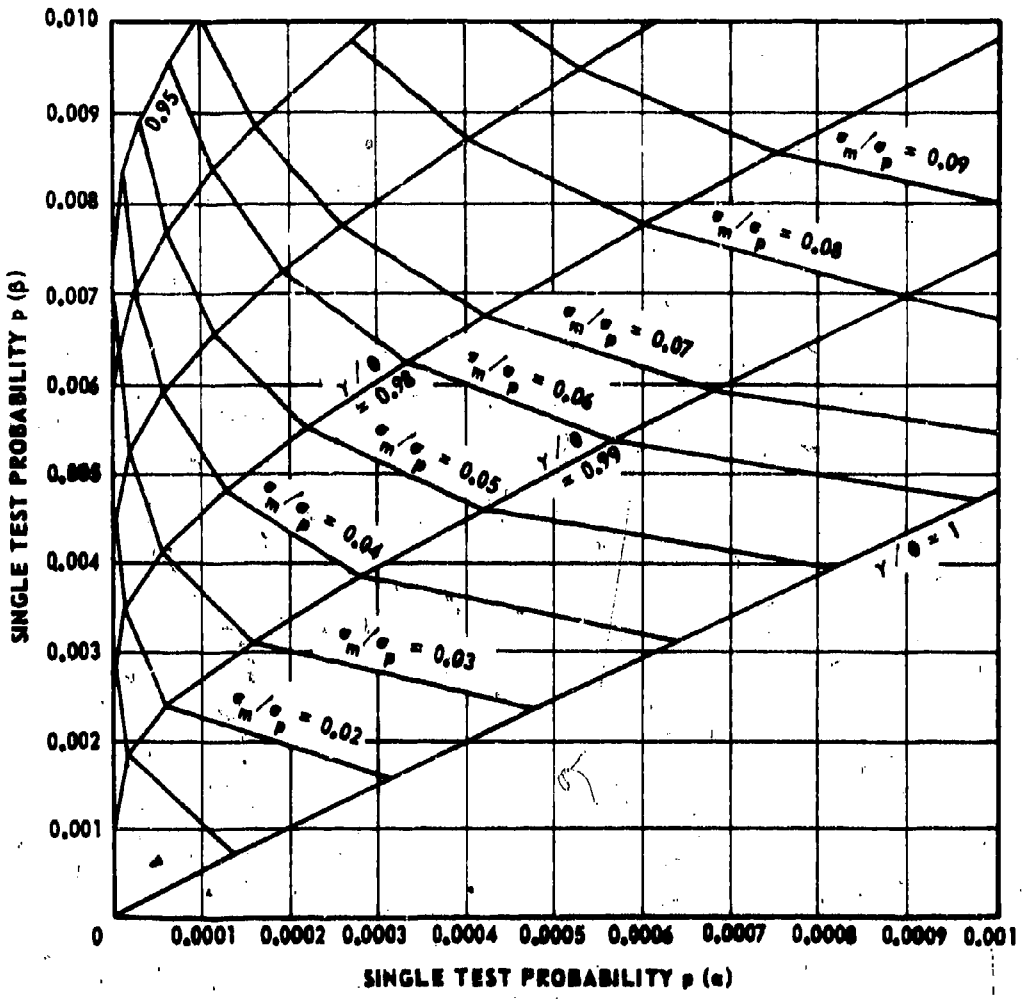


FIGURE 2. SINGLE TEST PROBABILITY VERSUS ACCURACY RATIOS AND DECISION LIMITS

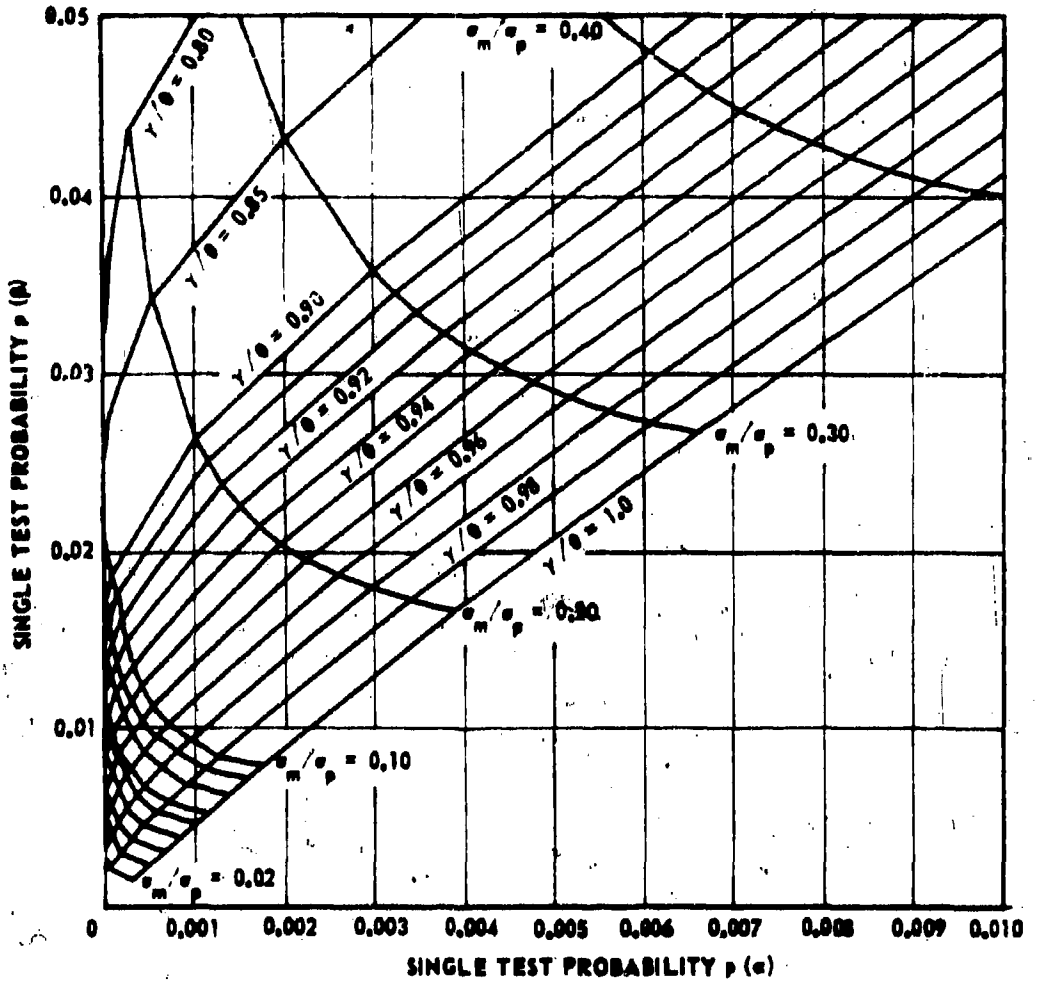


FIGURE 3. SINGLE TEST PROBABILITY VERSUS ACCURACY RATIOS AND DECISION LIMITS

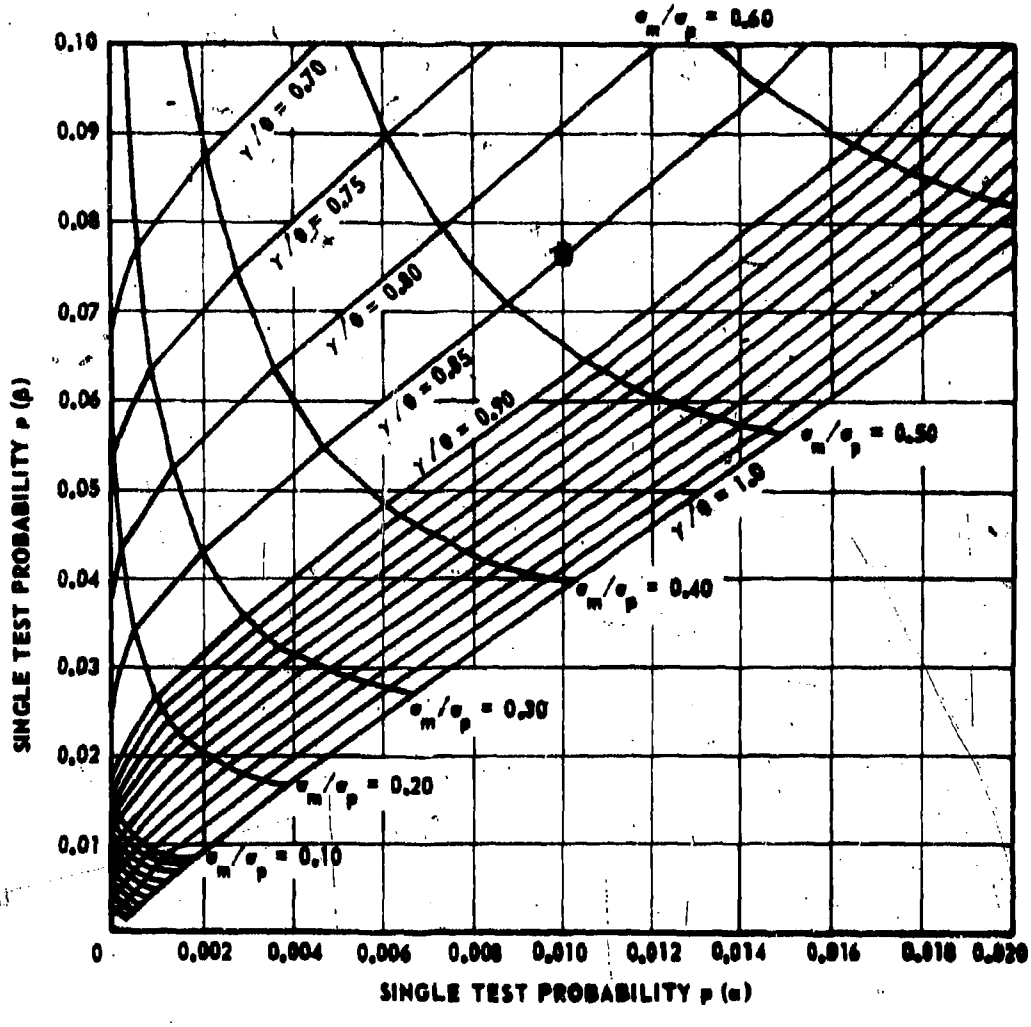


FIGURE 4. SINGLE TEST PROBABILITY VERSUS ACCURACY RATIOS AND DECISION LIMITS

Equations (14) and (15) were solved for a series of values of $p(\sigma)$, $p(\sigma)$ and n and are shown in plots of two different scales in Figures 5 and 6.

The more common way of expressing measurement accuracy is as a plus or minus percentage of full scale reading with a certain confidence. The standard deviations σ_m or σ_p can be expressed in percentage when the measurements are made at nearly full scale. The relationship between σ and percent is as follows:

$$\pm X\% = \left(\frac{\pm n\sigma}{F.S.} \right) 100 , \qquad (16.$$

where

$\pm X\%$ ≡ accuracy in percent full scale
σ ≡ standard deviation
n ≡ desired confidence level (i.e., 1, 2, 3, ... etc.)
F. S. ≡ full scale deflection of instrument.

A more complete and detailed discussion on the above derivation of error probability density functions and their relationship to test equipment may be obtained from Moon [5, 6]. The objective of this effort is to apply the mathematical models as shown to the design criteria of the LCSS-ETG equipment.

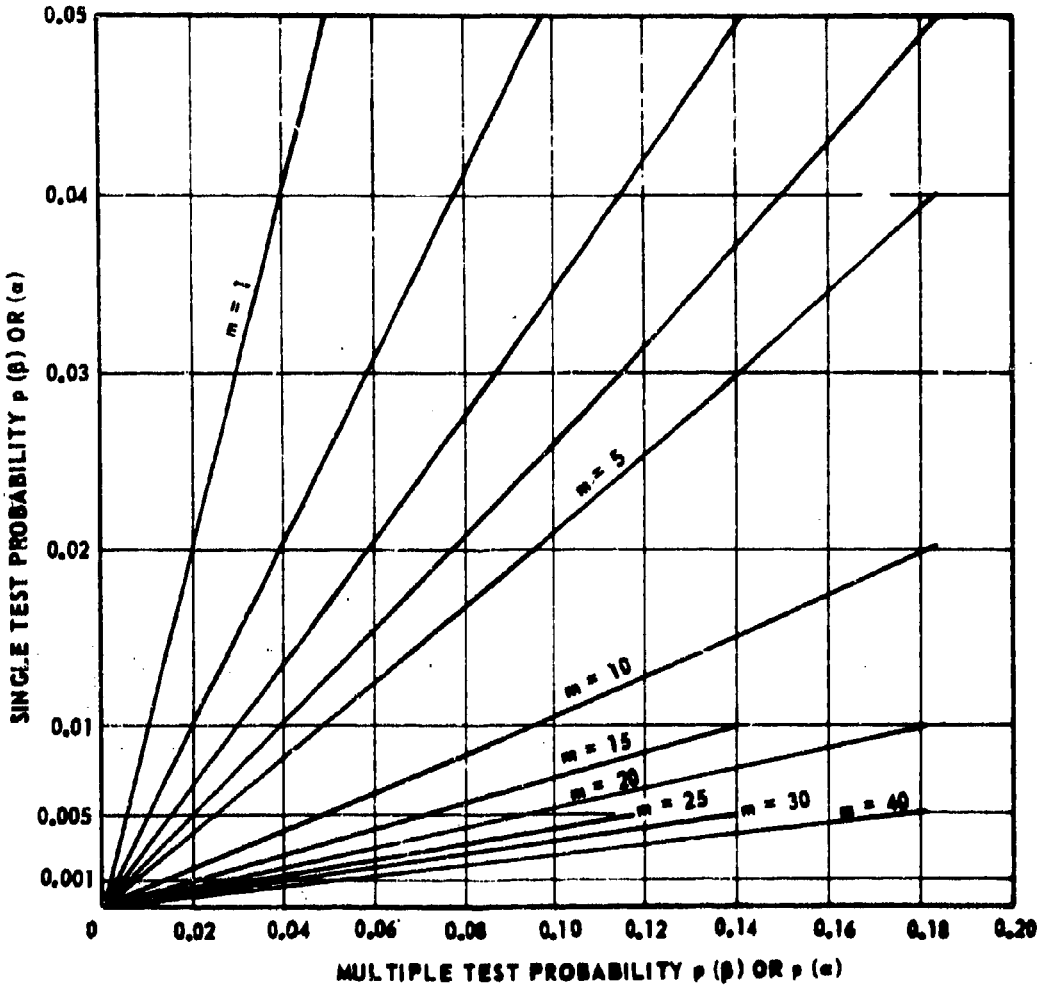


FIGURE 5. SINGLE TEST PROBABILITY [$p(\alpha)$, $p(\beta)$] VERSUS MULTIPLE TEST PROBABILITY [$P(\alpha)$, $P(\beta)$]

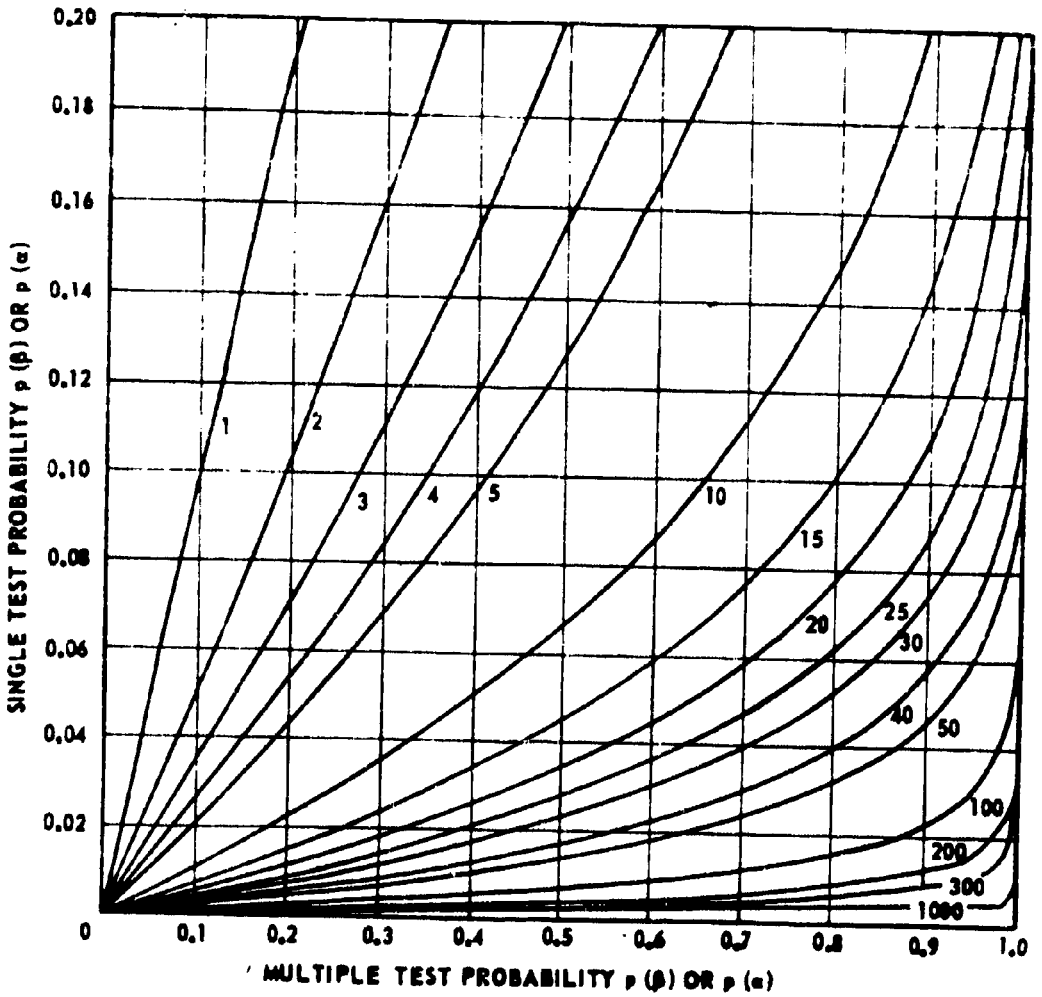


FIGURE 6. SINGLE TEST PROBABILITY [$p(\alpha)$, $p(\beta)$] VERSUS MULTIPLE TEST PROBABILITY [$P(\alpha)$, $P(\beta)$]

Section III. TEST PLAN

The purpose of the test program is to obtain a random sample of measurements to establish estimates of the accuracy, precision, and stability of the different measuring and stimulating systems of the LCSS-ETG over the full-scale ranges of the equipment. All values are, of course, only sample estimates based on limited samples of values from only two machines. Care should be taken when assuming that these two machines are representative of the population of machines.

1. Test Design

Two ETG sets were available for use. In view of the complexity of the ETG, it was decided to include in the test design the effects of time delay between the measurement command and the actual measurement. ETG specifications called for various time delays. In order to check for the effects of transients and drift upon short delays and long delays, respectively, it was decided to use the specified delay, and 3 times and 5 times the specified delay. The effects of repetitive measurements were also considered by including test time durations. The durations were established as four 1-hour tests and one 4-hour test periods per machine, and delay times. The test design is shown in Table II, from which it can be seen that each combination of machine, delay, and duration is tested for 4 hours for a total test time of 48 hours for all combinations.

In order to filter out as much test environment bias as possible, the test sequence was randomized. The randomization would tend to filter out

TABLE II. TEST DESIGN

Machine			2			5			Σ hr by duration
Tape delay			A	B	C	A	B	C	
Test run duration (Nom. hr)	hr	freq							
	1	4	4	4	4	4	4	4	24
	4	1	4	4	4	4	4	4	24
Σ hr by tape			8	8	8	8	8	8	48
Σ hr by machine			24			24			48

such effects as changes in calibrations due to periodic work variations, operators, time-of-day effects, etc. The objective of randomization is to increase the probability that the sample is random and increase the probability that inherent environmental biases are eliminated. The randomization was done using a random number generator coded for days, hours, machines, and delays. The randomized sequence as determined by the random number generator is shown in Table III. The machines are designated as 2 and 5 corresponding to their RCA serial numbers, ETG 3-002 and ETG 3-005, respectively. The delay times are designated as A, B, and C for the 1, 2, and 5 multiplier of specified delays, respectively. The numbers shown in parentheses in Table III refer to the test numbers used for identification in the computer programming. The numbers run sequentially from 1 through 48 corresponding to the total 48 hours of testing. With Tables II and III, all the data can be identified, classified, and grouped in any combination of test conditions for analysis.

2. Test Method

The test method was designed to simulate actual operational tests as performed on a weapon UUT by the English Language Program tapes. It was considered that this method would provide maximum data in a minimum amount of test time. Also, it was desired to leave the ETG sets available for other tests and uses as much as possible during their limited availability at Redstone Arsenal. It was designed to minimize operator error, biases, and interpretation simulating actual UUT test conditions.

A programmed semi-automatic test tape was compiled to make measurements and observations rapidly and automatically. The tapes were identical except for the programmed time delays between stimulus command and measurement execution. General instructions were provided the operator for running the tapes and all measurements were performed and printed automatically by the ETG. In programming the tapes, interrupts were inserted with the proper instructions to the operator where manual switching is required. It was specified that the operator, prior to running the test tape, would:

a) visually inspect the ETG machines for defective or missing components

b) check out the hook-up and main functional arrangement

c) perform all preliminary operational checks

d) make at least one successful run on the ETG with the RCA calibration and maintenance (C&M) tape

e) log all unusual conditions prior to and during the test tape run.

TABLE III. RANDOMIZATION OF LCSS TEST SCHEDULE

Date of Test	Hour of Test							Test Set 002			Test Set 005		
	0900	1000	1100	1200	1300	1400	1500	ΣA	ΣB	ΣC	ΣA	Σ B	Σ C
Feb.													
Tues 13		2C (1)								1			
Wed 14						2A (2)		1					
Thur 15			2B (3)						1				
Fri 16	2B (4)	2B (5)	2B (6)	2B (7)					1				
Mon 19							5A (8)				1		
Tues 20	5C (9)												1
Wed 21						2B (10)			1				
Tues 27	2C (11)		2A (12)					1		1			
Thur 29			5A (13)	5C (14)							1		1
March													
Fri 1	2C (15)	2C (16)	2C (17)	2C (18)		5B (19)				1		1	
Mon 4				5B (20)								1	
Fri 8		2A (21)	2A (22)	2A (23)	2A (24)			1					
Mon 11		2C (25)								1			
Tues 12					5C (26)								
Wed 13				5A (27)							1		
Thur 14		5A (28)	5A (29)	5A (30)	5A (31)						1		
Fri 15	5A (32)										1		
Mon 18	5B (33)	5C (34)	5C (35)	5C (36)	5C (37)							1	1
Tues 19					5C (38)								1
Wed 20	5B (39)	5B (40)	5B (41)	5B (42)								1	
Thur 21		5B (43)										1	
Tues 26					2B (44)				1				
Wed 27						2A (45)		1					

TABLE III. RANDOMIZATION OF LCSS TEST SCHEDULE (Concluded)

Date of Test	Hour of Test							Test Set 002			Test Set 003		
	0900	1000	1100	1200	1300	1400	1500	ΣA	ΣB	ΣC	ΣA	ΣB	ΣC
March Thur 28 Fri 29	2A (47)			2C (46) 2B (48)									
Σ Hours and Tapes	8	9	9	11	6	4	1	8	8	8	8	8	8
ΣΣ	48							24			24		

The standards used as references during the test program were located inside or adjacent to the ETG. The standards were provided and calibrated by the U. S. Army Metrology and Calibration Center. The following equipment was used as standards:

a)	D-C standard/D-C differential voltmeter	HP 710B
b)	Kelvin Vauley divider	ESI 722
c)	Electronic counter	HP 5245L
d)	Audio voltage standard	Holt Avs 323
e)	Ratio transformer	Gertsch PT-2
f)	Oscillator	HP 241 A
g)	RMS A-C differential voltmeter	John Fluke 931 A

The semi-automatic test tape has a general test procedure to program the appropriate ETG system to measure and print the corresponding values in the following sequence:

shorted input
external reference
self test reference
stimuli
external reference.

This sequence provides data to compute estimates:

measurement bias error (ϵ_b)
measurement accuracy error (ϵ_a)
stimulus setting error (ϵ_s)

all relative to a known external standard reference. The following FTG measurement parameters were written into the test program tapes as functional groups:

Test Operation Series Number	Parameter Measured
100	dc voltage
200	Resistance
300	Frequency
400	ac voltage (400 Hz)
500	ac voltage (50 Hz)

Test Operation Series Number	Parameter Measured
600	ac voltage (10 kHz)
700	ac voltage (1 MHz)
800	Pulse train

Each line on the programmed printout is referred to as a single observation. All observations as printed out are identified by their operation number and parameter. On the complete tape going through all series (100 through 800) there are 101 observations (i.e., measurements). During each hour of test, all measurements were replicated five times. Table IV is a compilation of the complete test tape sequence of measurements showing the parameters, test operations numbers, type test, ETG component, ETG full-scale range, test conditions and the expected measured value or standard value.

3. Test Measurements

A complete set of the test tape results are shown in Volume II, Appendix A. The data shown therein have been corrected for obviously bad data and printout errors. In the great majority the data are duplicates of the original test tapes as printed out by the ETG.

TABLE IV. ETG TEST OPERATION IDENTIFICATION CHART

Parameter	Test Operation	Type Test	ETG Component	ETG Full-scale Range	Test Conditions	Expected Measured Value (or STD)	Computations
dc Voltage	100	SI	DMM	10 Vdc		0	
	101	ER	DMM	10 Vdc	7.800 Vdc	7.800	
	102	STR	DMM	10 Vdc	-7.500 Vdc	-7.500	
	103	DC 1	DMM	10 Vdc	9.000 Vdc	9.000	
	104	DA 1	DMM	10 Vdc	9.000 Vdc	9.000	
	105	FR	DMM	10 Vdc	7.800 Vdc	7.800	
	106	SI	DMM	250 Vdc		0	
	107	ER	DMM	250 Vdc	223.00 Vdc	223.00	
	108	PSI - 6	DMM	250 Vdc	230.00 Vdc	230.00	
	109	ER	DMM	250 Vdc	223.00 Vdc	223.00	
	110	SI	DMM	1000 mVdc		0	
	111	ER	DMM	1000 mVdc	503.21 mVdc	503.21	
	112	DA2	DMM	1000 mVdc	900.00 mVdc	900.00	
	113	FR	DMM	1000 mVdc	503.21 mVdc	503.21	
Resistance	200	SI	DMM	10 kΩ		0	
	201	ER	DMM	10 kΩ	8.8198 kΩ	8.8198 kΩ	
	202	STR	DMM	10 kΩ	9.0 kΩ	9.000 kΩ	
	203	ER	DMM	10 kΩ	8.8198 kΩ	8.8198 kΩ	
	204	SI	DMM	100 Ω		0	
	205	ER	DMM	100 Ω	80.05 Ω	80.05	
	206	SI	DMM	1000 kΩ		0	
	207	ER	DMM	1000 kΩ	519.9 kΩ	519.9 kΩ	
Frequency	300	SI	W.C.	1000 kHz		0	
	301	FR	W.C.	1000 kHz	999,996 kHz	999,996 kΩ	
	302	STR	W.C.	1000 kHz	900,000 kHz	900,000 kΩ	
	303	FR	W.C.	1000 kHz	999,996 kHz	999,996 kΩ	
	304	SI	DMM	10 VRMS		0	
	305	ER	DMM	10 VRMS	5.625 VRMS, 100 Hz	5.625 VRMS	
	306	STR	DMM	10 VRMS	7.0 VRMS, 100 Hz	7.0 VRMS	
	307	ER	DMM	10 VRMS	5.625 VRMS, 100 Hz	5.625 VRMS	

Best Available Copy

TABLE IV. ETC TEST OPERATION IDENTIFICATION CHART (Continued)

Parameter	Test Operation	Type Test	ETG Component	ETG Full-scale Range	Test Conditions	Expected Measured Value (or STD)	Computations
ac Voltage, 400 Hz	404	SI	DMM	10 VRMS		0	
	405	ER	DMM	10 VRMS	5.628 VRMS, 400 Hz	0.317 VTRMS	$(5.628)^2/100 = 0.317$
	406	SI	DMM	10 VRMS		0	
	407	ER	DMM	10 VRMS	5.628 VRMS, 400 Hz	7.958 Vp	$(5.628)\sqrt{2} = 7.958$ Vp
	408	SI	DMM	10 VRMS		0	
	409	ER	DMM	10 VRMS	5.628 VRMS, 400 Hz	15.92 Vpp	$(5.628)\,2\sqrt{2} = 15.916$
	410	SG 1	DMM/WC	10 VRMS	9.000 VRMS, 400 Hz	9.000 VRMS	
	411	SG 1	DMM/WC	10 VRMS	9.000 VRMS, 400 Hz	0.081 VTRMS	$(9)^2/10^3 = 0.081$
	412	SG 1	DMM/WC	10 VRMS	9.000 VRMS, 400 Hz	0.9000 VRMS	SG 1 on 10:1 attenuator
	413	SG 1	DMM/WC	10 VRMS	9.000 VRMS, 400 Hz	0.400 kHz	
	414	SG 2	DMM/WC	10 VRMS	9.000 VRMS, 400 Hz	9.000 VRMS	
	415	SG P	DMM/WC	100 VRMS	90.00 VRMS, 400 Hz	90.00 VRMS	
	416	SG Q	DMM/WC	10 VRMS	9.00 VRMS, 400 Hz	9.00 VRMS	(90-deg delay)
	417	SG Q	DMM/WC	10 VRMS	9.00 VRMS, 400 Hz	9.00 VRMS	(180-deg delay)
	418	ER	DMM	10 VRMS	5.628 VRMS, 400 Hz	5.628 VRMS	
ac Voltage, 50 Hz	500	SI	DMM	10 VRMS		0	
	501	ER	DMM	10 VRMS	0.977 VRMS, 50 Hz	0.977 VRMS	
	502	SG 1		10 VRMS	1.1 VRMS, 20 Hz	1.1 VRMS	
	503	SG 1		10 VRMS	1.1 VRMS, 20 Hz	0.0200 kHz	
	504	ER	DMM	10 VRMS	0.977 VRMS, 50 Hz	0.955 VTRMS	$(0.977)^2 = 0.955$ VTRMS
	505	ER	DMM	10 VRMS	0.977 VRMS, 50 Hz	1.381 Vp	$(0.977)\sqrt{2} = 1.381$ Vp
	506	ER	DMM	10 VRMS	0.977 VRMS, 50 Hz	2.762 Vpp	$(0.977)\,2\sqrt{2} = 2.762$ Vpp
	507	SG 1		10 VRMS	1.1 VRMS, 20 Hz	1.21 VTRMS	$(1.1)^2/10^0 = 1.21$
	508	SG 1		10 VRMS	1.1 VRMS, 20 Hz	1.5565 Vp	$(1.1)\sqrt{2} = 1.5565$ Vp
	509	SG 1		10 VRMS	1.1 VRMS, 20 Hz	3.11 Vpp	$(1.1)\,2\sqrt{2} = 3.11$ Vpp
	510	ER	DMM	10 VRMS	0.977 VRMS, 50 Hz	0.955 VTRMS	$(0.977)^2/10^0 = 0.955$ VTRMS
	511	ER	DMM	10 VRMS	0.977 VRMS, 50 Hz	1.381 Vp	$(0.977)\sqrt{2} = 1.381$ Vp
	512	ER	DMM	10 VRMS	0.977 VRMS, 50 Hz	2.762 Vpp	$(0.977)\,2\sqrt{2} = 2.762$ Vpp

TABLE IV. ETC TEST OPERATION IDENTIFICATION CHART (Continued)

Parameter	Test Operation	Type Test	ETC Component	ETG Full-scale Range	Test Conditions	Expected Measured Value (or STD)	Computations
ac Voltage, 50 Hz	513	SI	DMM	250 VRMS		0	
	514	ER	DMM	250 VRMS	110.0 VRMS, 50 Hz	110.0 VRMS	
	515	SI	DMM	250 VRMS		0	
	516	ER	DMM	250 VRMS	110.0 VRMS, 50 Hz	121.0 VTRMS	$(110)^2/10^2 = 121.0$ VTRMS
	517	SI	DMM	250 VRMS		0	
	518	ER	DMM	250 VRMS	110.0 VRMS, 50 Hz	155.5 Vp	$(110)\sqrt{2} = 155.5$ Vp
	519	SI	DMM	250 VRMS		0	
	520	ER	DMM	250 VRMS	110.0 VRMS, 50 Hz	311.1 Vpp	$(110)\,2\sqrt{2} = 311.1$ Vpp
	521	SG P	DMM/WC	250 VRMS	110.0 VRMS, 20 Hz	110.0 VRMS	
	522	SG P	DMM/WC	250 VRMS	110.0 VRMS, 20 Hz	121.0 VTRMS	$(110)^2/10^2 = 121$ VTRMS
	523	SG P	DMM/WC	250 VRMS	110.0 VRMS, 20 Hz	155.5 Vp	$(110)\sqrt{2} = 155.5$
	524	SG P	DMM/WC	250 VRMS	110.0 VRMS, 20 Hz	311.1 Vpp	$(110)\,2\sqrt{2} = 311.1$
	525	SG P	DMM/WC	1000 kHz	110.0 VRMS, 20 Hz	0.020 kHz	
	526	ER		250 VRMS	110.0 VRMS, 50 Hz	110.0 VRMS	
	527	ER	DMM	250 VRMS	110.0	121.0 VTRMS	$(110)^2/10^2 = 121.0$
	528	ER	DMM	250 VRMS	110.0	155.5 Vp	$(110)\sqrt{2} = 155.5$
	529	ER	DMM	250 VRMS	110.0	311.1 Vpp	$(110)\cdot 2\sqrt{2} = 311.1$
ac Voltage, 10 kHz	601	ER	DMM	10 VRMS	0.614	0.614 VRMS	
	602	SG I		10 VRMS	1.1	1.1 VRMS	
	603	SG I		10 kHz	1.1	9.0 kHz	
	604	ER	DMM	10 VRMS	0.614 VRMS, 10 kHz	0.614 VRMS	
	605	ER	DMM	10 VRMS	0.614 VRMS, 10 kHz	0.8682 Vp	$(0.614)\sqrt{2} = 0.8682$ Vp
	606	ER	DMM	10 VRMS	0.614 VRMS, 10 kHz	1.736 Vpp	$(0.614)\,2\sqrt{2} = 1.736$ Vpp
	607	ER	DMM	250 VRMS	110 VRMS, 1 kHz	110 VRMS	
	608	ER	DMM	250 VRMS	110 VRMS, 1 kHz	121 VTRMS	$(110)^2/10^2 = 121$ VTRMS
	609	ER	DMM	250 VRMS	110 VRMS, 1 kHz	155.5 Vp	$(110)\sqrt{2} = 155.5$
	610	ER	DMM	250 VRMS	110 VRMS, 1 kHz	311.1 Vpp	$(110)\,2\sqrt{2} = 311.1$ Vpp
	611	SG P		250 VRMS	110 VRMS, 1 kHz	110 VRMS	
	612	SG P		250 VRMS	110 VRMS, 1 kHz	121.0 VTRMS	$(110)^2/10^2 = 121.0$ VTRMS
	613	SG P		10 kHz	110 VRMS, 1 kHz	1.0 kHz	
	614	ER	DMM	250 VRMS	110 VRMS, 1 kHz	110 VRMS	
	615	ER	DMM	250 VRMS	110 VRMS, 1 kHz	121.0 VTRMS	$(110)^2/10^2 = 121.0$ VTRMS
	616	ER	DMM	250 VRMS	110 VRMS, 1 kHz	155.5 Vp	$(110)\sqrt{2} = 155.5$ Vp
	617	ER	DMM	250 VRMS	110 VRMS, 1 kHz	311.1 Vpp	$(110)\,2\sqrt{2} = 311.1$ Vpp

TABLE IV. ETG TEST OPERATION IDENTIFICATION CHART (Concluded)

Parameter	Test Operation	Type Test	ETG Component	ETG Full-scale Range	Test Conditions	Expected Measured Value (or STD)	Computations
ac 1000 kHz	701	ER	DMM	10 VRMS	1.00 Vp, 100 kHz	1.00 Vp	
	702	ER	DMM	10 VRMS	1.00 Vp, 100 kHz	2.00 Vpp	
Pulse Train	801	ER	W.C.	1000 kHz	999.996 kHz	999.996 kHz	
	803	PGN	W.C.	1000 kHz	100 kHz (30 Vp)	99.9999 MS	
	804	PGN	DMM/W.C.	100 Vp	100 kHz (30 Vp)	30 Vp	Volts p in time period
	805	PGN	W.C.	1000 kHz	100 kHz (30 Vp)	99.9999 MS	Pulse rise time
	806	PGN	W.C.	1000 kHz	100 kHz (30 Vp)	0.005 MS	Pulse width
	807	PGN	W.C.	1000 kHz	100 kHz (30 Vp)	10 MS	Pulse delay
	809	ER	W.C.	1000 kHz	999.996 kHz	999.996 kHz	

Section IV. DATA ANALYSIS

1. General

The semiautomatic test tapes were run on the ETG-3 machines in accordance with the randomized sequence shown in Table III. On the tapes are 101 measurement commands. Each tape was replicated five times in each hour. Each machine was tested for 24 hours for a total of 48 hours. Therefore, the total number of observations is computed as follows:

$$N_T = 101 \text{ obs/rep.} \times 5 \text{ rep/hr} \times 48 \text{ hr} = 24,240$$

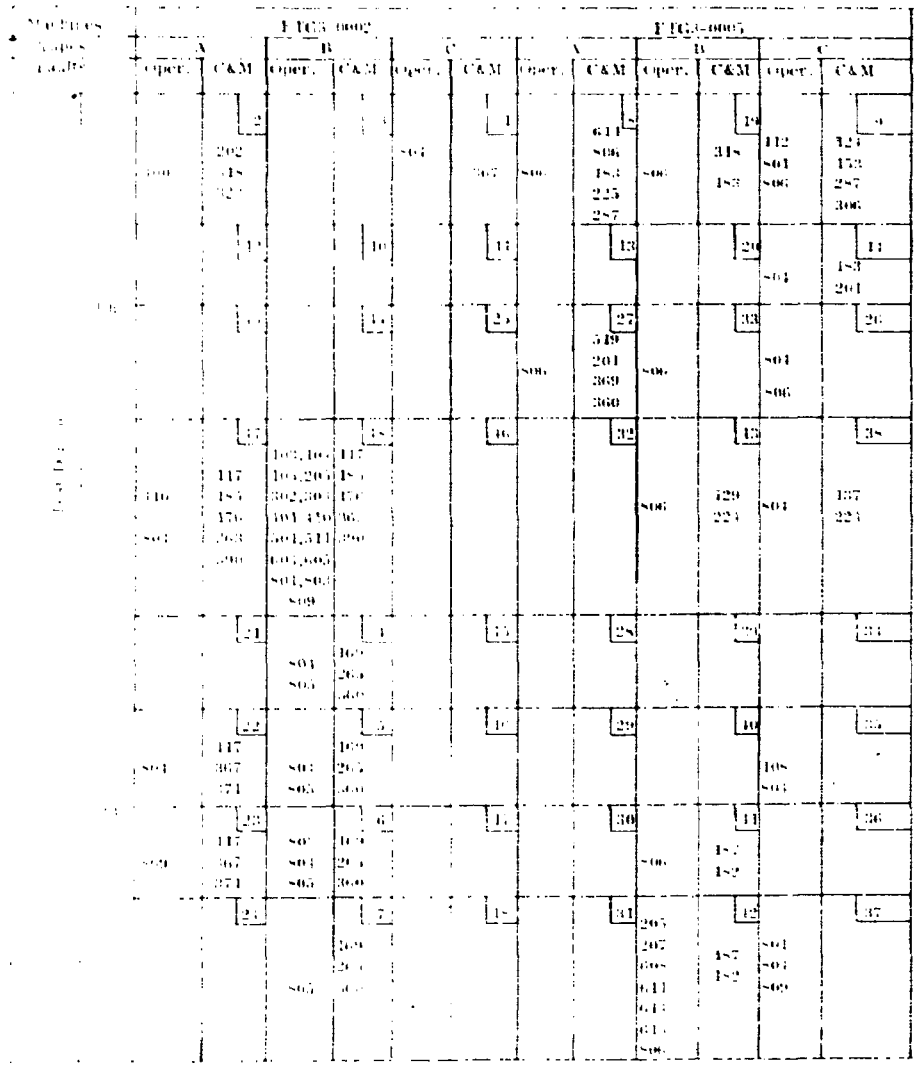
for all combinations of machines, time delays, durations and parameters. The 101 observations are broken down into a certain number for each parameter. A breakdown of the total number of observations for any combination is shown in Table V.

The measurements data were printed out by the ETG-3 printer on paper tapes. The data were coded for computer use, punched on input cards, and identified by test set numbers 1 through 48. The test set numbers included identification by machine, time delay, duration, date and hour of the day. The data used are subject to at least two sources of error. The first is machine printout errors. If a measurement was different from the last measurement of the same operation by less than a factor of two, it was included in the analysis as recorded. If the difference was greater than a factor of two, the measurement was excluded from the analysis and recorded as a machine "fault" or "dropout." All data replacing dropouts were estimated in accordance with established missing value procedures. Sample computations were made, including dropout values, to assess their effect on a computed statistic. The factor of two may not be the optimum factor but was selected to avoid any arbitrariness on determination of which values to exclude. The second source of error is the transfer of data from machine printout tapes to input cards. Obviously all errors were not eliminated and the results include the errors which could not be identified or which were overlooked. The data used for the computations are shown in Appendix B of Volume II. Lines are drawn under values estimated for dropouts. The total number of dropouts was 54 out of 24,240 observations. A list of dropouts by operation number is shown in Table VI, which also tabulates the dropouts by machine, delay time, and duration on an hourly basis. Shown for comparison are the operational faults on the C&M tapes which occurred at the time the test tapes were run. A description of the faults is listed below the table. Also shown is the ETG-3 parameter being tested by each operation. The 800 series of test operations (pulse train) had the largest

TABLE V. SCHEDULE OF OBSERVATIONS BY MACHINE, TIME DELAYS, DURATION AND PARAMETER

Machines		ETG3-0002						ETG3-0005						Σ by Parameter and Scale
Time Delays		A		B		C		A		B		C		
Durations		1	4	1	4	1	4	1	4	1	4	1	4	
Parameters and Scales	10 Vdc	120	=	=	=	=	=	=	=	=	=	=	=	1440
	250 Vdc	80	=	=	=	=	=	=	=	=	=	=	=	960
	1000 mVdc	80	=	=	=	=	=	=	=	=	=	=	=	960
	10 k ohm resistance	80	=	=	=	=	=	=	=	=	=	=	=	960
	100 ohm resistance	40												480
	1000 k ohm resistance	40												480
	1000 kHz frequency	80												960
	10 Vac 400 Hz	380												4560
	10 Vac 50 Hz	260												3120
	250 Vac 50 Hz	340												4080
	10 Vac 10 kHz	120												1440
	250 Vac 10 kHz	220												2640
	10 Vac 1000 kHz	40												480
	Pulse Train	140												1680
Σ	MTD	2020	=	=	=	=	=	=	=	=	=	=	=	
Σ	TD	4040		=		=		=		=		=		24,240
Σ	M	12,120						=						

TABLE VI. COMPILATION OF PROGRAM DROPOUTS AND C&M FAULTS



C&M Test Function Checks:

117 - test adapter (1A2) routing switches A1 and B1 and controller (1A1) switching data character SD5-2 and SD5-4 drivers.

124 - 1-megohm input impedance of digital multimeter (1A5) dc converter with dc standard output.

129 - digital multimeter (1A5) 0-9,999 K resistance measurement scale.

137 - resistance across capacitor C7. Test continues by checking +2.5 V dc limit on digital multimeter resistance converter output on 0-99.99 ohm and 0-999.9 ohm measurement scales.

153 - waveform converter (1A6) 0-99.99 V peak voltage measurement scale using probe and 7 V RMS output of self-test reference (1A7).

169 - test points connected to fixed resistors mounted on confidence and maintenance patchboard.

176-188 - check operation of 0-10 V dc power supply (1A8).

201 - 2 bits in pulse generator results counter 3 most significant decades. Also checks waveform converter results counter 3 least significant decades.

202 - changes oscillatory status level pulse repetition frequency to check waveform converter (1A6) results counter 2 bit in second decade.

224 - signal conditioner chopper drive frequency and digital multimeter 10 mVdc measurement range by measuring 500 mVdc output from 0-10 Vdc power supply (1A8).

265 - output of 0-10 Vdc power supply 3 (7A7) on MBA terminals before receipt of apply command and resistance at 0-10 Vdc power supply 3 (7A7-A1 output for a 21 ohm load.

287 - Measures 0-10 Vdc power supply 5 (1A18-A1 MBA output with supply programmed to 0 Vdc.

306 - 5-50 Vdc power supply (7A12) MBA output.

318 - frequency of signal generator (1A19) oscillator (SG1) output on MBA terminals. Then checks signal generator (SG1) and waveform converter (1A6).

323 - SG1 output frequency on MBA terminals.

360 - period of pulse generator (1A17) (PGN) main output. Test also checks unit apply function of PGN main output.

363 - pulse width of pulse generator (1A17) main output on MBA terminals.

367 - pulse width of pulse generator (1A17) main output.

369 - amplitude of pulse generator (1A17) main output.

371 - amplitude of pulse generator (1A17) main output. Test continues by checking amplitude of pulse generator main output.

398 - TPA11 by measuring continuity from RSA001 normally open contact to arm. RSA521 energized.

519 - (undetermined); 611 - (undetermined); 806 - (undetermined).

Operations: 100-105 10 Vdc; 106-109 250 Vdc; 110-113 100.0 mVdc; 200-203 10 K resistance; 204-205 0.1K resistance; 206-207 1000K resistance; 300-307 1000 kHz frequency; 308-318 10 VRMS 400 Hz; 508-512 10 VRMS 50 Hz; 513-529 250 VRMS 50 Hz; 601-606 10 VRMS 10 kHz; 607-617 250 VRMS 10 kHz; 701-702 10 VRMS 1000 kHz; 801-809 pulse train.

number of dropouts with 32. Test hour 48 had the largest number of dropouts by the hour with 15. A correlation of dropouts with C&M faults is evident. In the table the hourly test set numbers are in the upper right square. In test set 2, operation 400 was a dropout. Operation 400 measures 10 VRMS, 400 Hz ac voltage. The corresponding C&M faults, obtained from the machine daily log, which could be the cause, is 318 and 323 which checks the signal generator frequency. A complete analysis is beyond the scope of this effort, but would give additional insight into test set problems.

2. Statistical Computations

The main objective of this program was to determine an estimate of the standard error of measurement (measurement standard deviation) for the different parameters and scales of the machine. This is referred to earlier as σ_m. This statistic would be used in conjunction with the weapon system parameter standard deviation σ_p to obtain an estimate of the accuracy ratio. This ratio is needed to determine realistically the statistical probability of undetected defects and false alarms in the go/no-go chain. However, an attempt has been made to estimate other characteristics of the ETG-3.

When a measurement is made by the ETC-3, there are three primary sources of error. These are:

a) inherent machine bias error, ϵ_b

b) machine measurement accuracy error, ϵ_a

c) machine stimulus setting error, ϵ_s.

To illustrate, suppose the English Language Test Program Tape commands a certain voltage be applied to the UUT with a specified response to that stimulus to be measured. The circuits providing that stimulus may be effected by a previous operation (bias), the measuring device may sense the bias and read erroneously (accuracy) or the bias and accuracy may be acceptable but the voltage applied was not the value specified (stimulus setting). Naturally, combinations of all three may exist on any given measurement.

Estimates of bias error, ϵ_b, were obtained by programming a shorted input (SI) operation as shown in Table IV and measuring for a response. Estimates of measurement error, ϵ_a, were obtained by measuring an accurately known value from one of the external reference (ER) standards mentioned

previously and subtracting the bias error. Estimates of the stimulus setting error were obtained by programming a stimulus from one of the ETG-3 components and subtracting the bias and measurement errors.

The error estimates were computed with the following equations:

$$\epsilon_b = \frac{\sum_{i=1}^{N} \left(SI_i - EV \right)}{N} , \tag{17}$$

where

SI ≡ shorted input value
EV ≡ expected value
N ≡ number of responses.

$$\epsilon_a = \frac{\sum_{i=1}^{N} \left(ER_i - EV \right)}{N} - \epsilon_b , \tag{18}$$

where ER_i ≡ external reference measured value.

$$\epsilon_s = \frac{\sum_{i=1}^{N} \left(MV_i - EV \right)}{N} - \epsilon_a , \tag{19}$$

where MV ≡ measured value of ETG-3 component (e.g., DA-1, DC-1, STR). The above computed errors are the means of the differences between the actual (or expected) value and the measured (or unexpected) value.

The standard deviation of measurement $\left(\sigma_m \right)$ is computed by the following equation:

$$S_M^2 = \frac{\sum_{i=1}^{N} \left[(MV - EV)_i - \overline{(MV - EV)} \right]^2}{N - 1} , \tag{20}$$

which is the measurement variance. The standard deviation is simply the square root (S_M) of the equation and is designated S rather than σ since it is computed from a limited sample of data.

The statistics as computed with the above equations were computed in two ways. The first computations were made for each hour (test set) of testing. The purpose of this was to use the estimates of these statistics as response variables for a subsequent analysis of variance. The analysis of variance enables one to test for significant differences between factors. In addition to testing for significant differences between the main factors (i.e., machines, time delays and durations), tests can also be made for differences between second order or interaction effects (e.g., machines × time delay effects) as well as third order interaction effects. The second set of calculations was made using all observations for each parameter. These are referred to as the overall values of ϵ_b, ϵ_a, ϵ_s, and S_M. These calculations are more sensitive to subsequent tests due to the much larger sample size and greater degree of freedom.

The hourly computations are shown in Volume II, Appendix C. These values were computed with small sample size and degrees of freedom and are used only in the analysis of variance computations. The overall statistics are shown in Table VII. While the statistics ϵ_b, ϵ_a, and ϵ_s give some indication of the machine error by parameter and scale, the measurement standard deviation S_M is the important statistic from the standpoint of machine variability and determination of statistical control of desired probability levels of "undetected defects" and "false alarms," $p(\alpha)$ and $p(\beta)$.

3. Analysis of Variance (ANOVA)

The hourly computations of ϵ_b, ϵ_a, ϵ_s, and S_M^2 were used as response variables for analyses of variance of all parameters and scales. Hence, there were 4 × 14 = 56 analyses performed. The statistical models are

$$\epsilon_b(i, j, k, l) = \mu + A_i + B_j + C_k + AB_{ij} + AC_{ik} + BC_{jk} + ABC_{ijk} + e_{l(ijk)} \qquad (21)$$

and similarly for ϵ_a, ϵ_s, and S_M^2. This model assumes that each response of errors ϵ_b, ϵ_a, ϵ_s and variance S_M^2 is the algebraic sum of:

TABLE VII. COMPILATION OF OVERALL COMPUTED ERRORS, STANDARD DEVIATION AND PERCENTS BY PARAMETER AND SCALE

No.	Parameter and Scale	Computed Values					Specified		* Comp. % Value ÷ Specified Value
		ϵ_b	ϵ_A	ϵ_S	S_M	±% 3 S_M	S	±% (3 S)	
1	10 Vdc	0.000946	-0.00197	0.00106	0.02515	0.8	0.0013	0.04	
2	250 Vdc	0.147	-0.0798	1.77	2.975	3.6	0.0666	0.08	
3	1000 mVdc	0.349	-3.46	0.476	2.40	0.7	3.333	1.2	*
4	10 kΩ resistance	0.00327	-0.00171	0.00396	0.0098	0.3	0.0333	1.03	*
5	100 Ω resistance	0.895	-0.843	1.32	0.658	2.0	0.387	1.73	
6	1000 kΩ resistance	0.193	-1.13	0.758	0.747	0.2	3.33 k	1.00	*
7	1000 kHz frequency	0.00100	-0.0953	0.0485	1.52	0.5	0.0037 k	0.0011	
8	10 VRMS ac 400 Hz	0.370	-0.745	0.747	0.636	19.1	0.01	0.3	
9	10 VRMS ac 50 Hz	0.00453	0.421	0.00304	0.0867	2.6	0.01	0.3	
10	250 VRMS ac 50 Hz	2.93	-2.21	3.58	2.62	3.1	0.0666	0.08	
11	10 VRMS ac 10 kHz	0.00453	0.212	-0.0663	0.2088	6.3	0.01	0.3	
12	250 VRMS ac 10 kHz	2.93	-2.09	3.48	2.46	3.0	0.0666	0.08	
13	10 VPP ac 1.0 mHz	0.00453	0.01520	0.00453	0.0364	1.1	0.00108	3.25	*
14	Pulse train	0.925	-1.67	1.52	1.96	0.6	NA	NA	

a) a universal mean of the error μ (i.e., the true error or variance)

b) a machine effect on the error or the variance, C_k

c) a time delay effect on the error or the variance, B_j

d) a duration effect on the error or the variance, A_i

e) an interaction effect on the error or the variance, AB_{ij}, AC_{ik}, BC_{jk}, ABC_{ijk}

f) a random residual experimental error, $e_{l(ijk)}$

Since the model is fixed, none of the effects can be determined absolutely. They can be measured only as differential deviations, i.e.:

a) the A_i effects as deviations from μ

b) the B_j effects as deviations from μ

c) the C_k effects as deviations from μ

d) the AB_{ij}, AC_{ik}, BC_{jk} as deviations from the $A_i + B_j$, $A_i + C_k$, and $B_j + C_k$ respectively

e) the ABC_{ijk} as deviations from $A_i + B_j + C_k$.

Because of the large volume of data and the number (56) of analyses required to cover all combinations of parameters and statistics, the analyses were computer programmed. The program used was obtained from Edwin Bartee and was compiled by J. A. Svestka. A printout of the computer program is shown in Volume II, Appendix D, including dimensions and correspondence statements and subroutines. Each combination of machines, time delays, and duration, of which there are 12 (2 × 3 × 2), had four responses in each cell. Since each hour of test was replicated five times, four responses, which represents 4 hours of testing, was the average value for that hour. Table VIII is a completely coded layout of the data input to the analysis of variance. The table identifies the factors in the mathematical model [Equation (21)], the test set numbers, the date, and the hour. To illustrate, the four response values for the cell representing a 1-hour test duration (A_i), with the specified delay time (B_j), on the ETG-3-0002 machine (C_k) are data test sets 2, 12, 45, and 47 which are coded as 1, 2, 3, and 4 for input to the computerized analysis of variance. A computer printout of the input data to each analysis of variance and the ANOVA results are shown in Volume II, Appendix E.

TABLE VIII. COMPILATION OF HOURLY TESTS BY MACHINE, TAPE, DURATION, DAY AND HOUR

Machines (C_k)		ETG3-0002			ETG3-0005			Σ Duration
Tapes (B_j)		1	3	5	1	3	5	
Test Duration (A_i)	1 hr	2 [1] 2/14-11	3 [9] 2/15-11	1 [17] 2/13-10	8 [25] 2/19-15	19 [33] 3/01-14	9 [41] 2/20-9	24
		12 [2] 2/23-11	10 [10] 2/21-11	11 [18] 2/23-9	13 [26] 2/29-11	20 [34] 3/04-12	14 [42] 2/29-12	
		45 [3] 3/27-14	44 [11] 3/26-13	25 [19] 3/11-10	27 [27] 3/13-12	33 [35] 3/18-9	26 [43] 3/12-13	
		47 [4] 3/29-9	48 [12] 3/29-12	46 [20] 3/28-12	32 [28] 3/15-9	43 [36] 3/21-10	38 [44] 3/19-13	
	4 hrs	21 [5] 3/08-10	4 [13] 2/16-8	15 [21] 3/01-8	28 [29] 3/14-10	39 [37] 3/20-8	34 [45] 3/18-10	24
		22 [6] 3/08-11	5 [14] 2/16-9	16 [22] 3/01-9	29 [30] 3/14-11	40 [38] 3/20-9	35 [46] 3/18-11	
		23 [7] 3/08-12	6 [15] 2/16-10	17 [23] 3/01-10	30 [31] 3/14-12	41 [39] 3/20-10	36 [47] 3/18-12	
		24 [8] 3/08-13	7 [16] 2/16-11	18 [24] 3/01-11	31 [32] 3/14-13	42 [40] 3/20-11	37 [48] 3/18-13	
Σ Tapes		8	8	8	8	8	8	ΣΣ = 48
Σ Machine		24			24			

Code:

1 ← Code number for analysis of variance data input
2 ← Test set identification number
Date of test in 1968
2/14-14 ← Hour of test (e.g. 1400 or 2 pm)

4. Test of Hypotheses

A statistical hypothesis is an assumption about the population being sampled. It usually consists of assigning a value to one or more parameters of the population. A test of a hypothesis is simply a rule by which a hypothesis is either accepted or rejected. The rule is usually based on sample or test statistics used to test the hypotheses. The critical region of a test statistic consists of all values of the test statistic where the decision is made to reject or accept the hypothesis. Since hypothesis testing is based on observed sample statistics computed on N observations, the decision is always subject to errors. If the hypothesis is really true and is rejected by the sample statistic, a Type I error is committed. The probability of a Type I error is designated as α. If the hypothesis is accepted when it is not true, i.e., if some alternate hypothesis is true, then a Type II error has been made. The probability of a Type II error is designated β.

The overall values of ϵ_b, ϵ_a, and ϵ_s for all parameters were tested with the following hypotheses for the sample error statistics:

$$Ho_1: \epsilon_b = 0 \quad \alpha = 0.05$$

$$Ho_2: \epsilon_a = 0 \quad \alpha = 0.05$$

$$Ho_3: \epsilon_s = 0 \quad \alpha = 0.05 .$$

The alternate hypotheses are:

$$H1_1: \epsilon_b \neq 0$$

$$H1_2: \epsilon_a \neq 0$$

$$H1_3: \epsilon_s \neq 0 .$$

The computed estimates of the errors are the differences in means. The hypothesis (Ho) is that the errors are not significantly different from zero. This is based on the assumption that the universe mean of the errors is zero. If the basic hypothesis is rejected [at the $(1 - \alpha)$ level of confidence] the alternate (H_1) is accepted. However, if the basic hypothesis is accepted at a specified level of confidence there is still a chance that an error of the first kind has been made. In testing hypotheses pertaining to the universe mean the procedure is simplest if the standard deviation of the universe is known. In this case the sample errors are treated as having a normal distribution with

Best Available Cop

a mean equal to the universe mean; but, the universe standard deviation is unknown. Hence, the correct test statistic is the 't' statistic. The 't' statistic is used when the standard deviation must be estimated from the sample data. The 't' statistic is computed as follows:

$$'t'_{(\alpha)} = \frac{\bar{X} - \bar{X}'}{S/\sqrt{N}} \quad , \tag{22}$$

where

$\bar{X}$ = calculated mean
$\bar{X}'$ = universe mean
N = sample size
α = risk = 0.05

and

$$S = \left[\frac{\sum_{i=1}^{N} \left(X_i - \bar{X} \right)^2}{N - 1} \right]^{1/2} \tag{23}$$

If the computed value of 't' is less than the table value, for N - 1 degrees of freedom and $\alpha = 0.05$, then the basic hypothesis is accepted and the alternate is rejected. A computer printout of the results is shown in Volume II, Appendix F. A summary is shown in Table IX in the row labeled "'t' Test of Overall Values."

In the ANOVA the basic hypotheses are as follows:

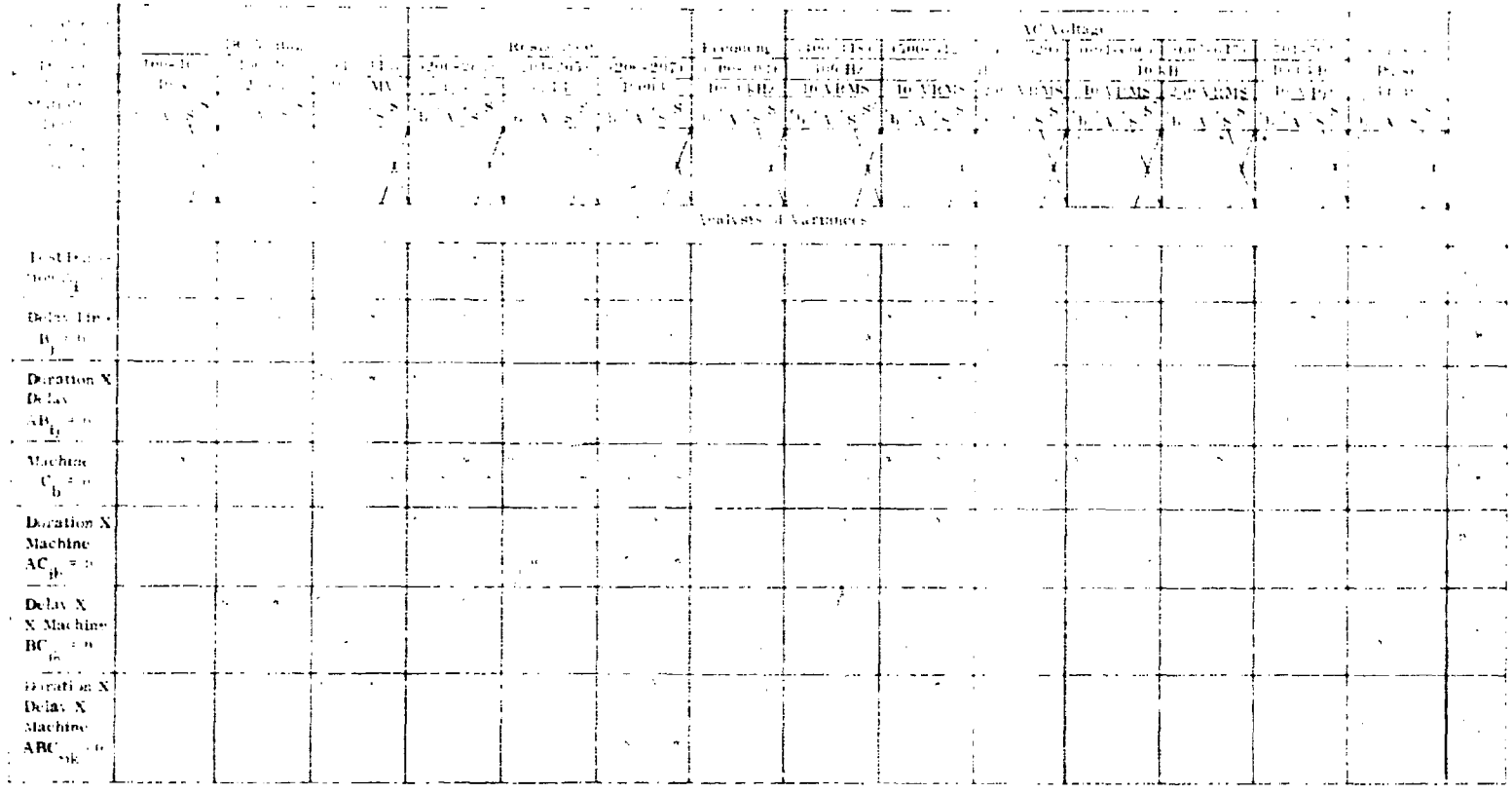
$$\text{Ho: } A_i = 0 \quad \alpha = 0.10$$

$$\text{Ho: } B_j = 0 \quad \alpha = 0.10$$

$$\text{Ho: } C_k = 0 \quad \alpha = 0.10$$

and similarly for the second order (ABij, etc.) effects and the third order (ABCijk) effects. The alternate hypotheses are:

TABLE IX. COMPILATION OF RESULTS OF TESTS OF HYPOTHESES ('t' TEST AND 'F' TESTS)



Indicates in 't' test that hypotheses: $\mu = 0$ is rejected and alternate hypotheses: $\mu \neq 0$ is accepted.

Indicates in ANOVA 'F' test that hypotheses: $A_i = 0$; $B_j = 0$; $C_k = 0$; etc. are rejected and alternate hypotheses: $A_i \neq 0$; $B_j \neq 0$; $C_k \neq 0$; etc. are accepted.

$$H_1: A_i \neq 0$$

$$H_1: B_j \neq 0$$

$$H_1: C_k \neq 0$$

and similarly for the second and third order effects.

The objectives of basic hypothesis tests are to determine that the main factors of duration (A_i), time delay (B_j) and machines (C_k) as well as interactions between main factors $(e.g., AB_{ij})$ do not significantly effect the response variables, ϵ_b, ϵ_a, ϵ_s, and S^2 and they are essentially zero, i.e., there is no treatment effect. The test statistic is the F distribution which is the ratio of two independent chi-square distributions. This means that the F distribution is the ratio of the mean squares between treatments to the mean squares (MS) within treatments or mean square for error.

$$F_{(\alpha)} = \frac{\text{MS (treatments)}}{\text{MS (error)}} , \tag{24}$$

where

$$\text{MS (treatments)} = \frac{\text{sum of squares}}{\text{degrees of freedom}} \tag{25}$$

$$\text{MS (error)} = \frac{\text{sum of squares}}{\text{degrees of freedom}} . \tag{26}$$

If the computed value of F is equal to or greater than the table value of F for the set level of confidence (α) and the proper degrees of freedom then Ho: is rejected and Hl: is accepted. A summary of the results of the ANOVA is shown in Table IX. The ANOVA printouts for each statistic and parameter are shown in Appendix G of Volume II.

Section V. DISCUSSION

This analysis has been primarily concerned with the estimation of the measurement variance and secondarily with the bias error, measurement error and stimulus setting error. The estimation of these statistics has been done with a sample of data taken from two machines. The analyses have also estimated the effects of time delays and duration of tests on these statistics. The statistics were computed for fourteen different measurement parameters and scales performed by the ETG-3 test sets. These parameters are not all of the functions of the machines and represent, at best, only a partial test program. The results represent estimates for a small sample of machines, factors, and measurements, and, should be accepted as such. The results shown in Tables VII and IX will be discussed briefly. The discussion will be categorized by machine parameters which are probably of greater interest to machine users than the sample statistics.

1. 10-Volt dc Scale

In Table VII, the estimated value of the bias error is 0.000946, the accuracy error is -0.00197 and the stimulus setting error is 0.00106. The standard deviation is 0.02515. Table IX shows that the hypothesis that the bias error is not significant Ho: $\epsilon_b = 0$ is rejected. The hypotheses that $\epsilon_a = 0$ and $\epsilon_s = 0$ were accepted in the 't' tests. In the ANOVA portion of Table IX, it is shown that machine effects were significant; i.e., Ho: $C_k = 0$ was rejected for both statistics ϵ_b and ϵ_s. In Table VII, the 3S level of standard deviation expressed as percent full scale is 0.8 as compared with the specified (assumed 3S) value of 0.04 percent. This is a factor of 20 higher than specified assuming the specified values are correct.

2. 250-Volt dc Scale

For this parameter, the hypotheses that ϵ_b and $\epsilon_s = 0$ were rejected indicating that there are significant bias and stimulus errors on this parameter. The ANOVA reveals that the contributing factors to ϵ_b and ϵ_s are main effects of delay time and machines and interactions between delay times and machines. The estimated values as shown in Table VII are $\epsilon_b = 0.147$, $\epsilon_a = -0.0798$, $\epsilon_s = 1.77$ and S = 2.975. The 3S level expressed as a percent = 3.6 as compared with the specified value of 0.08 percent, a factor of about 45 greater.

3. 1000-Millivolt dc Scale

The sample of data reveals that all errors were significantly different from zero. The ANOVA indicates that all factors significantly contribute to bias error. Delay times, machines and delay time-machine interactions contribute to ϵ_a. All factors except durations contribute to ϵ_s and delay times and machines contribute to the variance, S^2. The estimated values in Table VII are $\epsilon_b = 0.349$, $\epsilon_a = -3.46$, $\epsilon_s = 0.476$, and $S = 2.40$. The 3S level is 0.7 percent as compared to the specified value of 1.2 percent.

The rest of the parameters are left to the reader. An overall look at Table VII shows that the machines have met the specified standard error of measurement on the 1000-millivolt dc scale, the 10-kilohm resistance scale, the 1000-kilohm resistance scale and the 10-volt PP ac voltage scale at 1-megahertz frequency. The 100-ohm resistance scale is close with 2 percent as compared with 1.73 percent specified. The 10-volt ac 400-hertz scale is not good with a 19 percent as compared with 0.3 percent specified. The worst seems to be the 1.0 megahertz frequency parameter with 0.5 percent as compared with 0.0011 percent specified a factor of about 500 greater.

Table IX shows the bias error was significant for all parameters and scales. Accuracy error was significant for all parameters except 10 volts dc, 250 volts dc, and 1000-kilohertz frequency, i.e., 11 out of 14. Stimulus setting error was significant on all parameters except 10 volts dc, 1000-kilohertz frequency and 10 volts ac, 50 hertz, i.e., 11 out of 14.

The ANOVA portion of Table IX shows that test duration $\left(A_i\right)$ was detected as a significant effect on the responses of ϵ_b, ϵ_a, ϵ_s, and S^2 in only 4 analyses out of 56. The main effect, delay time $\left(B_j\right)$ was found to be significant 21 times out of 56. The main effect of machines $\left(C_k\right)$ was found to be significant 46 times out of 56. The second order effect, AB_{ij}, 6 times; AC_{ik}, 13 times; BC_{jk}, 11 times; and ABC_{ijk}, 8 times out of 56. Machine effects far outweigh the other main and interaction effects.

The estimates of the standard deviation of measurement $\left(S_M\right)$ as obtained from this sample will be used as an estimate of σ_m. Previously, the use of σ_m in the determination of an accuracy ratio was discussed. Also, the use of the accuracy ratio, decision limits and $p(\alpha)$ and $p(\beta)$ in the determination of the

statistical capability of the test sets for single tests. With the values of σ_m as obtained a few examples will be made with actual values of the weapon system components tolerances (σ_p). Also, decision limits will be assumed. The parameters 1000-millivolt dc scale, 10-kilohm resistance scale, 1000-kilohm resistance scale met the MIS-6000 specifications. These parameters and their values are used in comparison with missile specification values for components which would require testing on these scales. Also, one of the parameters which did not meet MIS-6000 specs, the 10-volt dc scale, is included for comparison purposes. The sample estimated values of the standard deviation (σ_m) for each of these parameters is divided by the component standard deviation (σ_p) to obtain an accuracy ratio (σ_m/σ_p). Decision limit ratios (ζ/θ) of 1.00, 0.95, and 0.90 were assumed for this example.

The values for each weapon system are actual values obtained from the system specifications. The TOW system tolerances are based on 2σ levels and the SHILLELAGH system tolerances are based on 3σ levels. The values of σ for both systems were obtained from the system tolerances accordingly. The values of $p(\alpha)$ and $p(\beta)$ were obtained from the computer printouts used to plot Figures 2, 3, and 4. The results are shown in Table X.

The system components are as shown in Table XI. Similar analyses can be made for other parameters and scales and weapon systems.

Table X demonstrates the relationships between accuracy ratios, decision limits, undetected defects and false alarms for single tests for actual system components. It also demonstrates the trade-offs and compromises available to hold certain $p(\alpha)$ and $p(\beta)$. A component of Shillelagh has a 50-percent tolerance of measurement on the 1000 millivolt dc scale which gives an accuracy ratio of 0.014 or about 70 to 1 when compared with sample σ_m found for the ETG on that scale. Therefore, the probability is that practically no undetected defects will get by the tests; however, some false alarms varying from 1 out of 1000 to 1 out of 100 depending on the decision limits will probably occur. An accuracy ratio of 0.504 (about 2 to 1) shown for TOW component on 10-volt dc scale will result in about 1.5 units out of 100 at decision limit of 1.00 to 1.1 units out of 100 at decision limit of 0.90 passing with "undetected defects." Similarly, 5.6 and 6.6 "false alarms" will occur out of 100 tests.

TABLE X. SINGLE TEST $p(\alpha)$ AND $p(\beta)$ FOR RATIOS OF CALCULATED ETG σ_m AND SYSTEM SPECIFIED σ_p

	Parameter and Scale	10 Vdc			1000 MVdc			10 kΩ Resistance			1000 kΩ Resistance		
ETG-3 STD Dev. σ_m		0.0252 V*			2.40 mV			0.0098 kΩ			0.747 kΩ		
TOW (2σ)	Sys. STD Dev. σ_p	0.050 V			5.00 mV			0.050 kΩ			5.00 kΩ		
	Accuracy Ratio σ_m/σ_p	0.504 (1.97 to 1)			0.48 (2.08 to 1)			0.196 (5.62 to 1)			0.149 (6.72 to 1)		
	Decision Limit (γ/θ)	1.00	0.95	0.90	1.00	0.95	0.90	1.00	0.95	0.90	1.00	0.95	0.90
	$p(\alpha)$≈**	0.015	0.013	0.011	0.014	0.011	0.008	0.003	0.0018	0.0009	0.0026	0.0015	0.0005
	$p(\beta)$≈**	0.056	0.059	0.066	0.054	0.056	0.060	0.0166	0.019	0.026	0.011	0.016	0.021
Shillelagh (3σ)	Sys. STD Dev. σ_p	0.233			167.0			0.0334			3.334		
	Accuracy Ratio σ_m/σ_p	0.108 (9.26 to 1)			0.014 (71.4 to 1)			0.29 (3.45 to 1)			0.224 (4.46 to 1)		
	Decision Limit (γ/θ)	1.00	0.95	0.90	1.00	0.95	0.90	1.00	0.95	0.90	1.00	0.95	0.90
	$p(\alpha)$**	0.0017	0.0003	0.00007	0.0	0.0	0.0	0.006	0.004	0.0029	0.0047	0.0027	0.0030
	$p(\beta)$**	0.0078	0.012	0.019	0.001	0.007	0.01	0.026	0.03	0.35	0.018	0.022	0.02

* This parameter and scale does not meet requirements of ETG MIS-6000 specifications based on the sample of data obtained in this effort.

** Multiply $p(\alpha)$ and $p(\beta)$ values in table by 100 to obtain units per 100 of undetected defects and false alarms.

TABLE XI. IDENTIFICATION OF SYSTEM COMPONENTS USED IN SAMPLE COMPUTATIONS OF TABLE X

Parameter	10 Vdc	1000 mVdc	10-kΩ Resistance	1000-kΩ Resistance
TOW	Yaw CSG Card, Part No. 10191893	Excitation Generator and Self Test Card, Part No. 10191908	Pitch Command Signal Generator, Part No. 10225421	Pitch Command Signal Generator, Part No. 10225421
Shillelagh	SDC Program, Test Step 304, Part No. 1500	Transmitter Signal Analyzer, Test No. 247	19 Vdc Regulator Card, G.M. System Test Set, Test No. 41	Lamp Driver Card, G.M. System, Test No. 74, Part No. 1618

Section VI. SUMMARY

This analysis effort has resulted in estimates of bias error, accuracy error, stimuli setting error and overall standard deviation of measurement (ETG precision) for 13 different combinations of parameters and scales that the test set is capable of measuring. The sample statistics are, of course, subject to sampling errors, hidden or undetected effects and to human errors in the test program and the computations. The test program could be improved and the statistical computations expanded for a more complete and detailed analysis which may reduce some of the inherent errors in the test program and computations. However, the results found from this sample of data give good indications of the following conditions of the ETG-3 test sets.

a) There is significant bias error (ϵ_b) on all parameters and scales tested.

b) There is significant accuracy error (ϵ_a) on 11 out of 13 parameters and scales. (10 Vdc, 250 Vdc, and 1 MHz frequency accepted).

c) There is significant stimulus setting error (ϵ_s) on 10 out of 13 parameters and scales. (10 Vdc, 1 MHz frequency, 10 VRMS, 50 Hz ac excepted).

d) A significant effect of the main factor test duration (A_i) was detected on ϵ_b two out of 13 parameters and on s^2 two out of 13 parameters.

e) A significant effect of the main factor delay time (B_j) was detected on ϵ_b nine out of 13 parameters, on ϵ_a three out of 13, on ϵ_s seven out of 13, and on s^2 two out of 13.

f) A significant effect of the main factor machines (C_k) was detected on ϵ_b 12 out of 13, on ϵ_a ten out of 13, on ϵ_s 13 out of 13, and on s^2 11 out of 13.

g) The ETG is meeting the MIS 6000 specification for precision on only four parameters out of 13 considered or about 30 percent.

h) On the basis of the weapon system tolerances used in the sample comparisons, the ETG will not be able to hold the specified "across the board" 1 out of 100 probabilities for single checkouts except where the system tolerances are broad resulting in a high accuracy ratio. Multiple sequential test probabilities will be worse.

1) The present situation of incompatibility is between

   1) across-the-board 1 out of 100 probabilities

   2) MIS 6000 specified LCSS precision

   3) unreasonably close tolerances for weapon system UUT's

   4) English Language Test Program requirements which must be resolved in a comprehensive manner.

The pulse train is not included in the parameters as no suitable specified capability for the pulse train was found and also due to the excessive percentage of the total dropouts (60 percent) found on the pulse train portion of the test program.

The test set functional dependence cannot be commented on even though it was considered to have had only 56 command dropouts out of 24,240 commands; future tests could possibly obtain an estimate of functional dependence in light of the presently obtained estimates of precision. That is, the factor of 2 variation from an expected value used in this analysis could now be tightened up to obtain a better estimate of functional dependence. This would probably lower the variability and functional dependence.

A more complete test program is considered highly desirable. It should be performed on a continuing basis by the LCSS Project Office or the prime contractor. Tests should be conducted in the field to obtain estimates of the effects of other factors such as temperature, humidity, pressure, dust, etc., on the ETG-3 performance. Such a program to be properly executed would require a considerable level of effort; but, it would be a notable achievement and contribution in the area of evaluation of the test capability of complex test equipment.

REFERENCES

1. Multisystem Test Equipment System Description Report, 1 and 2, RCA Aerospace Systems Division, Burlington, Massachusetts, Report No. ATE-L-45, 15 October 1966.

2. Mirick, H. L., "A Statistical Approach to Test Equipment Reliability," Journal of the Electronics Division of the American Society for Quality Control, November 1962.

3. Grubbs, F. E., and Coon, H. J., "On Setting Test Limits Relative to Specification Limits," Industrial Quality Control, 10 March 1954, p. 15.

4. Duncan, A. J., Quality Control and Industrial Statistics, Richard E. Irwin, Inc., Homewood, Illinois, 1959.

5. Moon, W. D., "Periodic Checkout and Associated Errors," IEEE Transactions on Aerospace, II, No. 2, April 1964.

6. Moon, W. D., "Predicting System Checkout Error," Electro-Technology, January 1964, p. 46.

7. Bartee, E. M., Engineering Experimental Design Fundamentals, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1968.

8. Hicks, C. R., Fundamental Concepts in the Design of Experiments, Holt, Rinehart, and Winston, New York, N. Y., 1964.

ESTIMATION OF LAUNCH VEHICLE PARAMETERS FOR THE GIVEN MODEL

$$y = \theta_1 e^{\theta_2 t} \sin(\theta_3 t)$$

John W. Howerton and D. Ray Campbell
Redstone Arsenal, Alabama

SUMMARY

This paper presents analysis techniques of a launch vehicle of an antitank missile system. Since the development of a mathematical model from an analysis of components subsystems' responses must be checked against overall performance, it was assumed that the vehicle responded as a second order differential equation. The solution of this equation is fitted to the experimental data.

The parameters θ_1, θ_2, θ_3 are estimated for the model $y = \theta_1 e^{\theta_2} \sin(\theta_3 t)$ and given data points $\left(T_h, Y_h\right)$, $h = 1, 2 \ldots N$. Several techniques of estimation are used. The following methods are included:

(1) Prony's Exponential Approximation

(2) Least Squares Polynomial — Taylor Series

(3) Differential Correction

(3) Gradient-Descent

(5) Modified Newton-Gauss.

A comparison of the techniques is presented and a "best" method of estimation is selected.

This article has been reproduced photographically from the authors' manuscript.

Preceding page blank

BACKGROUND

During the investigation of a particular antitank missile system, it appeared that launcher motion may have an effect on the missile trajectory. This phenomenon results from the fact that the missile is command guided and the tracker is mounted rigidly on the launcher. The program requires that a method be developed for estimating the inertial restoring, and damping parameters of a vehicle from measured motion of the vehicle during the launch phase of the flight. It will then be possible to determine more readily if the vehicle meets the desired performance requirements and to trace any degradation which may occur after a number of hours of use in the field.

The mathematical model utilized for this investigation is a damped sine wave which comes from the solution of a second order differential equation. For the purpose of this investigation the mathematical model has been assumed to be correct and is of the form: $y = \theta_1 e^{\theta_2 t} \sin(\theta_3 t)$.

Best Available Copy

I. INTRODUCTION

If a scatter diagram in the x, y plane indicates that a straight line will not fit a set of points satisfactorily because of the nonlinearity of the relationship, it may be feasible to fit a simple curve that will yield a satisfactory fit. Since an investigator always strives to explain relationships as simply as possible, with the restriction that his explanation be consistent with previous knowledge, he will prefer to use a simple type of curve. It follows, therefore, that the type of curve to use will depend largely on the amount of theoretical information one has concerning the relationship and, also, convenience.

In the problem under study, it was assumed that a second order differential equation described the motion of the vehicle. The data recorded $D = \left\{ \left(t_h, y_h\right) | h = 1, 2, \ldots N \right\}$, which was the displacement, y_h, from equilibrium position at time, t_h. Thus, the solution to the differential equation,

$$Y = \theta_1 e^{\theta_2 t} \sin \theta_3 t \tag{1}$$

is to be fitted to D by determining

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}$$

so as to minimize the residuals.

An attempt is made to solve directly the least squares problem resulting from (1) and D. Let

$$Q(\theta) = \sum_{h=1}^{N} \left(y_h - \theta_1 e^{\theta_2 t_h} \sin \theta_3 t_h \right)^2 . \tag{2}$$

By taking partials of $Q(\theta)$ with respect to each of the parameters and setting the results equal to zero, the following nonlinear system results:

$$\frac{\partial Q(\theta)}{\partial \theta_1} = \theta_1 \sum_{h=1}^{N} e^{2\theta_2 t_h} \sin^2\left(\theta_3 t_h\right) - \sum_{h=1}^{N} y_h e^{\theta_2 t_h} \sin\left(\theta_3 t_h\right) = 0$$

$$\frac{\partial Q(\theta)}{\partial \theta_2} = \theta_1 \sum_{h=1}^{N} t_h e^{2\theta_2 t_h} \sin^2\left(\theta_3 t_h\right) + \sum_{h=1}^{N} y_h t_h e^{\theta_2 t_h} \sin\left(\theta_3 t_h\right) = 0$$

$$\frac{\partial Q(\theta)}{\partial \theta_3} = \theta_1 \sum_{h=1}^{N} t_h e^{2\theta_2 t_h} \sin\left(\theta_3 t_h\right) \cos\left(\theta_3 t_h\right) - \sum_{h=1}^{N} t_h y_h e^{\theta_2 t_h} \cos\left(\theta_3 t_h\right) = 0. \tag{3}$$

Although the above system can be reduced to a two-dimensional system, the two-dimensional system indicated that the direct approach is not feasible.

To take a slightly different approach, the Laplace transform of $Q(\theta)$ was taken. The partial derivatives of the resulting expression, $L\{Q[(\theta), t]\}$, with respect to each of the three parameters were taken and the results set equal to zero. The exponential and the sine function are suppressed in the resulting system, but the rational functions involved proved unmanageable and this approach was also abandoned.

Also attempts at simplification by taking the logarithm of (1) and (2) proved futile.

Since the direct approach to the problem would not yield a solution, several indirect methods were applied. These methods were:

(1) Proney's Exponential Approximation

(2) Taylor Series — Least Squares Polynomial

(3) Modified Newton-Gauss

(4) Steepest Descent — Method of Optimum Gradients

(5) Differential Correction.

Based on machine time, number of iterations and minimum of residuals, some methods gave better results than others. In Section III, advantages and disadvantages of each method are given and a "best" method is chosen.

II. DISCUSSION OF METHODS

In this section a detailed discussion of the four indirect methods is given.

A. Prony's Exponential Approximation

From the set D, four equally spaced (with respect to time) points, (t_1, y_1), (t_2, y_2), (t_3, y_3), (t_4, y_4) are chosen. Let

$$A_1 = \frac{\theta_1}{2i} \qquad A_2 = \frac{-\theta_1}{2i}$$

$$a_1 = \theta_2 + \theta_3 i \qquad a_2 = \theta_2 - \theta_3 i.$$

Then $\theta_1 e^{\theta_2 t} \sin \theta_3 t = A_1 e^{a_1 t} + A_2 e^{a_2 t}$. Prony's theory states that e^{a_1} and e^{a_2} satisfy the equation,

$$r^2 + C_1 r + C_2 = 0 \tag{4}$$

when (4) is the characteristic equation of the assumed difference equations

$$C_2y_1 + C_1y_2 + y_3 = 0$$

$$C_2y_2 + C_1y_3 + y_4 = 0 \quad . \tag{5}$$

Thus, from (4) and (5), a_1 and a_2 are determined. Next, the system

$$y_1 = A_1e^{a_1t_1} + A_2e^{a_2t_1}$$

$$y_2 = A_1e^{a_1t_2} + A_2e^{a_2t_2} \tag{6}$$

is solved to find A_1 and A_2. Since θ_1, θ_2, and θ_3 are given in terms of a_1, a_2, A_1, and A_2, θ_1, θ_2, and θ_3 can be found.

b. Taylor Series — Least Squares Polynomial

Motivated by the analysis of the scatter diagram of the data, a third degree polynomial $P(t) = a_0 + a_1t + a_2t^2 + a_3t^3$ was least squares fitted to the data. Thus, a_0, a_1, a_2, and a_3 are determined. Next, $y(t) = \theta_1e^{\theta_2t}\sin\theta_3t$ is expanded in a Taylor series about $t = 0$ and the series is truncated after the first four terms. From the least squares polynomial and the truncated Taylor series, coefficients of equal powers of t are equated yielding:

$$a_0 = 0$$

$$a_1 = \theta_3\theta_1$$

$$a_2 = \theta_3\theta_2\theta_1$$

$$a_3 = \frac{1}{6}(\theta_3\theta_2^2\theta_1 + 2\theta_2^3\theta_3\theta_1 - \theta_3^3\theta_1) \quad . \tag{7}$$

Since a_0, a_1, a_2, and a_3 are known, (7) can be solved to obtain θ_1, θ_2, and θ_3:

$$\theta_2 = \frac{-a_2}{a_1}$$

$$\theta_3 = \sqrt{\frac{3\theta_2^2 - 6a_3}{a_1}}$$

$$\theta_1 = \frac{a_1}{\theta_3} \quad . \tag{8}$$

Unfortunately, considerable error may be introduced by the least squares fit of the polynomial and by the truncation of the Taylor series after four terms. To improve our estimate of θ, the following technique is applied. From the set D, the set D*CD is selected such that $(t^*, y^*) \epsilon D^*$ and $(t, y) \epsilon D$ and $t = t^* \rightarrow y^* \geq y$. The Taylor series — least squares polynomial method is applied to the set D* and we obtain

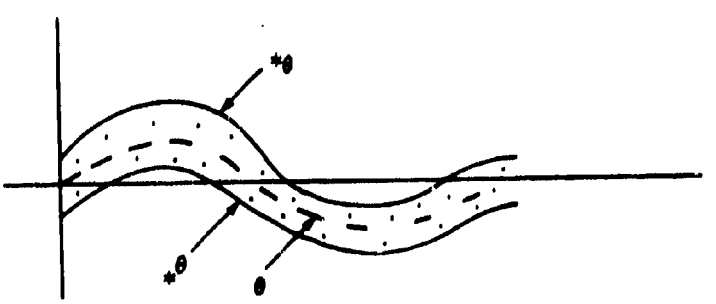
$$ {}^*\theta = \begin{pmatrix} {}^*\theta_1 \\ {}^*\theta_2 \\ {}^*\theta_3 \end{pmatrix} . $$

Now select $D_* CD$ such that $(t_*, t_*) \epsilon D_*$ and $(t, y) \epsilon D$ and $t = t_* \rightarrow y_* \geq y$. Again apply the Taylor series — least squares polynomial method and obtain

$$ {}_*\theta = \begin{pmatrix} {}_*\theta_1 \\ {}_*\theta_2 \\ {}_*\theta_3 \end{pmatrix} . $$

Actually, what we want to do now is to "scan the interval between ${}^*\theta$ and ${}_*\theta$" to get the best estimate for θ.

Ideal Picture



Let

$$ {}_0\theta_i = \min\left\{{}_*\theta_i, {}^*\theta_i\right\} \quad i = 1, 2 $$

$$ {}_0\theta_3 = \max\left\{{}_*\theta_3, {}^*\theta_3\right\} . $$

Note: These choices are made to help insure that we start with "lower" curve

Let

$$ \Delta_i = \frac{{}_0\theta_i - {}^*\theta_i}{K} . $$

K is some preassigned constant that determines step length.

Let

$$ \Delta = \begin{pmatrix} \Delta_1 \\ \Delta_2 \\ -\Delta_3 \end{pmatrix} . $$

For notational purpose, let

$$ \theta_i = {}_j\theta_i , \qquad i = 1, 2, 3 , \qquad j = 0, 1, 2, \ldots K $$

$$ {}_j\theta = \begin{pmatrix} {}_j\theta_1 \\ {}_j\theta_2 \\ {}_j\theta_3 \end{pmatrix} \quad j = 0, 1, 2, \ldots, K. $$

Note: j is the number of the step in the "scanning" procedure.

Let

$$ Q\left({}_j\theta\right) = \sum_{h=1}^{N} \left[y_h - {}_j\theta_1 e^{{}_j\theta_2 t_h} \sin\left({}_j\theta_3\right) \right]^2 . $$

Now $Q\left({}_j\theta\right)$ is computed, beginning with $j = 0$

Let ${}_{j+1}\theta = {}_j\theta + \Delta$ and $Q\left({}_{j+1}\theta\right)$ is computed. If $Q\left({}_{j+1}\theta\right) > Q\left({}_j\theta\right)$, set ${}_{j+1}\theta = {}_j\theta + \Delta\left(\frac{1}{2}\right)^m$ beginning with $m = 1$.

Again we compute $Q\left({}_{j+1}\theta\right)$. If $Q\left({}_{j+1}\theta\right) > Q\left({}_j\theta\right)$, we put $m = 2$ and repeat the above (if necessary for $m = 3, 4, 5$). If $Q\left({}_{j+1}\theta\right) < Q(\theta)$ for $m = 1, 2, 3, 4$, we set ${}_{j+1}\theta = \theta_j + \Delta$ and repeat the original procedure. Then we use $m = 5$; even though we may not have $Q\left({}_{j+1}\theta\right) < Q\left({}_j\theta\right)$, we put ${}_{j+1}\theta = {}_j\theta + \Delta$ and repeat the original procedure.

It is desirable that θ and the residuals be printed out in each step of the search so as to gain some insight of the relationship between θ and the residuals.

C. Modified Newton-Gauss

In a complete description of the first iteration, we are trying to solve

$$Q(\theta) = \sum_{h=1}^{N} \left(y_h - \theta_1 e^{\theta_2 t} \sin \theta_3 t\right)^2 = 0$$

for

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} .$$

To this end, an initial estimate

$${}_0\theta = \begin{pmatrix} {}_0\theta_1 \\ {}_0\theta_2 \\ {}_0\theta_3 \end{pmatrix}$$

is made. Next $\theta_1 e^{\theta_2 t} \sin(\theta_3 t)$ is expanded in a Taylor series about ${}_0\theta$ and only the first two terms of the series are used. The truncated Taylor series replaces $\theta_1 e^{\theta_2 t} \sin \theta_3 t$ in $Q(\theta)$. Then, the partial derivatives with respect to θ_1, θ_2, and θ_3 of $Q(\theta)$ are computed and set equal to zero. This results in a system of equations that is linear in $\left(\theta_i - {}_0\theta_i\right) i = 1, 2, 3$. Thus, if the coefficient matrix of the system is invertible, we can solve for

$$\Delta\theta = \begin{pmatrix} \theta_1 - {}_0\theta_1 \\ \theta_2 - {}_0\theta_2 \\ \theta_3 - {}_0\theta_3 \end{pmatrix}.$$

There are cases in which the coefficient matrix may not be invertible. This is certainly true if the parameter surface is flat. Thus, a singular coefficient matrix, in this case, is indicative of a non-unique solution. In cases in which a unique solution does exist, the addition of second partials in computing the correction, $\Delta\theta$ helps to insure the non-singularity of the coefficient matrix.

Effectively, we have a new estimate for θ. (This is where the differential correction method, alias Newton's Method, begins a new iteration.)

Next, we optimize the magnitude of the correction $\Delta\theta$. The expression ${}_0\theta + v\Delta\theta \quad 0 \leq v \leq 1$ is considered.

The term Q is evaluated at ${}_0\theta$, ${}_0\theta + \frac{1}{2}\Delta\theta$, and ${}_0\theta + \Delta\theta$. A parabola is then passed through these three points and the minimum of the parabola is calculated in terms of v. Then $Q({}_0\theta + v\Delta\theta)$ is computed. The θ associated with the $\min\left\{Q({}_0\theta), Q\left({}_0\theta + \frac{1}{2}\Delta\theta\right), Q({}_0\theta + \Delta\theta), Q({}_0\theta + v\Delta\theta)\right\}$ is chosen and called ${}_1\theta$. If $Q({}_1\theta) < Q({}_0\theta)$ the initial estimate of θ, ${}_0\theta$ is replaced by ${}_1\theta$ and the

entire procedure is repeated. If this is not the case, the domain of v is diminished $\left(\text{usually by } \frac{1}{2}\right)$ and the optimization of the magnitude of the correction $\Delta\theta$ is repeated the required number of times to produce $Q({}_1\theta) < Q({}_0\theta)$, or the domain of v is sufficiently small and we terminate the procedure.

NOTE: The reason that the check of the $\min \left\{ Q({}_0\theta), Q\left({}_0\theta + \frac{1}{2}\Delta\theta\right), Q({}_0\theta + \Delta\theta), Q({}_0\theta + v\Delta\theta) \right\}$ is considered in that the minimum of the parabola does not necessarily occur at that value of θ that will produce the minimum Q.

D. The Method of Steepest Descent — Optimum Gradient

Again we consider the expression

$$Q(\theta) = \sum_{h=1}^{N} \left[y_h - \theta_1 e^{\theta_2 t_h} \sin\left(\theta_3 t_h\right) \right]^2 . \tag{9}$$

An initial estimate is made; call it ${}_0\theta$. The gradient, ∇Q, of Q is computed at ${}_0\theta$. Since the gradient points in the direction of maximum increase of Q, the negative of the gradient will point in the direction of greatest decrease of the function. Now the gradient is normalized by dividing each component of the gradient by the maximum of the absolute values of the components.

We next optimize the step-length in the direction of steepest descent by considering the function

$$g(\alpha) = Q\left[{}_1\theta - \alpha \nabla Q\left({}_1\theta\right) \right] ; \tag{10}$$

we find that value of α that will make Q a minimum.

Now

$$g'(\alpha) = -\nabla Q\left({}_1\theta\right) \cdot \nabla Q\left[{}_1\theta - \alpha \nabla Q\left({}_1\theta\right) \right] . \tag{11}$$

By setting $g'(\alpha) = 0 \rightarrow \nabla Q\left({}_1\theta\right)$ and $\nabla Q\ {}_1\theta - \alpha \nabla Q\ {}_1\theta$ are orthogonal to each other for the value of α that makes Q a minimum.

We now compute the function values of g beginning at g(o) and continuing [illegible] until the slope of g changes from negative to non-negative or until g(α-1), whichever comes first. Suppose we have the change of [illegible] the magnitude of Q at N. Then define

$$b = N$$

$$a = N - 1,$$

[illegible] we will pass a cubic through [illegible] and determine the value of α that produces a minimum in the cubic. As in the parabolic interpolation in the Newton-Gauss program, a check is made to see if we are decreasing the magnitude of Q.

Since the "best" α is determined, we put

$$\theta_{i+1} = \theta_i - \alpha \nabla Q\left(\theta_i\right)$$

and compute

$$Q\left(\theta_i\right) \quad \text{and} \quad Q\left(\theta_{i+1}\right);$$

when the changes in Q are negligible over four iterations, the process ceases.

F. Method of Differential Correction

$$Y = f(t, \theta_1, \theta_2, \theta_3). \tag{12}$$

We want this formula to be a good fit to the data $\left(t_h, y_h\right)(h = 1, \ldots N)$. The residuals are given by

$$\left\{\begin{array}{l} R_1 = f(t_1,\theta_1,\theta_2,\theta_3) - Y_1 \\ R_2 = f(t_2,\theta_1,\theta_2,\theta_3) - Y_2 \\ \cdot \quad \cdot \cdot \cdot \cdot \cdot \quad \cdot \\ R_h = f\left(t_n,\theta_1,\theta_2,\theta_3\right) - Y_h \end{array}\right\} \tag{13}$$

where $y_h (h = 1, 2, \ldots n)$ are the given (observed) values from the original data.

Let ${}_0\theta_1, {}_0\theta_2, {}_0\theta_3$ be an initial guess. Now we need to correct this guess by some incremental amount, say α, β, γ such that

$$\left\{\begin{array}{l} \theta_1 = {}_0\theta_1 + \alpha \\ \theta_2 = {}_0\theta_2 + \beta \\ \theta_3 = {}_0\theta_3 + \gamma \end{array}\right\} \tag{14}$$

will yield a better fit to our data.

If we substitute the values of (14) into the residuals (13) and transpose the y_h, we have the following:

$$R_h + y_h = f\left(t_i, {}_0\theta_1 + \alpha, {}_0\theta_2 + \beta, {}_0\theta_3 + \gamma\right). \tag{15}$$

Expanding the right-hand side by Taylor's theorem we get

$$R_h + y_h = f\left(t_h, {}_0\theta_1, {}_0\theta_2, {}_0\theta_3\right) + \alpha\left(\frac{\partial f_h}{\partial \theta_1}\right)_0 + \beta\left(\frac{\partial f_h}{\partial \theta_2}\right)_0 + \gamma\left(\frac{\partial f_h}{\partial \theta_3}\right)_0$$

$$+ \text{higher order terms in } \alpha, \beta, \gamma \tag{16}$$

where $\left(\frac{\partial f}{\partial y}\right)_0 \equiv$ the value of the partial derivative $\frac{\partial f}{\partial y}$ at

$$t = t_h, \quad \theta_1 = {}_0\theta_1, \quad \theta_2 = {}_0\theta_2, \text{ and } \theta_3 = {}_0\theta_3 . \tag{17}$$

Our first approximation is obtained from

$$Y' = f(t, {}_0\theta_1, {}_0\theta_2, {}_0\theta_3)$$

so that we have

$$f\left(t_h, \theta_1, \theta_2, \theta_3\right) = Y'_h \quad . \tag{18}$$

This can also be put into equation (16). Now let

$$r_h = Y'_h - Y_h \quad . \tag{19}$$

Ignoring the higher order terms the residuals now have the form

$$r_h = \alpha\left(\frac{\partial f_h}{\partial \theta_1}\right)_0 + \beta\left(\frac{\partial f_h}{\partial \theta_2}\right)_0 + \gamma\left(\frac{\partial f_h}{\partial \theta_3}\right)_0 + r_h \tag{20}$$

which are linear in α, β, γ. Therefore, we may determine the corrections by the method of least squares.

III. RESULTS AND CONCLUSIONS

These five methods of estimating θ were used on data taken from actual test firings in which we had four test modes (a standard and three modifications). Plots of typical data are shown in Figures 1 through 4. The results of all five methods are shown in Tables I through IV.

Some advantages and disadvantages of each method are given in Table V. It turns out that the "best" method to use depends on the tools that one has on hand. For example, if one has to estimate the parameters by use of only paper and pencil, he naturally would chose Prony's method. If one had access to a small computer, he might chose the Taylor-least squares approach. Of the five methods discussed in this paper, the method of steepest descent, and the method of modified Newton-Gauss are the most accurate in terms of the smallest residuals.

TABLE I. STANDARD DATA

	θ_1	θ_2	θ_3	Residual
Prony's	6.9	-1.14	7.3	374
Taylor's modified	6.4	-2.8	7.6	94
Differential correction	Did not converge			
Steepest descent	6.5	-2.5	7.9	91
Modified Newton-Gauss	6.5	-2.5	7.9	91

TABLE II. MOD 1 DATA

	θ_1	θ_2	θ_3	Residual
Prony's	2.64	-1.24	10.3	171
Taylor's modified	2.4	-3.7	10.0	88
Differential correction	2.07	-1.85	12.3	68
Steepest descent	2.07	-1.85	12.3	68
Modified Newton-Gauss	2.09	-1.84	12.3	68

TABLE III. MOD 2 DATA

	θ_1	θ_2	θ_3	Residual
Prony's	5.3	-0.7	10.46	407
Taylor's modified	4.14	-2.5	8.92	177
Differential correction	Did not converge			
Steepest descent	4.7	-1.5	9.4	102
Modified Newton-Gauss	4.8	-1.5	9.4	101

TABLE IV. MOD 3 DATA

	θ_1	θ_2	θ_3	Residual
Prony's	8.6	-0.6	7.2	820
Taylor's modified	8.0	-2.1	7.7	328
Differential correction	Did not converge			
Steepest descent	8.5	-1.4	7.3	153
Modified Newton-Gauss	Relative min due to guess 1609			

TABLE V. ADVANTAGES AND DISADVANTAGES

	Easily Hand-fit	Easily Adapted to Small Computer	Converges to Relative Minimum	Relative Accuracy	Sensitive to Guess
Prony's	yes	yes	N/A	4	N/A
Taylor's - least squares	no	yes	yes	3	N/A
Differential-correction	no	no	no	2	yes
Steepest descent	no	no	yes	1	no
Modified Newton-Gauss	no	no	yes	1	no

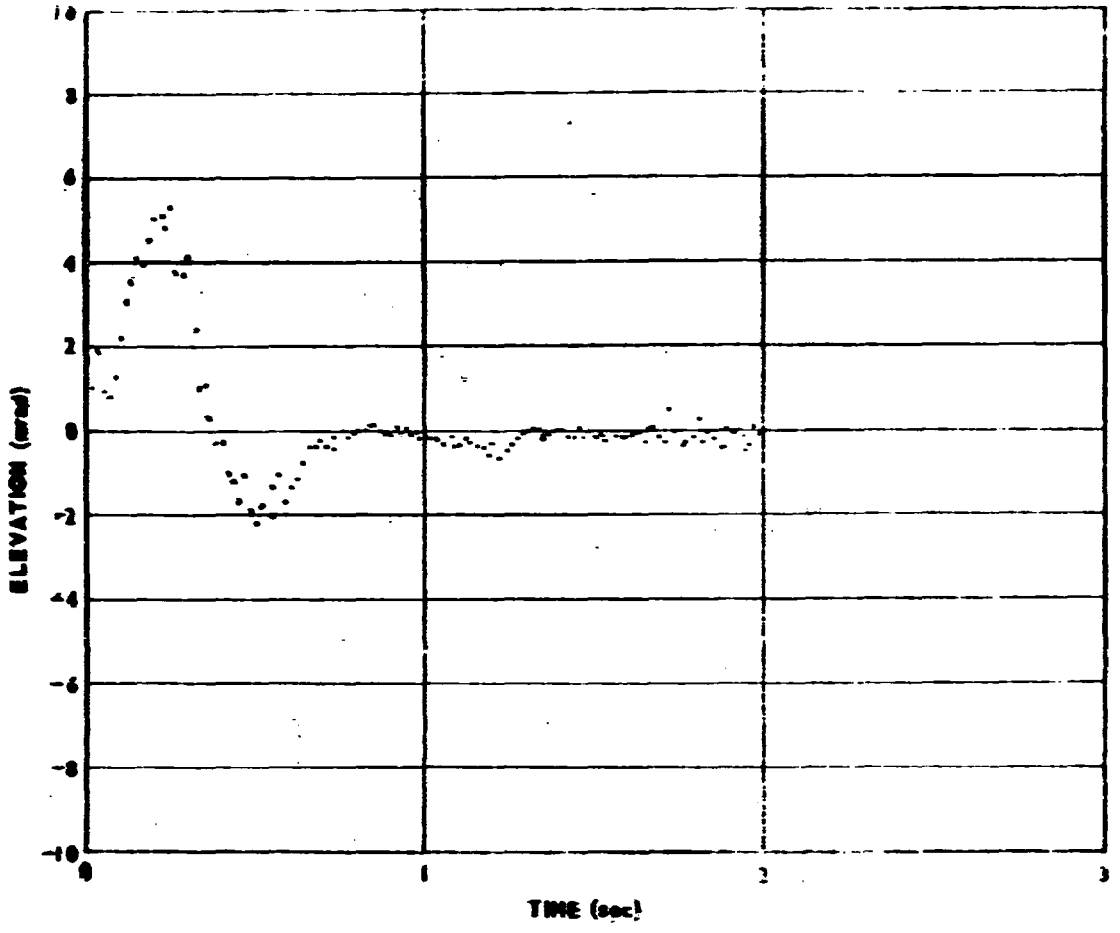


FIGURE 1. STANDARD VEHICLE

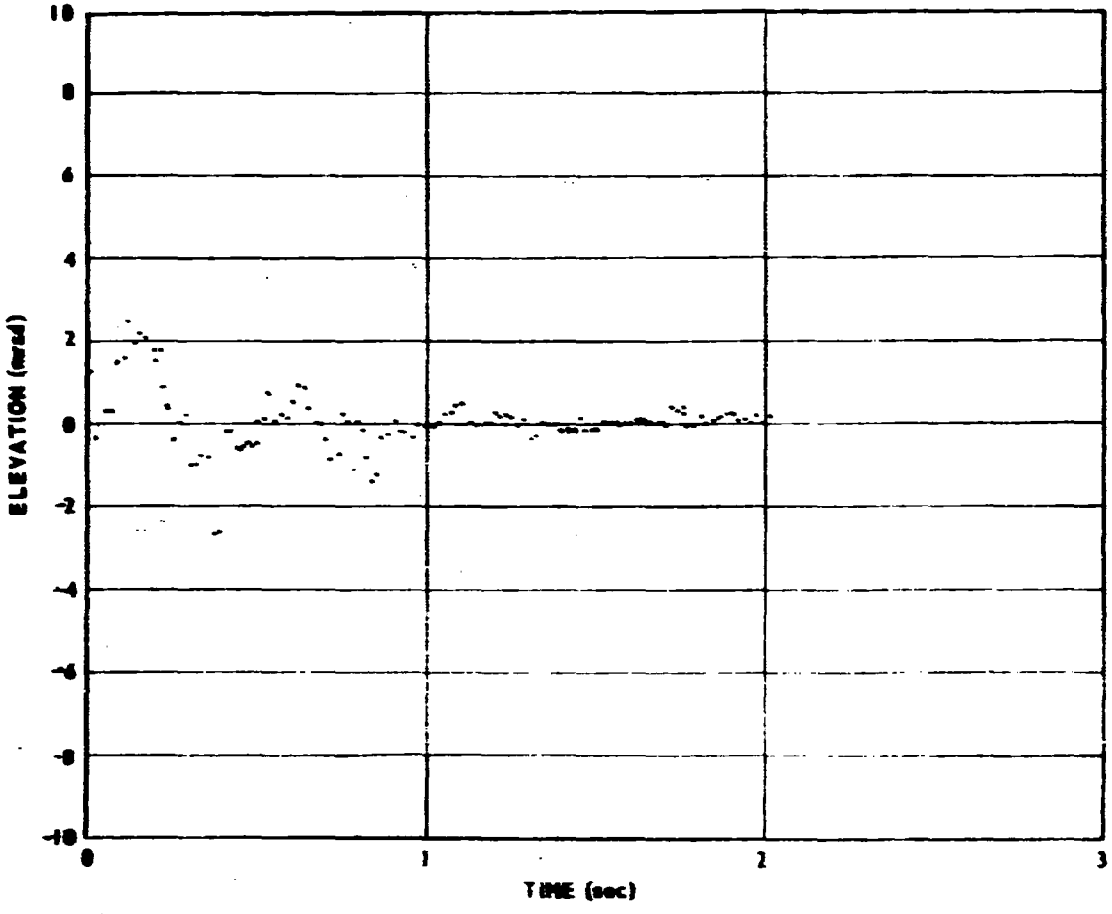


FIGURE 2. MODIFICATION 1

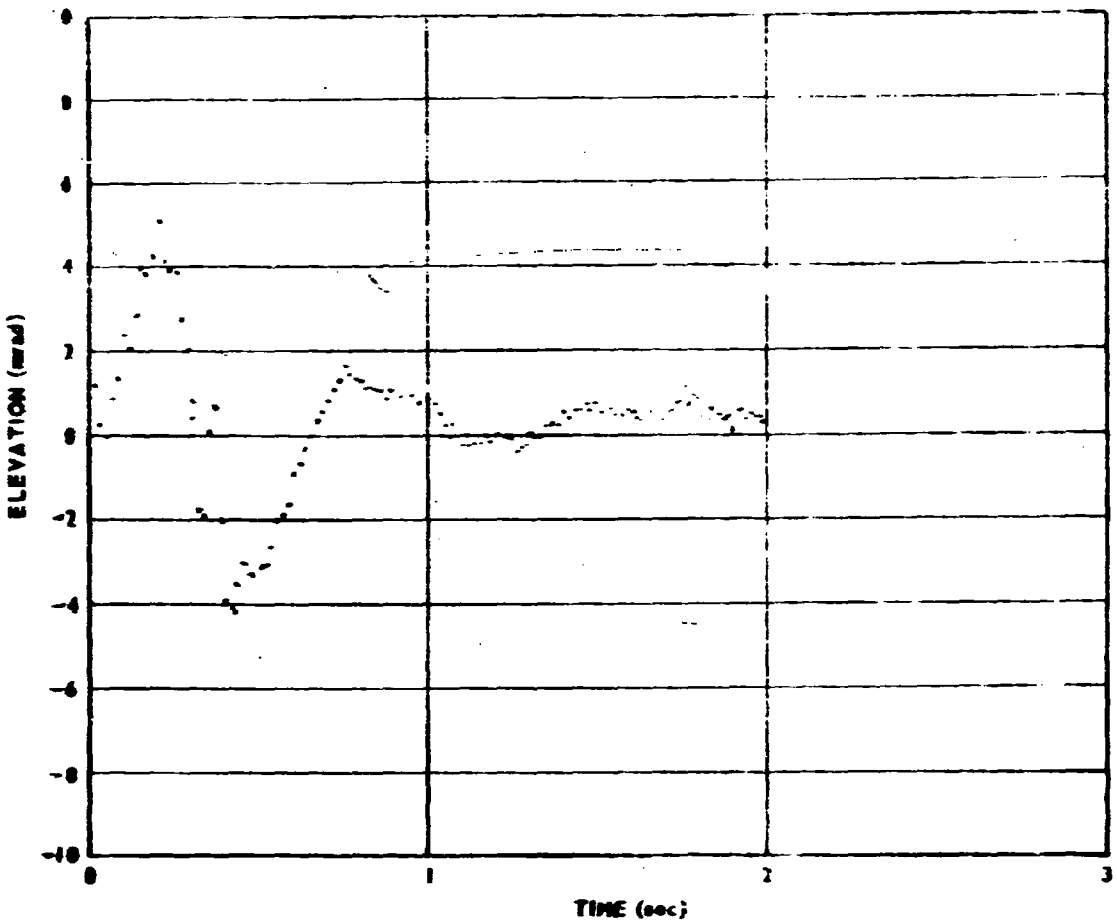


FIGURE 3. MODIFICATION 2

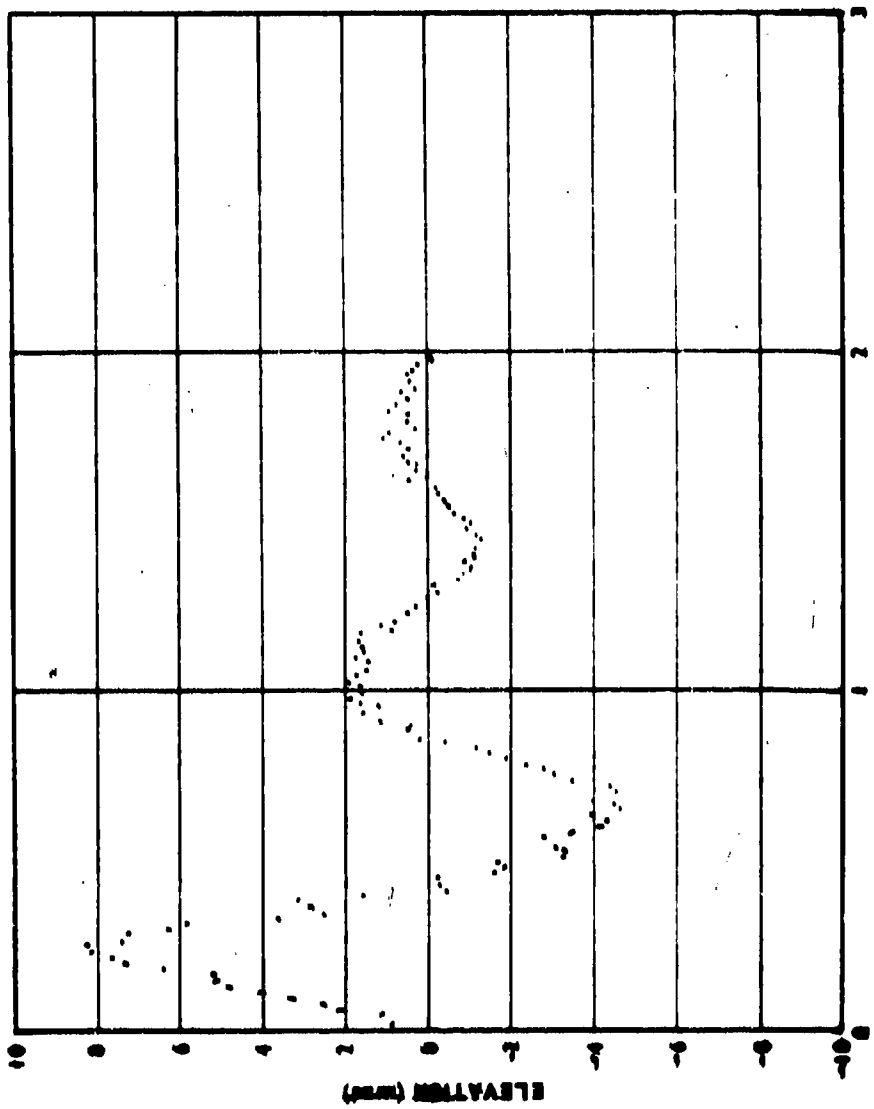


FIGURE 4. MODIFICATION 3

REFERENCES

1. Nielsen, K. L., Methods of Numerical Analysis, The MacMillian Company, New York, 1956.

2. Bendat, J. S., Piersol, A. G., Measurement and Analysis of Random Data, John Wiley & Sons, Inc., New York.

3. Hartley, H. O., Technometrics, The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares, Vol. 7, No. 2, May 1961.

4. Kelly, L. G., Handbook of Numerical Methods and Applications.

A METHOD OF IMPROVING THE ESTIMATION OF A VARIANCE

John Gurland
University of Wisconsin

and

J. S. Mehta
Temple University

1: Introduction.

Consider a situation where an experimenter has an unbiased estimator s_1^2 of the population variance σ_1^2 from a normal distribution. Let us further suppose that another independent unbiased estimator s_2^2 of the population variance is also available. The two estimators may have been based on samples taken at different times and places. Because of the circumstances it may be known that s_1^2 estimates σ_1^2 while s_2^2 may be estimating σ_2^2 ($\gtrless \sigma_1^2$). This shift in the variance, if any, may have taken place because of the time lapse between obtaining independent samples or it could be due to the shift in places. It is obvious that if $\sigma_1^2=\sigma_2^2$ then the two estimators can be pooled to obtain a "better estimator" of σ_1^2. On the other hand, if $\sigma_2^2\neq\sigma_1^2$ then one may estimate σ_1^2 by s_1^2 alone. A preliminary test of the hypothesis H_o: $\sigma_2^2=\sigma_1^2$ can be carried out by utilizing the $F(=s_2^2/s_1^2)$ statistic and if the hypothesis is rejected then one uses s_1^2 to estimate σ_1^2 otherwise a pooled estimator of σ_1^2 is obtained by pooling s_1^2 and s_2^2 appropriately. The estimator designated here as U has been obtained by following this approach, which is similar to the approach used by Bancroft [1].

Another method of estimating σ_1^2 is to use weights which are continuous functions of F. The estimator S described below is constructed in this manner and turns out to be more effective than the estimator U.

Let $x_{11}, x_{12}, \ldots, x_{1N}$ and $x_{21}, x_{22}, \ldots, x_{2N}$ be independent samples from normal populations with unknown variances σ_1^2 and σ_2^2 respectively, and let $k = \sigma_2^2/\sigma_1^2$. It is required to estimate σ_1^2. Define as usual

$$s_1^2 = \frac{1}{N-1}\sum_{j=1}^{N}(x_{1j}-\bar{x}_1)^2 \quad ; \quad s_2^2 = \frac{1}{N-1}\sum_{j=1}^{N}(x_{2j}-\bar{x}_2)^2$$

We consider an estimator T of weighted sums of s_1^2 and $\frac{s_1^2 + s_2^2}{2}$ of the form

$$T = \psi s_1^2 + (1-\psi) \frac{s_1^2 + s_2^2}{2} \qquad (1)$$

where ψ is some function of F. The reason for considering T as an estimator of σ_1^2 is that s_1^2 is an estimator based only on the first sample and

$$\frac{s_1^2 + s_2^2}{2}$$

is an estimator based on combining the two samples when $\sigma_1^2 = \sigma_2^2$. As an estimator of σ_1^2 we consider a subclass S of T which reduces to the form

$$S = s_1^2 \left[\frac{\alpha_1 - \alpha_2 F + \alpha_3 F^2}{\beta_1 - \beta_2 F + \beta_3 F^2} \right] \qquad (2)$$

where the constants of α_i and β_i are given below. This estimator has also been considered elsewhere (Mehta and Gurland [6] but we include it here for comparison.

The estimator U mentioned above is also a subclass of T with the weight function ψ given by

$$\psi = \begin{cases} 0 & \text{for } 1/F_o < F < F_o \\ 1 & \text{otherwise} \end{cases} \qquad (3)$$

The constant F_o is determined by the level of significance of the preliminary test for equality of variances.

The behaviour of S and U will be investigated in regard to expected mean square and relative bias.

2. The estimator U

The estimator U is of the form T with the weight function ψ defined by (3) above. It is based on the random outcome of a preliminary test of whether $\sigma_1^2 = \sigma_2^2$. If the preliminary test rejects equality of variances then s_1^2 is employed as the estimator; but if it does not reject equality of variances then the average of s_1^2 and s_2^2 is used as the estimator of σ_1^2. This estimator is similar to one used by Bancroft [1] except that we use a two-sided test for equality of variances whereas he uses a one-sided test.

First we consider the expected mean square of U and compare it with that of s_1^2 which utilizes only the first sample. Since s_1^2 is unbiased we define the efficiency of U as

$$\text{Eff } U = \frac{\text{Var } s_1^2}{E(U-\sigma_1^2)^2} \tag{4}$$

Subsequently we shall also consider the relative bias of U given by

$$\frac{E(U) - \sigma_1^2}{\sigma_1^2} \tag{5}$$

In order to obtain expressions for the efficiency and bias of U we need to evaluate the first two moments of U.

It can be shown by straight-forward integration that

$$E(U) = \sigma_1^2 \left[1+\frac{1}{2}\left(I_{x_0}\left(\frac{N+1}{2},\frac{N-1}{2}\right)-I_{x_0}\left(\frac{N+1}{2},\frac{N-1}{2}\right)\right)+\frac{1}{2}\left(I_{x_0}\left(\frac{N-1}{2},\frac{N+1}{2}\right)-I_{y_0}\left(\frac{N-1}{2},\frac{N-1}{2}\right)\right)\right] \quad (21)$$

$$x_0 = \frac{kF_0}{1+kF_0} \qquad y_0 = \frac{1}{1+kF_0}$$

$$B(a,b)\, I_{t_0}(a,b) = \int_0^{t_0} t^{a-1}\,(1-t)^{b-1}\,dt$$

Similarly the variance of U is given by

$$E(U^2) = \frac{\sigma_1^4}{4(N-1)^2}\left[(N^2-1)\left(I_{x_0}\left(\frac{N+3}{2},\frac{N-1}{2}\right)-I_{y_0}\left(\frac{N+3}{2},\frac{N-1}{2}\right)\right)+\right.$$

$$2(N-1)^2 k\left(I_{x_0}\left(\frac{N+1}{2},\frac{N+1}{2}\right)-I_{y_0}\left(\frac{N+1}{2},\frac{N+1}{2}\right)\right)+$$

$$\left.(N^2-1)\,k^2\left(I_{x_0}\left(\frac{N-1}{2},\frac{N+3}{2}\right)-I_{y_0}\left(\frac{N-1}{2},\frac{N+3}{2}\right)\right)\right]+$$

$$\sigma_1^4\left[\frac{N+1}{N-1}\left(I_{y_0}\left(\frac{N+3}{2},\frac{N-1}{2}\right)-I_{x_0}\left(\frac{N-1}{2},\frac{N+1}{2}\right)+1\right)\right]$$

where x_0, y_0 and $I_{t_0}(a,b)$ have been defined in (22). Thus the efficiency of U can be written as follows

$$\text{Eff } U = \frac{\text{Var}\,(s_1^2)}{E(U^2)-[E(U)]^2+[\text{Bias } U]^2}$$

where

$$\text{Bias } U = E(U) - \sigma_1^2$$

and $E(U)$, $E(U^2)$ are given above.

3. Some computed values of Efficiency and Relative Bias of U.

In Tables 1-5 are given calculated values of the efficiency of U_2 for sample sizes 3,5,7,9,11, for the range $0.1 \leq k \leq 10.0$, and for values of the constant F_0 which correspond to 1%,5%,10%,20% and 50% levels of significance. Examination of these tables reveals that there is a gain of efficiency for some values of k and a loss for other values of k. Furthermore the extent of these gains or losses depends on sample size. For $k \leq 1$ there is generally a gain in efficiency but the magnitude of this gain decreases with increase in sample size. As a matter of fact for small values of k, e.g. around 0.1, there is even a loss of efficiency as N becomes larger.

As far as values of $k > 1$ are concerned there is, without exception, a general loss of efficiency as manifested for all the sample sizes considered. The relationship of this loss with the values of k and N is more complicated than for the case $k \leq 1$ considered above. For some values of $k > 1$, e.g. $k = 2$, the loss in efficiency becomes more pronounced as N increases while for other values of $k > 1$, e.g. $k = 10$, it becomes less pronounced.

For all values of k and N considered the relation of efficiency to the value of the constant F_0 follows a definite pattern. Whatever be the sample size, the efficiency for values of $k > 1$ increases with decrease of F_0 (or equivalently with increase of level of significance of the preliminary test). On the other hand for $k \leq 1$ the trend is reversed, that is to say, the efficiency decreases with increase of the level of significance of the preliminary test.

Best Available Copy

On the basis of efficiency only, one would be tempted to select an estimator of this class with a low level of significance of the preliminary test involved. However, when this level is low the relative bias is high. In Table 10 is presented the relative bias of U for sample sizes 5,6,7,9 and 11. The value of F_0, for each sample size, corresponds to a 20% level of significance of the preliminary test of the hypothesis H_0: $k = 1$. Other levels of significance for the preliminary test are possible, of course, and as a matter of fact in Tables 1-5 we have already discussed the efficiency of U for values of F_0 corresponding to levels 1%,5%,10%,20% and 50%. Now for levels greater than 20% the relative bias will be smaller but at the same time the gain in efficiency will also be smaller. On the other hand for levels less than 20% the efficiency will be higher but the relative bias will also be higher. At this particular level, namely, 20%, the relative bias is reasonably well controlled for $0.1 \leq k \leq 1.0$ and at the same time there are gains of efficiency, at least in a subset of this range. On refering to Table 10 we note that the maximum relative bias in this range of k for sample size $N = 11$ is 6% and for $N = 3$ it is 14%.

4. The estimator S

In the general estimator T defined by (2) let us regard ψ for the moment as a constant, and minimize the expected mean square error of T with respect to ψ. The minimum is reached when

$$\psi = \frac{N-3-2k(N-1)+k^2(N+1)}{(N+1)-2k(N-1)+k^2(N+1)}$$

If we substitute this value for ψ in the estimator T and replace k, which is, of course, unknown, by c+dF, where c and d are arbitrary constants, we obtain the estimator S given by (2), where

$$\alpha_1 = (N-1)(1-2c) + (N+1)c^2 \qquad \beta_1 = (N+1)(1+c^2)-2(N-1)c$$

$$\alpha_2 = 2\{(N-1)d-(N+1)cd-1\} \qquad \beta_2 = 2d\{(N-1) + (N+1)c\}$$

$$\alpha_3 = (N+1)d^2 \qquad \beta_3 = (N+1)d^2$$

The relevant underlying details involved in obtaining the above estimator S are outlined in the paper by Mehta and Gurland [4].

Estimation of k by a simple function such as c + dF has been applied similarly in other contexts (cf.[2],[3]). The constants c and d must be appropriately chosen, and the results of certain choices will appear in Tables 6, 7, 8, 9, 11, 12 considered below.

As in the case of U we require the first two moments of S in oder to evaluate its efficiency and relative bias. For odd values of the sample size N these moments can be expressed as a finite series of integrals which can be evaluated by reduction. The precise form of these moments appears in the work by Mehta and Gurland [6] cited above.

5. Some computed values of Efficiency and Relative Bias of S

In examining the behaviour of the estimator S we employ the same criteria of efficiency and relative bias defined above as in (4), (5), for the estimator U. The behaviour of the estimator S has been considered previously in [6], but for convenience of making comparisons with the estimator U we sketch these results here briefly. In table 6 the efficiency is shown

when values of the constants c and d have been selected to emphasize the range $1 \leq k \leq 10$. Except for very small sample sizes, such as N = 3, the gain in efficiency is not substantial. In fact, a loss in efficency begins to occur for sample size 7 and for larger values of k.

In Table 7 the efficiency is shown corresponding to constants c and d which emphasize the range $0.1 \leq k \leq 1.0$. There is a considerable gain in efficiency for this range, especially for small sample size. This gain, however, is accompanied by a large relative bias.

In Table 8 the efficiency is given for estimator S where the constants c and d are chosen so that the relative bias remains numerically below 10% for the range $0.1 \leq k \leq 10.0$. Table 9 indicates the efficiency that results when constants c and d are chose to hold the relative bias numerically below 5%. In this case the gain in efficiency is slight, especially for larger sample sizes.

In Tables 11 and 12 are presented the relative bias of the estimators in the class S for which the efficiency has been discussed in Tables 8 and 9. From Table 11 it is evident that the relative bias of the estimators [illegible] there is much smaller. In fact for N = 11 the maximum relative bias [illegible] and for N = 3 it is 5.8%. The gain in efficiency, however, as indicated in Table 9 is only very slight for the range $0.1 \leq k \leq 1.0$; [illegible] slight loss of efficiency for the range $[illegible] \leq 10.0$. From the point of view of efficiency this estimator is not so attractive; however, it is of interest if one is mainly interested in controlling the bias.

It is evident from Table 11 that the maximum relative bias for the whole range $0.1 \leq k \leq 10$ and for all the sample sizes considered is 16%. It is also evident that for many of these values of k for the sample sizes considered this relative bias is very small. This control of relative bias together with the gain in efficiency discussed previously would indicate that this member of the class S merits consideration as an estimator of σ_1^2.

5. Comparison of the behaviour of estimators U and S

In comparing the behaviour of S and U it is necessary to keep in mind that the parameter k can assume values greater than or less than one. It is evident that we can find members of the class S which in most of the range $0.1 \leq k \leq 10.0$ are more efficient than members of U .. For example on comparing the estimator S in Table 8 corresponding to N = 9 and the estimator U in Table 4 corresponding to a preliminary test at a 20% level we observe that for all values of $k > 1$ the efficiencies of S are very much higher than those of U. For the range $0.1 \leq k \leq 1.0$ the efficiencies of S exceed those of U except for values of k in the subset $0.7 \leq k \leq 1.0$ in which subset the efficiencies of U are only slightly greater than those of S. The comparative behaviour of S and U for other sample sizes considered follows a similar pattern. Generally speaking therefore in the whole range $0.1 \leq k \leq 10.0$ the estimator S appears preferable as far as efficiency is concerned.

Let us now consider the relative bias of these estimators. Values of the relative bias of S and U are given in Tables 11 and 10 respectively. On comparing these biases corresponding to sample size N = 9, for example, the relative bias of S in the range $1.0 < k \leq 10.0$ is very much less than that of U , while for k in the range $0.1 \leq k \leq 1.0$ the relative bias of S remains less

than that of U [illegible] subset [illegible] $\leq k \leq 0.3$. The [illegible] behaviour [illegible] follows a similar pattern, [illegible] is much less than that of U while for values of $k \geq 1$ the relative bias of S remains less than that of U except [illegible] subset [illegible].

[illegible] U [illegible] in order to keep the relative bias below a [illegible] is chosen with a preliminary test at 20% level the results indicate that if we regard efficiency and relative bias over the whole range $0.1 \leq k \leq 10.0$ the estimator S is preferable.

[illegible] are also possible by selecting different members of these classes of estimators: however, we believe, the basis of comparison considered here is the most generous towards U . If, for example, we consider the class U with a level of 50% for the preliminary test, its relative bias is improved but the efficiencies, with very slight exceptions, are all less than those of S. On the other hand if we consider the class U with a preliminary test at a low level, for example 1% the efficiencies for $k < 1$ generally exceed those of S: however, the disadvantages of such a U would be overwhelming because its relative bias is prohibitively large and its loss of efficiency for $k > 1$ is terrible.

Other estimators S besides those given in Table 6 might also have been considered for the comparison. In Tables 6 and 7, for example, estimators are presented which emphasize the ranges $1.0 < k \leq 10.0$ and $0.1 \leq k \leq 1.0$ respectively. On the other hand the estimators considered in Tables 9 and 12 control the relative bias within a maximum of 5%. For the whole range $0.1 \leq k \leq 10.0$, however, the estimator considered in Table 8 is worth recommending because the relative bias is reasonably well controlled and there are noticeable gains of efficiency.

Examples:

(1) The following example illustrates how to obtain S and U in a practical situation. For this we have drawn a sample of size 7 from N(3,1) and named it as the first sample. The unbiased estimate of $\sigma_1^2 = 1$ as given by s_1^2 from this sample is $s_1^2 = 1.653$. We draw another sample of size 7 from N(5,0.36) and designate it as the second sample. The unbiased estimate of $\sigma_2^2 = 0.36$ as given by s_2^2 is $s_2^2 = 0.448$. Consequently $F = s_2^2/s_1^2 = 0.271$ and the hypothesis that $k = 1$ is rejected at the 20% level of significance. Thus $U = 1.653$. On the other hand we obtain

$S = 1.527$ if we restrict the relative bias to be less than 10%.
$= 1.623$ if we restrict the relative bias to be less than 5%.
$= 1.011$ if we use the estimator S which emphasizes the range $0.1 \leq k \leq 1.0$.
$= 0.969$ if we use the estimator S which emphasizes the range $1.0 \leq k \leq 10.0$.

In all the cases we note that S is nearer the true value of unity than U.

(2) The example considered here differs from that of (1) in that here the value of the ratio k is greater than 1. Suppose now we have a second sample also of size 7 from N(4,4). The unbiased estimate of $\sigma_2^2 = 4$ is $s_2^2 = 4.305$. The value of $F = s_2^2/s_1^2$ is now 2.603 which is not significant at 20% level of significance and consequently the null hypothesis that $k = 1$ is not rejected. Therefore in this case the estimate $U = \frac{1.653 + 4.305}{2} = 2.979$.

On the other hand we obtain

$S = 1.660$ if we restrict the relative bias to be less than 10%.
$= 1.660$ if we restrict the relative bias to be less than 5%.
$= 1.881$ if we use the estimator S which emphasizes the range $0.1 \leq k \leq 1.0$.
$= 1.663$ if we use the estimator S which emphasizes the range $1.0 \leq k \leq 10.0$.

In all the cases we note that the estimator S is nearer the true value than U.

Table 1: Efficiency of B_2 for N=3

k	F_o=199 1%	F_o=[illegible] 5%	F_o=[illegible] 10%	F_o=9.0 20%	F_o=3.0 50%
10	0.025	0.042	[illegible]	0.171	0.705
9	0.031	0.050	[illegible]	0.181	0.711
8	0.039	0.059	[illegible]	0.194	0.725
7	0.051	0.073	[illegible]	0.213	0.738
6	0.070	0.097	[illegible]	0.241	0.757
5	0.102	0.140	[illegible]	0.284	0.749
4	0.162	0.145	[illegible]	0.357	0.777
3	0.295	0.[illegible]	[illegible]	0.[illegible]	0.[illegible]
2	0.673	0.[illegible]	[illegible]	0.791	0.[illegible]
1	1.941	1.[illegible]	1.[illegible]	1.344	1.[illegible]
0.9	2.122	1.[illegible]	1.[illegible]	1.384	1.[illegible]
0.8	2.283	1.[illegible]	1.[illegible]	1.409	1.[illegible]
0.7	2.410	1.[illegible]	1.[illegible]	1.412	1.[illegible]
0.6	2.485	2.[illegible]	1.[illegible]	1.396	1.[illegible]
0.5	2.493	2.[illegible]	1.[illegible]	1.343	1.[illegible]
0.4	2.430	1.[illegible]	1.[illegible]	1.269	1.[illegible]
0.3	2.294	1.[illegible]	1.[illegible]	1.174	1.[illegible]
0.2	2.086	1.[illegible]	1.[illegible]	1.067	0.[illegible]
0.1	1.764	1.[illegible]	1.[illegible]	0.976	0.[illegible]

Table 2: Efficiency of [illegible]

k	F_o=23.1[illegible] 1%	F_o=9.60 5%	F_o=6.[illegible] 10%	[illegible]
10	0.02[illegible]	0.0[illegible]	0.1[illegible]	[illegible]
9	0.03[illegible]	[illegible]	0.1[illegible]	[illegible]
8	0.[illegible]	[illegible]	[illegible]	[illegible]
7	0.06	0.1[illegible]	0.17[illegible]	[illegible]
6	0.09[illegible]	0.[illegible]	[illegible]	[illegible]
5	0.08[illegible]	[illegible]	0.21[illegible]	[illegible]
4	0.13[illegible]	[illegible]	[illegible]	[illegible]
3	0.24[illegible]	[illegible]	0.[illegible]	[illegible]
2	0.68[illegible]	[illegible]	[illegible]	[illegible]
1	1.91[illegible]	1.[illegible]	1.[illegible]	[illegible]
0.9	2.07	1.[illegible]	1.[illegible]	[illegible]
0.8	2.15[illegible]	1.[illegible]	1.[illegible]	[illegible]
0.7	2.21[illegible]	[illegible]	1.[illegible]	[illegible]
0.6	2.1[illegible]	[illegible]	1.[illegible]	[illegible]
0.5	2.0[illegible]	1.5[illegible]	1.[illegible]	1.[illegible]
0.4	1.8[illegible]	1.[illegible]	1.[illegible]	1.[illegible]
0.3	1.58[illegible]	1.[illegible]	1.[illegible]	1.[illegible]
0.2	1.[illegible]	1.[illegible]	0.[illegible]	1.[illegible]
0.1	0.[illegible]	0.[illegible]	0.[illegible]	0.[illegible]

108

Table 3: Efficiency of B_2 for N=7

k	F_o=11.07 1%	F_o=5.82 5%	F_o=4.28 10%	F_o=3.05 20%	F_o=1.78 50%
10	0.041	0.147	0.298	0.584	0.94[illegible]
9	0.044	0.142	0.281	0.553	0.9[illegible]
8	0.048	0.140	0.267	0.522	0.927
7	0.054	0.141	0.257	0.494	0.913
6	0.065	0.149	0.254	0.471	0.897
5	0.083	0.163	0.261	0.459	0.881
4	0.120	0.199	0.289	0.467	0.869
3	0.210	0.286	0.369	0.525	0.87[illegible]
2	0.516	0.569	0.627	0.729	0.930
1	1.898	1.645	1.460	1.257	1.046
0.9	2.039	1.724	1.506	1.277	1.048
0.8	2.102	1.736	1.502	1.268	1.04[illegible]
0.7	2.063	1.670	1.442	1.228	1.037
0.6	1.925	1.534	1.335	1.161	1.024
0.5	1.713	1.35[illegible]	1.[illegible]	1.083	1.008
0.4	1.460	1.163	1.062	1.004	0.994
0.3	1.197	0.990	0.948	0.946	0.986
0.2	0.953	0.873	0.890	0.932	0.988
0.1	0.802	0.884	0.938	0.973	0.997

Table 4: Efficiency of [illegible] for N=9

k	F_o=7.5[illegible] 1%	F_o=4.43 5%	F_o=3.[illegible] 10%	F_o=2.5[illegible] 20%	F_o=1.6[illegible] 50%
10	0.066	0.256	0.475	0.762	0.[illegible]
9	0.064	0.234	0.436	0.722	0.[illegible]
8	0.065	0.215	0.395	0.677	0.[illegible]
7	0.067	0.206	0.360	0.[illegible]	0.[illegible]
6	0.074	0.192	0.330	0.582	0.[illegible]
5	0.087	0.193	0.315	0.541	0.[illegible]
4	0.116	0.213	0.317	0.517	0.[illegible]
3	0.191	0.280	0.372	0.539	0.[illegible]
2	0.465	0.559	0.598	0.706	0.[illegible]
1	1.897	1.626	1.448	1.247	1.[illegible]
0.9	2.014	1.692	1.491	1.262	1.[illegible]
0.8	2.035	1.674	1.457	1.243	1.[illegible]
0.7	1.935	1.569	1.371	1.190	1.[illegible]
0.6	1.737	1.[illegible]	1.[illegible]	1.113	1.[illegible]
0.5	1.485	1.[illegible]	1.[illegible]	1.[illegible]	0.[illegible]
0.4	1.221	1.022	0.922	0.958	0.[illegible]
0.3	0.992	0.885	0.888	0.922	0.[illegible]
0.2	0.804	0.[illegible]	0.878	0.935	0.[illegible]
0.1	0.792	0.917	0.959	0.986	0.[illegible]

Table 5: Efficiency of R_k for N=11

k	$F_{10\%}$	$F_{5\%}$	$F_{2.5\%}$	$F_{2.5\%}$	$F_{50\%}$
10	0.108	0.408	[illegible]	0.881	0.991
9	0.099	0.361	0.674	0.847	0.988
8	0.092	0.317	0.547	0.804	0.983
7	0.088	0.279	0.482	0.751	0.975
6	0.079	0.249	0.426	0.693	0.962
5	0.095	0.231	0.386	0.626	0.945
4	0.117	0.232	0.358	0.572	0.921
3	0.179	0.279	0.381	0.560	0.900
2	0.426	0.437	0.567	0.691	0.922
1	1.680	1.615	1.433	1.239	1.042
0.9	1.994	1.671	1.464	1.250	1.043
0.8	1.976	1.627	1.422	1.222	1.037
0.7	1.826	1.488	1.315	1.158	1.025
0.6	1.585	1.296	1.171	1.074	1.008
0.5	1.313	1.097	1.028	0.990	0.992
0.4	1.056	0.932	0.915	0.930	0.983
0.3	0.852	0.829	0.860	0.913	0.984
0.2	0.737	0.825	0.887	0.946	0.993
0.1	0.831	0.949	0.978	0.993	0.999

Table 6: Efficiency of S [illegible]

k	N=3 c=-2.0 d=3.5	N=5 [illegible]	N=7 [illegible]	N=9 [illegible]	N=11 [illegible]
10	1.106	[illegible]	1.029	[illegible]	[illegible]
9	1.119	[illegible]	[illegible]	[illegible]	[illegible]
8	1.136	[illegible]	1.061	[illegible]	[illegible]
7	1.157	[illegible]	1.083	[illegible]	[illegible]
6	1.185	[illegible]	1.122	[illegible]	[illegible]
5	1.225	[illegible]	1.170	[illegible]	[illegible]
4	1.283	[illegible]	1.229	[illegible]	[illegible]
3	1.377	[illegible]	1.312	[illegible]	[illegible]
2	1.551	[illegible]	1.477	[illegible]	[illegible]
1	1.931	[illegible]	1.896	[illegible]	1.[illegible]
0.9	1.988	[illegible]	1.961	1.[illegible]	[illegible]
0.8	2.047	1.721	1.[illegible]	1.117	[illegible]
0.7	2.106	[illegible]	[illegible]	1.[illegible]	[illegible]
0.6	2.160	1.796	1.671	1.[illegible]	[illegible]
0.5	2.200	1.[illegible]	1.[illegible]	1.173	[illegible]
0.4	2.213	1.[illegible]	1.[illegible]	1.[illegible]	[illegible]
0.3	2.182	1.[illegible]	1.[illegible]	1.[illegible]	[illegible]
0.2	2.087	1.[illegible]	1.[illegible]	1.[illegible]	[illegible]
0.1	1.932	1.[illegible]	1.[illegible]	1.[illegible]	[illegible]

Table 7: Efficiency of S (Emphasizes range $-1 \le k \le 1.0$)

k	N=3 c=0.49 d=0.49	N=5 c=0.49 d=0.49	N=7 c=0.49 d=0.49	N=9 c=0.49 d=0.49	N=11 c=0.49 d=0.49
10	0.760	0.654	0.646	0.664	0.683
9	0.777	0.657	0.641	0.652	0.671
8	0.800	0.664	0.637	0.642	0.655
7	0.831	0.677	0.638	0.633	0.643
6	0.875	0.698	0.646	0.636	0.628
5	0.939	0.736	0.676	0.637	0.627
4	1.038	0.804	0.713	0.667	0.642
3	1.206	0.936	0.809	0.753	0.709
2	1.531	1.234	1.080	1.013	0.949
1	2.221	1.948	1.802	1.822	1.798
0.9	2.311	2.027	1.947	1.892	1.865
0.8	2.[illegible]	2.[illegible]	1.[illegible]	1.[illegible]	1.[illegible]
0.7	2.[illegible]	2.[illegible]	1.[illegible]	1.985	1.[illegible]
0.6	2.[illegible]	2.[illegible]	1.[illegible]	1.[illegible]	1.[illegible]
0.5	2.[illegible]	2.[illegible]	1.[illegible]	1.[illegible]	1.[illegible]
0.4	2.[illegible]	1.[illegible]	1.[illegible]	1.[illegible]	1.[illegible]
0.3	2.454	1.[illegible]	1.[illegible]	1.[illegible]	1.151
0.2	2.[illegible]	1.[illegible]	1.[illegible]	1.111	1.012
0.1	2.[illegible]	1.[illegible]	1.[illegible]	0.[illegible]	0.[illegible]

Table 8: Efficiency of [illegible] (optimum relative [illegible] Max [illegible]

k	N=3 c=-2.0 d=3.5	N=5 c=-2.0 d=5.0	N=7 c=-2.0 d=[illegible]	N=9 c=1.81 d=1.64	N=11 c=[illegible] d=[illegible]
10	1.052	1.[illegible]	[illegible]	0.[illegible]	[illegible]
9	1.059	1.[illegible]	[illegible]	[illegible]	[illegible]
8	1.067	1.[illegible]	[illegible]	[illegible]	[illegible]
7	1.078	1.[illegible]	[illegible]	[illegible]	[illegible]
6	1.092	1.[illegible]	[illegible]	[illegible]	[illegible]
5	1.109	1.[illegible]	[illegible]	[illegible]	[illegible]
4	1.133	1.[illegible]	[illegible]	[illegible]	[illegible]
3	1.166	1.[illegible]	[illegible]	[illegible]	[illegible]
2	1.209	1.[illegible]	[illegible]	[illegible]	[illegible]
1	1.[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]
0.9	1.[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]
0.8	1.[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]
0.7	1.[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]
0.6	1.[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]
0.5	1.[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]
0.4	1.[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]
0.3	1.[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]
0.2	1.[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]
0.1	1.[illegible]	1.[illegible]	[illegible]	[illegible]	[illegible]

	N=9	N=10	N=10	N=11
k	c=2.5 d=3.5	c=2.5 d=3.5	c=1.81 d=1.81	c=1.7 d=1.7
10	-.014	-.009	0.008	0.00
9	-.015	-.009	0.006	0.00
8	-.017	-.0	0.00	0.00
7	-.020	-.0	0.007	0.006
6	-.024	-.0	0.007	0.00
5	-.029	-.0	0.00	0.00
4	-.036	-.017	0.00	0.00
3	-.046	-.027	0.005	0.00
2	-.061	-.046	0.001	0.00
1	-.086	-.09	-.016	-.011
0.9	-.09	-.09	-.017	-.01
0.8	-.09	-.09	-.021	-.01
0.7	-.09	-.09	-.026	-.02
0.6	-.0	-.0	-.03	-.02
0.5	-.0	-.1	-.04	-.03
0.4	-.0	-.0	-.05	-.04
0.3	-.09	-.0	-.062	-.0
0.2	-.09	-.0	-.08	-.0
0.1	-.09	-.06	-.10	-.1

References

[1] Bancroft, T.A. (1944): On biases in estimation due to the use of preliminary tests of significance. Ann. Math. Stat. 15, 190-204

[2] Mehta, J.S. and Gurland, John (1969): Combination of unbiased estimators of the mean which consider inequality of unknown variances. Journal of the Amer. Stat. Assoc. 64, 1042-1055.

[3] Mehta, J.S. and Gurland, John (1970): On combining unbiased estimators of the mean. To appear in Technical Bulletin.

[4] Mehta, J.S. and Gurland, John (1969): On utilizing information from a second sample in estimating variance. Biometrika 56, 527-532.

A CLASS[illegible] OF BIVARIATE [illegible]

[illegible]
National Institutes of Health

I INTRODUCTION. Eisenhart (1947) distinguished two uses of analysis of variance which he designated as Type I and Type II. Type I provides a test of significance of the difference between estimates of population means. Type II provides a test for estimates of population variances. Eisenhart's treatment covered the general case of analysis of variance--but involved two important types of restrictions. First the "residual error" was assumed homogeneous with zero expected value. Second, all other parameters in Type I were assumed to have zero variances, and in Type II to have zero means. In the mixed model, parameters could be of either form, but, individually, where the means are not assumed zero the variances are and vice versa.

The present paper is limited to the case of two classes (the bivariate case) but removes both of these restrictions. This leads to a greatly enlarged variety of types and to a close parallelism with bivariate correlation. Two special cases, not previously treated, are discovered and appropriate formulae derived.

II MATHEMATICAL MODEL. Given

$$x = \alpha + \beta + \gamma + \delta + \varepsilon + \ldots$$

we have

$$E(x) = \bar{\alpha} + \bar{\beta} + \bar{\gamma} + \bar{\delta} + \bar{\varepsilon} + \ldots$$

$$V(x) = V(\alpha) + V(\beta) + V(\gamma) + V(\delta) + V(\varepsilon) + \ldots$$

This article has been reproduced photographically from the author's manuscript.

Preceding page blank

[illegible]

[illegible]

$$x = \bar{\alpha} + \bar{\beta} + \bar{\gamma} + \ldots + \alpha' + \beta' + \gamma' + \ldots,$$

[illegible] in Eisenhart's model, the parameters either have zero expected values or zero variances, but not both. In an exactly parallel way we may express a second variable y as

$$y = \bar{\alpha}^* + \bar{\beta}^* + \bar{\gamma}^* + \ldots + \alpha^* + \beta^* + \gamma^* + \ldots.$$

For [illegible] purposes it will be permissible to group the constant and/or the variable parameters in x and y into some lesser number adequate for the purpose in hand. If in particular, x and y are distributed in a bivariate normal, then the distribution factors into the product of one involving the means only, and the other involving the variances and covariances. For our purposes, it will be enough to ignore the distribution of the means, and to express x and y as the sum of two variables of the form

$$\begin{pmatrix} x = \tau + \zeta \\ y = \tau + \eta. \end{pmatrix} \tag{1}$$

If this is done, the variance-covariance (dispersion) matrix

$$\begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix} \tag{2}$$

becomes

$$\begin{pmatrix} \sigma_1^2+\sigma^2 & \sigma^2 \\ \sigma^2 & \sigma_2^2+\sigma^2 \end{pmatrix} \tag{3}$$

where

$$\sigma^2 = V(\tau);\ \sigma_1^2 = V(\zeta);\ \sigma_2^2 = V(\eta).$$

If ρ in (2) is negative*, then (1) is of the form

$$\begin{pmatrix} x = \tau + \zeta \\ y = -\tau + \eta \end{pmatrix} \tag{4}$$

and (3) becomes

$$\begin{pmatrix} \sigma_1^2+\sigma^2 & -\sigma^2 \\ -\sigma^2 & \sigma_2^2+\sigma^2 \end{pmatrix} \tag{5}$$

so that comparison of the two models can be obtained from (2) and (3).

III CORRELATION MODEL. Viewed as a purely mathematical object, the various special estimating and testing problems in the literature and certain simple extensions can be classified on the basis of restrictions on the elements of matrix (2) as follows:

A. Estimate ρ, σ_x^2 and σ_y^2. Test $\rho = 0$. This is the most common situation. The t-test applies; most naturally as a test of ρ.

B. Given σ_x^2 and σ_y^2, test $\rho = 0$. The t-test with infinite degrees of freedom, the normal test, for $\rho = 0$ applies.

C. Given $\rho = 0$, test $\sigma_x^2 = \sigma_y^2$, and estimate the common variance. This is the now classic case of estimating and testing equality of two independent variances.

D. Given $\sigma_x^2 = \sigma_y^2 = \sigma_c^2$, estimate ρ and σ_c^2, test $\rho = 0$. This example was treated by DeLury (1938).

E. Test $\rho = 0$, $\sigma_x^2 = \sigma_y^2$. This is the compound symmetry problem of Mauchly (1940).

* This possibility seems to have previously been overlooked (Anscombe).

[illegible] but known, test $\sigma_x^2 = \sigma_y^2$. This [illegible] Pitman (1938). Morgan (1939) showed that it has [illegible] greater power than the test in G.

G. [illegible] and test $\sigma_x^2 = \sigma_y^2$ whatever the value of ρ. This test was independently supplied by Pitman (1939) and Morgan (1939). It is described in Snedecor and Cochran (1967), Section 7.12.

Now suppose X is the sum of two independent random variables τ and ζ and similarly Y is the sum of τ and η, as given in equation (1). Then ρ, the correlation coefficient of X and Y is greater than or equal to zero and the matrix V was shown above to be

$$V = \begin{pmatrix} \sigma_1^2 + \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma_2^2 + \sigma^2 \end{pmatrix} \qquad (3)$$

The various possible tests in this formulation are:

H. Test $\sigma^2 = 0$. Estimate σ^2, σ_1^2, and σ_2^2. This is a test of the correlation of X and Y, expressed in the language of common and specific variances.

I. Given σ_1^2 and σ_2^2, test $\sigma^2 = 0$. This is the components of variance analogue of B above. The same test applies.

J. Given $\sigma^2 = 0$, estimate and test $\sigma_1^2 = \sigma_2^2$. This is the usual case of repeated measurements on constant but not necessarily equal standards; more generally, the test for equality of two independent variances as in C.

K. Given $\sigma_1^2 = \sigma_2^2 = \sigma_C^2$, estimate σ_C^2 and σ^2. Test $\sigma^2 = 0$. This is DeLury's problem expressed in the language of variance components. But whereas in its correlation formulation the problem appears to arise infrequently, in its components of variance formulation it is Eisenhart's Model II for Analysis of Variance.

L. Test $\sigma^2 = 0$, $\sigma_1^2 = \sigma_2^2$. This is Mauchly's problem expressed in the language of variance components. The components of variation interpretation of this test would be applicable wherever (a) observations occur in pairs, (b) the variance of the first member of each pair is to be compared with the variance of the second member. Thus in a paired comparison experiment, the residuals of the first members of each pair could be compared with the second as a test of the mathematical model underlying the design.

M. Test $\sigma_1^2 = 0$. This tests the legitimacy of treating the precision of one of two instruments, standards, or techniques as subject to no (negligible) error. The test is supplied in Maloney and Rastogi.

N. Given σ^2 known, test $\sigma_1^2 = \sigma_2^2$.

... Anderson and Bahadur, from which the following ... The likelihood criterion is

$$\lambda^{2/n} = \frac{\sum_i (X_i - \bar{X})^2 \sum_i (Y_i - \bar{Y})^2 - n^2\sigma^4}{\left[\sum_i (X_i - \bar{X})^2/2 + \sum_i (Y_i - \bar{Y})^2/2\right]^2 - n^2\sigma^4}. \tag{23}$$

For large n, $-2 \log \lambda$ is a chi-square r.v. with one d.f. (see Wilks, 1963, Chapter 13). Therefore, we reject the hypothesis H_o if $-2 \log \lambda \geq \chi^2_\alpha$ where χ^2_α is upper 100 percentile point of chi-square random variable with one degree of freedom.

Equation (23) can be written in a form which will be useful below and is mnemonic as well. Since $S_x S_y$ is the geometric mean of S_x^2 and S_y^2 and $\frac{1}{2}(S_x^2 + S_y^2)$ is their arithmetic mean, equation (23) becomes:

$$\lambda^{2/n} = \frac{(GM)^2 - e^2\sigma^4}{(AM)^2 - e^2\sigma^4} \tag{24}$$

writing GM for geometric mean, AM for arithmetic mean, and $e = n/(n-1)$.

Returning to equation (23) if, in particular $\sigma^2 = 0$, the likelihood ratio becomes

$$\lambda^{2/n} = \frac{S_x^2 S_y^2}{\left(\frac{S_x^2 + S_y^2}{2}\right)^2} = \frac{(GM)^2}{(AM)^2} = \frac{4F}{(1+F)^2} \tag{25}$$

where $F = S_x^2 / S_y$ is distributed as Snedecor's F r.v. with $(n-1, n-1)$ d.f., since, when $\sigma^2 = 0$, X and Y are independent r.v.'s.

The test based on λ is to reject H_0 if $0 < \lambda < k$, which is equivalent to $0 < F < k_1$ or $k_2 < F$, where $k_1 < 1$. As F is always taken to be greater than 1, the rule becomes reject H_0 if $F > k'$ at the chosen probability level, i.e., our test reduces to the ordinary F test when it is known that the population variance is zero.

Comparison of equations (24) and (25) exhibits the effect on the test of the existence and magnitude of population variance. Equation (24) is

$$\lambda^{2/n} = \frac{(GM)^2}{(AM)^2} \cdot \frac{(AM)^2}{(GM)^2} \cdot \frac{(GM)^2 - e^2\sigma^4}{(AM)^2 - e^2\sigma^4}$$

$$= \frac{(GM)^2}{(AM)^2} \cdot \frac{(AM)^2}{(AM)^2} \cdot \frac{(GM)^2 - e^2\sigma^4}{(GM)^2 - e^2\sigma^4} \cdot \frac{(AM)^2}{(GM)^2}$$

$$\leq \frac{4F}{(1 + F)^2} \qquad \text{(using 25)}$$

since $AM \geq GM$ for any set of positive numbers. It follows that, if a standard F test is applied to the variance estimates for the two instruments or procedures as if the effect of population variance were zero (equation (25)) the test will sometimes accept when the correct test (equation (23)) might reject. Conversely, when equation (25) is appropriate, discrimination will be sharper than in a test situation where population variance is present so that equation (23) must be used. In addition equation (23) can be used to gain insight into the benefit to be derived, hence into the care and expense which is justified, when the population variance is reduced; whether by

[illegible] test [illegible]s, animals, solutions or other material in an assay, or [illegible] ambient conditions or test practice.

[illegible] and test $\sigma_1^2 = \sigma_2^2$ whatever the value of σ^2. This statement of the problem differs from that of Morgan and Pitman in that here $\sigma_1^2 + \sigma^2 > \sigma^2$, whereas in their case the range of the main diagonal elements in (1) is $0 < \sigma_a^2 + \sigma^2 < \infty$. That their test is the likelihood ratio test in case 0 has been shown elsewhere.

The [illegible] be assembled into a table that brings out the symmetry between corresponding correlational and components of variance forms of expression.

Various Tests of Precision

MODEL					
CORRELATION			COMPONENTS OF VARIANCE		
Case	Assumption	Test	Case	Assumption	Test
A		$\rho = 0$	H		$\sigma^2 = 0$
B	σ_x^2, σ_y^2	$\rho = 0$	I	σ_1^2, σ_2^2	$\sigma^2 = 0$
C	$\rho = 0$	$\sigma_x^2 = \sigma_y^2$	J	$\sigma^2 = 0$	$\sigma_1^2 = \sigma_2^2$
D	$\sigma_x^2 = \sigma_y^2$	$\rho = 0$	K	$\sigma_1^2 = \sigma_2^2$	$\sigma^2 = 0$
E		$\rho = 0; \sigma_x^2 = \sigma_y^2$	L		$\sigma^2 = 0; \sigma_1^2 = \sigma_2^2$
			M		$\sigma_1^2 = 0$
F	$\rho \neq 0$(known)	$\sigma_x^2 = \sigma_y^2$	N	$\sigma^2 \neq 0$(known)	$\sigma_1^2 = \sigma_2^2$
G		$\sigma_x^2 = \sigma_y^2$	O		$\sigma_1^2 = \sigma_2^2$

Test of significance for variance-covariance parameters of a bivariate relation according to all possible non-trivial parameter restrictions. For all correlational models, the corresponding components of variance model yields the same test (not the same estimates) as its correlational analogue. No correlational model exists corresponding to item M.

REFERENCES

DeLURY, D.B. [1938].Note on correlations. Annals of Mathematical Statistics, 9: 149-151.

EISENHART, C. [1947]. The assumptions underlying the analysis of variance. Bioretrics, 3: 1-21.

FINNEY, D.J. [1938].The distribution of the ratio of estimates of the two variances in a sample from a normal bivariate population. Biometrika, 30: 190-92.

MALONEY, C.J. and HASTINGS, SURESH C. [1970].Significance test for grubb's estimates. Submitted to a technical journal.

MAUCHLY, JOHN W. [1940].A significance test for ellipticity in the harmonic dial. Terrestrial Magnetism and Atmospheric Electricity, 45: 145-148.

MORGAN, W.A. [1939].A test for the significance of the difference between the two variances in a sample from a normal bivariate population. Biometrika, 31: 13-19.

PITMAN, E.J.G. [1939].A note on normal correlation. Biometrika, 31: 9-12.

SNEDECOR, G.W. and W.G. COCHRAN [1967].Statistical Methods, sixth edition, Iowa State University Press.

WILKS, S.S. [1963].Mathematical Statistics, John Wiley and Sons, New York.

ANSCOMBE, F.J. [1948].Contribution to the discussion on D. G. Champernowne's "Sampling theory applied to autoregressive sequences". Journal of Royal' Statistics Society B, 10: 239.

COMPUTERIZED QUALITY CONTROL AS APPLIED TO UPPER ATMOSPHERE DATA

Oskar M. Essenwanger
U. S. Army Missile Command
Redstone Arsenal, Alabama

ABSTRACT. Any observational program, even if carried out with the best available instrumentation and carefulness to avoid instrumental error, may contain erroneous data introduced by preparation and transmission of data. Thus a good concept of quality assurance must precede any data analysis to avoid distortion and bias of results by erroneous records.

Three groups of analytical methods of quality assurance are discussed, inconsistencies, interrelationship of data, and frequency distributions. These methods have been developed at the Army Missile Command for screening radiosonde data by high speed computers. The goal is flagging of erroneous or suspicious records that these may be corrected.

Checking procedures include tests for trivial errors such as duplication, wrong sequence, missing data, special checks on identification numbers, etc. Other procedures utilize data interrelationships, in this special case the vertical structure of the atmosphere. Further checks employ screening of maxima and minima by exceedance criteria derived from the frequency distribution. The Weibull distribution has proven especially useful in this last phase of the checking procedure. Some pitfalls and limitations in the utilization of evaluation criteria are discussed.

The remainder of this article has been reproduced photographically from the author's manuscript.

I. INTRODUCTION

It is a generally accepted fact that all raw data from sample surveys and experiments contain errors. Even if an observational program has been carefully prepared and is carried out with the best available instrumentation which the program can reasonably afford, some errors are always present. They may be caused by instrumental deficiencies or inaccuracies of unqualified observers but can also be introduced by preparation or transmission.

Analysis of observational data can be no better than the quality of the available data. Thus a careful attention to quality assurance of the data must precede any data analysis. This vital part of any investigation should be a major concern to all investigators. Its main purpose is to avoid distortion of the analysis resulting from erroneous observations. This goal will determine the magnitude of the effort to be put into a quality control program and will influence the methods selected for quality assurance. Some results can be evaluated for soundness by qualified professionals and then a quality check could be omitted. The complexity of the atmosphere or the amount of the end product (such as tables of matrices or computer produced maps, etc.) made it virtually impossible in our case to judge correctness of the results afterwards.

Of course, one cannot make good data out of bad records, but a so-called "editing" process can make data more useful for analytical purposes. No editing program can eliminate the small random error. It is

the big mistake, such as a 10 or 20 degree error in temperature, which needs correction. Since the influence of erroneous records on results increases as the length of the observational series of record decreases, the need for quality assurance is the greater the shorter the record.

In some instances censoring of frequency distributions by eliminating extreme values may save elaborate screening procedures. This cannot be applied, however, if one of the analysis goals is the study of extremes. In the case of radiosonde data a second reason against censoring can be pointed out. Because of vertical consistency, data elimination at one level without attempt of correction may lead to discontinue the ascent from the censored level up. This may further reduce the already decreasing number of observations with altitude and may leave very few data reaching a top level of km, for example. Thus the cure is worse than the disease.

The availability of high speed computers has opened a new field in applying quality control methods and many methods considered too elaborate and cumbersome without computer use can now be employed without difficulties.

Some of the few basic principles, which reappear and can be commonly applied, may be demonstrated from the Army Missile Command's screening program of radiosonde data.

Since the author's detailed article is already scheduled for publication, (See 1969c) only some basic principles will be presented here.

11. CHECKING PROCEDURES[1]

Any automated screening procedure must be so designed that particular (consistent) errors as well as inconsistent errors can be recognized. This goal is rendered more difficult by the requirement that screening procedures should have a simple logic for computerized treatment.

1. Trivial Errors Check

Under this first category fall all errors which are easily recognized, and in many instances an automatic correction can be made. The errors can be divided largely into three groups: coding errors, data and limit checks.

In the first group one may encounter errors such as wrong location number, incorrect elevation, false identification code, mistakes in coding the type of observation, erroneous time, etc.

A second group comprises checks for completeness (missing data), duplication and sequence of the records. If it is intended to supplement the original data by automatic fill-in procedures, they can be incorporated in this phase of the screening procedure or at a later date.

The last group is the limitation violation, e.g. data are outside established tolerance limits or physical boundaries. For example in our case the dew point temperature cannot be greater than the air temperature and the wind direction cannot exceed 16 compass points or 360 degrees.

The examples given for the above error groups are some guidelines and are not exhaustive. They serve only as a demonstration for the type

of errors. Clearly, a procedure for trivial error checks evidently depends largely on the type of available data. The correction of the deficiency may also vary, e.g., if dealing with one station, an automatic correction could be made for wrong station code. In other instances elimination of the data may be necessary. If one had to establish a map by computer, this may be the only way to reduce the effect of large errors, while for other analyses time and personnel may be available to go through flagged observations and to painstakingly check their validity.

2. Error Checking by Adjacent Data

In this group inconsistencies are checked against adjacent data or a field of data in the horizontal (map or equations), vertical (cross-sections or equations), or by time relationship. The checking procedure depends largely on established physical or derived empirical laws. Again, procedures aim at flagging suspicious values by computer methods or correcting them if such procedures can be established. Tolerance limits of differences between two or more observations must be derived first.

a. Horizontal Checks

This type of checking process can be applied if computerized maps are available or become the end product or if records for neighboring stations for the same period of record are given. Under physical laws one may understand conditions like the gradient wind relationship etc. Empirical relationships between neighboring stations or thresholds of

tolerable differences between stations could be established. Whether these empirical relationships are derived in tabular form or as analytical expressions is not important, except that it is more convenient to work with mathematical statements for which computer programming is usually very simple. Changes of errors are considerably lower as opposed to table inputs, especially when these have more than one entry.

b. Vertical Relationship

The U. S. Army Missile Command's procedure of screening radiosonde data relies heavily on vertical relationships. Cross-sections could be used, but only if they are readily available or calculation of the cross-section by computer methods is the goal. The author does not know of any program at the present where space cross-sections have been utilized for data control. Time-sections have been employed by Canfield et al. (1966).

Our program checks two groups of elements, thermodynamic quantities and wind. In the thermodynamic portion the lapse rate between two consecutive observations at different altitudes is computed and compared with the dry adiabatic lapse rate. This method has proven quite efficient and satisfactory, as usually any error in pressure or temperature will show up eventually in a superadiabatic lapse rate either at the tested data pair or at the next step. For example, assume a 10^{o} negative error in the temperature. If it is the higher of two altitudes, it creates a superadiabatic lapse rate. If the error is positive, one

would have an inversion for this step of the program and the record is not flagged. The next step, however, would give a superadiabatic lapse rate.

The last observation of a radiosonde ascent cannot be checked by this method, as there is no other observation to compute the lapse rate. This last point could be checked by tolerance limits or other tools.

It should be added that superadiabatic lapse rates are not automatically eliminated by our program. The cause can be manifold. There may exist the unusual case of a true superadiabatic lapse rate in nature. One may have a temperature or pressure error or the data can be out of sequence by erroneous pressure. Thus all "suspicious" data are flagged and checked by a qualified meteorologist.

Since this simple tool worked so well for the thermodynamic parameters a similar principle was sought for the wind. In the beginning wind data were checked by the frequency distribution of wind shear with techniques established by Essenwanger et al. (1961). This is usually cumbersome and expensive, as computations of frequency distributions are generally costly. The difficulty in establishing a unique relationship was recognized by Finger et al. (1965) who established vertical shear limits for wind checks in tabular form for a few thresholds of layer thickness. However, their method requires detailed criteria depending on layer thickness, wind speed, and difference of direction or speed of

two kind of [illegible]. The author (1969) has derived a relationship between the vector shear (Δv) and the shear interval (Δh)

$$\Delta v = a_o (\Delta h)^{a_1} \quad (1)$$

The exponent for extreme value was found to be 1/3 (see also Essenwanger and Reiter 1969a). For use in our program eqn. (1) had to be modified to accommodate normalized shear interval, thus $\Delta v = V_e \Delta h$, resulting in

$$V_e = a (\Delta h)^{-\frac{1}{3}} \quad (2)$$

Where V_e denotes the total vector shear. With $a = 2.5$, a reasonable threshold V_e (m sec^{-1} per interval) is found. All values exceeding V_e are flagged.

Equation (2) expresses a unique relationship similar to the lapse rate as a convenient and simple tolerance criteria.

c. Time Series

All elements showing some form of time relationship could be checked by methods taking advantage of this relationship. It does not matter whether the time relationship is periodic or aperiodic. However, in all time related checking procedures the time relationship must be established first.

In case of periodic variations it is quite convenient to represent records by a Fourier series and check an expected versus an observed value. A tolerance limit for a maximum (absolute) difference

from the expected value may be determined by statistical methods of error theories. Sometimes it may be quite sufficient and suitable to use subjective tolerance limits.

If an aperiodic time relationship (e.g. persistence) has been found, tolerance criteria for time differences can be employed. In all cases an expected value is tested against the observation.

A time checking procedure can be applied, even if no functional relationship can be found. Although time differences may be randomly distributed, a tolerance criterion can be developed similar to that described in a later chapter on frequency distributions. If the difference exceeds a certain magnitude, it may indicate an erroneous observation.

7. Frequency Distribution Checks

Although methods described in the previous sections should catch the bulk of errors, some mistakes may slip through. Let us assume that the surface observation of a radiosonde ascent is missing. Vertical consistency could not discover this mistake. Although it could have been flagged in the trivial error check. Other examples can be given where vertical consistency existed, but the total ascent was either too warm or cold. These errors can be checked against a frequency distribution.

In the Army Missile Command's earlier screening procedure preliminary frequency distributions were established, with printout of the first five maxima and minima, mean and standard deviation. Visual inspection of the frequency distribution then revealed isolated observations. Vertical profiles for the maxima and minima were drawn and suspicious records could be detected by irregularities in profiles.

This process was time consuming, and not too many erroneous ascents were discovered, since the majority of corrections had been made. Nevertheless, all frequency distributions had to be inspected. This phase of the program was costly, too, as frequency distributions had to be grouped by small class intervals to detect isolated records and class intervals shifted from month to month or by altitude. This phase of the program was modernized by utilizing only mean and standard deviation and selecting suspicious values by a predetermined threshold x_{th} to be exceeded only a certain percentage of the time. This eliminates the establishment of frequency distributions and reduces the printout as only flagged observations appear.

It is evident that correct as well as incorrect observations will be flagged and printed out, as one should expect a number of observations exceeding the threshold x_{th} in agreement with the selected percentage figure. Unfortunately there is no easy way to separate the two groups by computer methods, as large deviation can be caused by extreme weather events. All cases must be judged by their own merits. It is reiterated that acceptance, correction or deletion of an observational record depends largely on the purpose of any analysis and existing possibilities. We have found it quite convenient to make available for any flagged value the threshold for 99% and the frequency of occurrence which the flagged observation would have in a theoretical distribution law. These values are helpful guidelines for evaluation, but are generally not sufficient by themselves for a decision that the observation is erroneous or not. It should further be pointed out that censoring of the frequency distribution cannot be applied in our particular case. Especially extreme value data analysis is part of the subsequent research topics. Censoring would not solve the quality assurance problem.

a. Gaussian Distribution

The critical problem is the determination of the threshold x_{th}, outside of whose boundary observations should be flagged. In statistical terms, one has to select a certain point of the cumulative distribution on one or both sides of this curve. The computation of the cumulative distribution is cumbersome for most types of distribution laws

as it involves integrating frequency density functions. In the Army Missile Command's earlier version empirical cumulative distributions were computed to secure close agreement with the observed frequency. This had the advantage of the frequency curve being independent from the statistical type, or the mean and the standard deviation of the distribution. Later this was replaced by establishment of frequency distributions, which display less complexity in computer programming.

If the element follows an approximate Gaussian normal distribution, one could determine the threshold by

$$x_{th} = \bar{x} \pm a\sigma \qquad (3)$$

where $\bar{x}$ is the mean value, σ the standard deviation and the coefficient "a" would be determined by the desired percentage exceedance, e.g. a = 3.0 for .135% of the observations beyond that point. All observations above x_{th} would then be flagged and printed out.

Since the relationship between the cumulative distribution and the standard deviation is known for the Gaussian distribution, the establishment of thresholds should not create any problem for meteorological elements following this distribution law. Gaussian laws apply to most thermodynamic quantities.

b. The Weibull Distribution

If Eqn. (3) were applied to meteorological elements not in agreement with the Gaussian law, one would have either too many flagged observations or not enough, depending on the deviation. Since the

relationship between standard deviation and cumulative distribution for other types of distributions is complex and generally cannot be found in simple tables, the ideal solution would be a cumulative frequency law versatile enough to adjust to a variety of types with good approximation. Thus we applied the Weibull distribution with considerable success in our screening procedure.

The Weibull distribution is defined as a cumulative type

$$F(x) = 1 - e^{-\left(\frac{x - \gamma}{\theta}\right)^{\beta}} \qquad (4)$$

with γ, θ and β as the reference, scale and shape parameter, respectively. Any percentage $F(x)$ can be related to x_{th} by the modification of equation (4) to

$$x_{th} = \theta \sqrt[\beta]{\ln P} + \gamma \qquad (5)$$

where

$$P = 1 / (1 - F(x)) \qquad (5a)$$

The estimation of the parameters is the only difficulty left. Maximum likelihood estimation for all three parameters cannot be performed analytically and solution is very time consuming. Thus the utilization of the maximum likelihood method for three parameters would have increased costs compared to frequency distributions. Simpler methods exist when $\gamma = 0$ (see Kao, 1958 or Menon, 1963), however, the

assumption $\gamma = 0$ would reduce the flexibility of adjustment for the Weibull distribution and would limit the ability to fit the frequency distribution. Since the major goal in the checking procedure is the establishment of a threshold value x_{th}, the reader may find a parameter estimation by moments, developed by the author (1968, 1969b) quite convenient.

$$\gamma_1 = (c - 3ab + 2a^3)/(b - a^2)^{3/2} \tag{6a}$$

γ_1 denotes the skewness, the ratio of the third moment (reference mean) to the cube of the standard deviation, $\gamma_1 = \mu_3/\sigma^3$

Since a, b and c depend on β only, a computer solution of (6a) is relatively easy or tables can be used (see Essenwanger, 1968, 1969b)

$$a = \Gamma(1 + 1/\beta) \tag{7a}$$

$$b = \Gamma(1 + 2/\beta) \tag{7b}$$

$$c = \Gamma(1 + 3/\beta) \tag{7c}$$

With a known, the other parameters become

$$\theta^2 = \sigma^2/(b - a^2) \tag{6b}$$

and

$$\gamma = \bar{x} - \theta \cdot a \tag{6c}$$

The three moments of the distribution must be known for application of eqns. 5, 6, and 7. In two cases two moments are sufficient.

As shown by the author (1968) the γ_1 can be approximated by

$$\gamma_1 = 1.4047E - .0646\sigma + .0987 \qquad (8a)$$

for wind and by

$$\gamma_1 = 3.1223E - .3680\sigma - .4515 \qquad (8b)$$

for the total vector wind shear

$$E = \epsilon_3/\sigma^3 = \bar{x}\,(1 + 3d + 2d^2)/\sigma^3 \qquad (9a)$$

with

$$d + 1 = \sigma^2/\bar{x} \qquad (9b)$$

The second case employs the Weibull distribution for elements whose distributions follow the Gaussian law.

Thus β can be determined a priori. Eqn. (6a) gives $\beta = 3.60$. If the squared difference of the Gaussian and the Weibull distribution at steps of half a standard deviation σ within $\pm 3.5\sigma$ is computed and summed up, a minimum is discovered at $\beta = 3.55$. Table 1a exhibits the frequencies for the Gaussian and the Weibull distribution (cumulative at left and density at right) for $\beta = 3.55$. All differences are less than 1%. The last columns in both sections contain the differences for the β, if selection is made for the smallest possible maximum deviation of any frequency within $\pm 3.5\sigma$ range.

Of more importance may be the agreement between the x-values as these are used to establish the flagging limit of Eqn. (5). The

Table 1. Comparison of Weibull Distribution With Gaussian Distribution

a. Frequency (in %)

	Cumulative Frequency					Frequency Density			
		β = 3.55		β = 3.62		β = 3.55			β = 3.33
x/σ	Gauss	Weibull	Difference	Difference	x/σ	Gauss	Weibull	Difference	Difference
-3.5	.023	-	.023	.023					
-3.0	.135	.004	.131	.128					
-2.5	.621	.313	.308	.273	< -2.5	.621	.313	.308	.258
-2.0	2.275	2.099	.176	.110	-2.5 to -2.0	1.654	1.786	-.132	-.153
-1.5	6.681	7.041	-.360	-.409	-2.0 to -1.5	4.406	4.942	-.536	-.583
-1.0	15.866	16.664	-.798	-.779	-1.5 to -1.0	9.185	9.623	-.438	-.676
- .5	30.854	31.413	-.559	-.462	-1.0 to - .5	14.988	14.749	.239	-.034
.0	50.000	49.799	.201	.321	- .5 to .0	19.146	18.386	.760	.674
.5	69.146	68.453	.693	.758	.0 to .5	19.146	18.655	.491	.668
1.0	84.134	83.621	.513	.495	.5 to 1.0	14.988	15.167	-.180	.096
1.5	93.319	93.259	.060	-.001	1.0 to 1.5	9.185	9.638	-.453	-.302
2.0	97.725	97.895	-.170	-.221	1.5 to 2.0	4.406	4.637	-.231	-.353
2.5	99.379	99.524	-.145	-.169	2.0 to 2.5	1.654	1.629	.025	-.060
3.0	99.865	99.926	-.061	-.067	> 2.5	.621	.475	.146	-.173
3.5	99.977	99.992	-.015	-.017					

Table 1. (Continued)

b. x-scale (in σ Units)

	Two-Sided				One-Sided		
	β = 4.36		β = 4.26	β = 5.53	β = 5.54	β = 2.97	β = 2.96
Gauss	Weibull	Difference	Difference	Difference	Difference	Difference	Difference
-4.0	-3.467	.533	.585	.002	-.002	1.365	1.371
-3.5	-3.255	.245	.306	-.147	-.150	.955	.961
-3.0	-2.933	.067	.105	-.217	-.219	.603	.608
-2.5	-2.535	-.035	-.015	-.217	-.219	.326	.329
-2.0	-2.074	-.074	-.063	-.166	-.167	.131	.133
-1.5	-1.562	-.062	-.058	-.089	-.089	.016	.017
-1.0	-1.026	-.026	-.027	-.013	-.013	-.035	-.036
- .5	- .488	.012	.009	.041	.042	-.043	-.044
.0	.036	.036	.033	.064	.064	-.026	-.027
.5	.537	.037	.036	.051	.051	-.001	-.001
1.0	1.013	.013	.014	.001	.001	.025	.025
1.5	1.463	-.037	-.033	-.081	-.082	.045	.044
2.0	1.888	-.112	-.104	-.194	-.194	.055	.055
2.5	2.290	-.210	-.197	-.332	-.333	.054	.057
3.0	2.680	-.320	-.311	-.494	-.495	.043	.046
3.5	3.062	-.438	-.441	-.674	-.676	.022	.028
4.0	3.389	-.611	-.583	-.869	-.871	-.002	.004

minimum sum of the squared differences between the Gaussian and the Weibull distribution for half units of σ within $\pm 4\sigma$ lies at $\beta = 4.36$ (Table 1b). It should be noticed, however, that the differences for the two-sided fit increase towards the marginal classes. This is a handicap, but is acceptable since it is preferable to flag more than the expected number of observations rather than less. The threshold could be adjusted, too. Again, if a minimum of the absolute deviation is desired, one would select $\beta = 4.26$ with deviations smaller than 0.6σ at the ends.

Since it is known whether an observation is below or above the mean value, a one sided fit solves the problem of poor agreement towards the ends. Good approximation for the minimum threshold can be obtained with a β of 5.53 or 5.54, while one may select a β of 2.96 or 2.97 for the maximum end. The differences are displayed in the right portion of Table 1b.

The advantage in using the Weibull distribution for flagging instead of the concept of the normal distribution lies in the easy computation of related frequency values for the flagged observation with Eqn. (5) and (5a). This eliminates any tabular input as necessary for Eqn. (3) and one program can be applied to all types of frequency distributions.

c. Elements With Various Types of Distributions

Thermodynamic quantities and wind can be treated with techniques as outlined previously. The Weibull distribution is very flexible and thus can be utilized for the purpose of flagging for numerous elements. Some distributions may display untolerable discrepancies.

Transformation of scale sometimes helps, such as a logarithmic progression of visibility data. This must be left to the individual analyst. The Weibull distribution is very flexible and transformation can usually be avoided.

6. Some Remarks of Caution

It is reiterated that no quality assurance program can make good data out of bad records. These programs can only contribute to an "editing" of data, after which the larger errors (hopefully) have been eliminated. Since these large errors can bias any statistical or computer result, the correction of these obvious mistakes is necessary. It must be cautioned, however, that correction methods cannot be geared to an expected analysis result, as all observations contradictory to an assumed hypothesis to be tested by these data are then eliminated. Correction methods must be independent of subsequent analysis. One cannot check persistence, for example, if the majority of data have been filled in by methods derived from persistence.

The editing process by "experts" is usually cumbersome, but correction methods by computers must be carefully designed. Where consistency equations can be obtained, methods for random error corrections can be developed With the complexity of the atmosphere it is difficult, however, to pinpoint unequivocally differences between a rare event and an obvious mistake.

Any correction method should be based upon known or derived principles of error sources. Sometimes data are rectified which later prove correct in the light of expanded knowledge.

Establishment of threshold values is arbitrary. Threshold values must be designed to catch all the large errors without the burden

of reviewing too many data by the expert. Any time a large pile of flagged data appears for a particular data sample, a search for a systematic error should precede any detailed correction operation. This systematic error can then be corrected before other computer runs are made. Sometimes a big bulk of printout can be caused by improper selection of the thresholds. Then an adjustment will give reasonable amounts.

It should be further mentioned that selection of thresholds succeeds for unlimited distributions only. It would be absurd, for instance, to flag all calms in surface wind distributions or all dry records for precipitation data. Elements with U-shaped distributions could in general not be checked by frequency methods.

III. CONCLUSIONS

Analytical methods for the editing of observational data have been divided into three major groups, the checking of inconsistencies (trivial errors), the procedures employing a set of data with interrelationship, and utilization of frequency distributions. The methods presented may serve as a guideline and cannot be exhaustive, as the complexity of the atmosphere with its different meteorological parameters necessitates individual techniques depending on the treated element. The three described groups of error checks are common with any quality assurance program.

It is repeated that editing of data cannot replace a carefully carried out observational program with adequate instrumentation. One can assure, however, that large mistakes and systematic errors from various sources are discovered and any bias of the results due to erroneous data is largely reduced. The small random error cannot ordinarily be eliminated.

Although the data may have gone through quality assurance programs several times before they reach the investigator, it is nevertheless advisable to resubmit the data to a screening procedure. Editing of data by other investigators or installations does not automatically guarantee that the received data are free of mistakes.

The goal of the editing process should not be to correct nature and reject data which do not fit into a predetermined model or hypothesis,

whether or not the data should be based upon known or discovered sources of error only. If the latter is kept in mind, analytical methods of quality assurance will serve their useful purpose.

REFERENCES

Canfield, N. L., Smith, J. E., and Vaughan, W. W., 1966: Progress in Circumventing Limitations of Upper Wind Records, J. App. Meteor. 5, No. 3, 301 - 303.

Essenwanger, O. M., Vaughan, W. W. and Bradford, R., 1961: On Verification of Upper Air Winds by Vertical Shears and Extremes, Mon. Wea. Rev., 89, 197 - 204.

Essenwanger, O. M., 1963: On the Derivation of Frequency Distributions of Vector Wind Shear Values for Small Shear Intervals. Geofis. Pura Appl., 56, 216 - 224.

Essenwanger, O. M., 1968: On Deriving 90 - 99% Wind and Wind Shear Thresholds from Statistical Parameters. Proc. of the Third Conf. on Aerospace Meteor., New Orleans, publ. by Am. Meteor. Soc., Boston, Mass.

Essenwanger, O. M. and Reiter, E. R., 1969a: Power Spectrum, Structure Function, Vertical Wind Shear and Turbulence in Troposphere and Stratosphere. Archiv. Meteor., Geoph., Bioklim, in print 1969.

Essenwanger, O. M., 1969b: Fitting of the Weibull Distribution with Non-Zero Location Parameter, AMICOM Techn. Report, in publication.

Essenwanger, O. M., 1969c: Analytical Procedures for the Quality Control of Meteorological Data. Proc. Symp. Meteor. Observations and Instrum., (published by Am. Meteor. Soc., Boston).

Finger, F. G., Woolf, H. M. and Anderson, C. E., 1965: A Method for Objective Analysis of Stratospheric Constant Pressure Charts. Mon. Wea. Rev., 93, 619 - 638.

Kao, J. H. K., 1958: Computer Methods for Estimating Weibull Parameters in Reliability Studies. Inst. of Radio Engr. Trans. on Rel. and Qual. Cont. PGRQC, 15 - 22.

Menon, M. V., 1963: Estimation of the Shape and Scale Parameters of the Weibull Distribution. Technometrics, 5, No. 2, 175 - 182.

ABSTRACT

A STATISTICAL MODEL FOR THE ANALYSIS OF SIMULTANEOUS TWO-STATION IONOSPHERIC SOUNDINGS

Dr. Erwin Biser Mr. Richard D'Accardi
US Army Electronics Command, Fort Monmouth, NJ

1. Ionospheric sounder data characteristics change as the distance between two sounder stations is increased from 0-500 Km. It is therefore desirable to know about the degree of correlation one can expect between vertical incidence (single station) data and oblique incidence (two-station) data. It will be shown that a single ionospheric sounder (ionosonde) operating in the vertical incidence mode can provide useful data over an area of 60 Km radius.

2. Experimentation was performed in the 2-16 MHz frequency range using two ionosondes, one as a fixed terminal and the other as a mobile terminal. Each terminal made scheduled soundings every ten minutes from 0530 to 1730 hours for ten days. While the fixed terminal was transmitting and receiving its own signal, the mobile terminal would simultaneously receive the same transmission; likewise for the mobile with respect to the fixed terminal. As each ionosonde transmitted and received in the vertical incidence mode, the other sounder, receiving the same transmission, completed the oblique ionospheric mode. (An oblique mode or path is one between two stations space a distance apart; a vertical mode or path occurs when either station receives its' own transmission.)

3. The experiment was designed primarily for a paired difference model, that is, the pairing of data occurred as planned by the experiment. The data were also analyzed by a paired comparison method to focus on the gain of information achieved with the paired difference or randomized block design, and to show that vertical incidence and oblique incidence ionosonde data are good estimators of each other over short distances.

4. The application of a similar method of analysis will hopefully be used in future experimentation to substantiate a high degree of correlation between vertical and oblique incidence soundings over field army distances (0-300 Km).

This article has been reproduced photographically from the authors manuscript.

Preceding page blank

The objective of the present data analysis is to show that daily ionospheric soundings taken at vertical incidence (VI) are very nearly the same as oblique incidence (OI) soundings taken over a 60 Km path (see Figure 1). We are interested in formulating hypothesis tests to determine whether or not the Vertical Incidence data (population I) is nearly the same or is, in fact, identical to the Oblique Incidence data (population II). The analysis investigates a total of 85 daily measurements of critical frequencies performed over a nine-day period, taken every ten minutes from 0530 hours to 1930 hours, for a 60 Km path (see Figures 6, 7). This yielded nine observations of critical frequency per time slot. Samples of raw data appear in Figures 2, 3, 4, 5. In order to test whether or not a given hypothesis is supported by a set of data, we devised a rule of procedure dependent upon certain calculations obtained from a sample of the data, and decided to accept or to reject the hypothesis formulated[3]. Two experiments, E_1 and E_2 were used in comparing the means of population I (VI) and those of population II (OI). The homogeneity of variance was tested by the use of the F test, where σ_o^2 was compared to σ_v^2.

To test homogeneity of variance, the variances of the vertical and oblique incidence data were paired. The 85 grouped values were:

$$(1) \qquad T = \left\{ \left(\frac{S_o^2}{S_v^2}\right)_1 , \left(\frac{S_o^2}{S_v^2}\right)_2 , \ldots\ldots , \left(\frac{S_o^2}{S_v^2}\right)_{85} \right\}$$

Tests of hypotheses for the equality of two variances were formulated as follows:

$$(2) \qquad H_o:\sigma_o^2 = \sigma_v^2 \qquad \text{vs.} \qquad H_1:\sigma_o^2 \neq \sigma_v^2$$

or:

$$H_o: \frac{\sigma_o^2}{\sigma_v^2} = 1 \qquad \text{vs.} \qquad H_1: \frac{\sigma_o^2}{\sigma_v^2} \neq 1$$

The rejection region is: $\frac{\sigma_o^2}{\sigma_v^2} \geq k$, where k is found by specifying the significance level $\alpha = .01$. The following probability function describes the relationship:

$$(3) \qquad Pr\left[\frac{\sigma_o^2}{\sigma_v^2} \geq k \;\middle|\; \frac{\sigma_o^2}{\sigma_v^2} = 1\right] = \alpha$$

Under the null hypothesis H_o, $\frac{\sigma_o^2}{\sigma_v^2}$ has an F distribution with $(n-1)$, $(n-1)$ degrees of freedom, which results in $k = F\left[(n-1), (n-1); (1-\frac{\alpha}{2})\right]$ and the rejection regions are:

$$(4) \qquad \frac{\sigma_o^2}{\sigma_v^2} \geq F[(n-1), (n-1); (1-\tfrac{\alpha}{2})]$$

$$\frac{\sigma_o^2}{\sigma_v^2} \geq F[(n-1, (n-1); \tfrac{\alpha}{2}] \qquad \text{for } \alpha = .01$$

If these inequalities are satisfied by S_o^2 and S_v^2, then we can conclude that the estimated variances are significantly different at $\alpha = .01$ level of significance. That is to say, H_o is rejected when:

(5) $$F[(n-1),\ (n-1);\ \frac{\alpha}{2}\ ] \geq \frac{\sigma_o^2}{\sigma_v^2} \geq F[(n-1),\ (n-1);\ (1-\frac{\alpha}{2})\ ]$$

By letting S_o^2 and S_v^2, the sample variances, estimate σ_o^2 and σ_v^2, we form the F ratio:[1]

(6) $$F = \frac{\frac{S_o^2}{\sigma_o^2}}{\frac{S_v^2}{\sigma_v^2}}$$

which has the F distribution with (n-1), (n-1) degrees of freedom. F involves the ratio $\frac{\sigma_o^2}{\sigma_v^2}$ but is independent of σ_o^2 and σ_v^2, therefore:

(7) $$Pr\left[F[(n-1),(n-1);\frac{\alpha}{2}] \leq \left(\frac{\frac{S_o^2}{\sigma_o^2}}{\frac{S_v^2}{\sigma_v^2}}\right) \leq F[(n-1),(n-1);(1-\frac{\alpha}{2})]\right] = 1-\alpha$$

In testing the means the observations were grouped into 85 values for each experiment:

(8) $$E_1 = \{\bar{d}_1, \bar{d}_2, \ldots\ldots, \bar{d}_{85}\}$$

where $d_i = x_o - y_v$, and $\bar{d} = (1/n)\sum_1^9 d_i$, $i = 1, \ldots\ldots, 9$ per time slot.

(9) $$E_2 = \{\delta_1, \delta_2, \ldots\ldots, \delta_{85}\}$$ where: $\delta_j = \bar{x}_{o_j} - \bar{y}_{v_j}$ per time slot

nd, $\bar{x}_{o_j} = (1/n) \sum_{1}^{9} x_{o_i}$ per time slot

$\bar{y}_{v_j} = (1/n) \sum_{1}^{9} y_{v_i}$ per time slot

These sets consist of the mean values of the differences between OI and VI, (E_1), and the differences between the means of the two populations, (E_2), repeatedly taken at the same time daily for nine days. These two sets are assumed to be normally distributed with variance σ_d^2 and σ_p^2 so that the means of differences, $\bar{d}_i$, and the difference between the means, δ_i, are normally distributed. Since σ_d^2 and σ_p^2 are unknown, take estimates of the variances for each time slot for E_1 and E_2 are:

(10) $$Sd_i^2 = \frac{1}{n-1} \sum_{i=1}^{9} (d_i - \bar{d})^2 \quad \text{for } E_1, \text{ where}$$

$n = 9$ samples per time slot.

$\bar{d}$ = mean of the differences between OI and VI per time slot.

d_i = difference between f_{OI} and $f_{VI} = x_{o_i} - y_{v_i}$.

(11) $$\text{For } E_2, \quad S_p^2 = \frac{(n_1-1)\, S_o^2 + (n_2-1)\, S_v^2}{n_1+n_2-2} = \frac{S_{x_o}^2 + S_{y_v}^2}{2}$$

where $n_1 = n_2$, and:

$$S_{x_o}^2 = \frac{1}{n-1} \sum_{1}^{9} (x_{o_i} - \bar{x}_o)^2$$

$$S_{y_v}^2 = \frac{1}{n-1} \sum_{i} (y_{v_i} - \bar{y}_v)^2$$

S_p^2 = pooled estimate of variance for E_2

n = 9 samples per time slot

$\bar{x}_o$, $\bar{y}_v$ = means of OI and VI populations per time slot.

x_{o_i}, y_{v_i} = OI and VI data per time slot.

The t-statistics employed are:[2]

(12a) $$t_{n-1} = \frac{\bar{d}_i \sqrt{n}}{S_{d_i}} \quad \text{for } E_1, \text{ and}$$

(12b) $$t_{n^*} = \frac{\bar{x}_o - \bar{y}_v}{S_p \sqrt{\frac{2}{n}}} = \frac{\delta_{ov}}{S_p \sqrt{\frac{2}{n}}} \quad \text{for } E_2,$$

where $n=9$ samples per time slot, and $n^* = n_1 + n_2 - 2 = 16$ degrees of freedom

$S_p = \sqrt{S_p^2}$, the pooled standard deviation for E_2,

$S_{d_i} = \sqrt{S_{d_i}^2}$, the standard deviation for E_1 .

Therefore, the populations are t-distributed with (n-1), and n* degrees of freedom. The first experiment or "paired" difference test, E_1, concerned itself with analyzing the means of the differences between OI and VI data. The second experiment or "paired" comparison test, E_2 was

concerned with analyzing the difference between the means of the two populations. For E_1 the following hypothesis was formulated:

(13a) $\quad H_o: \mu_o = \mu_v \quad$ vs. $\quad H_1: \mu_o \neq \mu_v$

where $\quad \bar{d}_i = (1/n) \sum_{j=1}^{9} (x_{o_i} - y_{v_i}) \quad$ for each time slot.

For E_2 the following hypothesis was formulated:

(13b) $\quad H_o: \mu_o = \mu_v \quad$ vs. $\quad H_1: \mu_o \neq \mu_v$

where $\quad \delta = \bar{x}_o - \bar{y}_v \quad$ for each time slot,

and $\quad \bar{x}_o = (1/n) \sum_{i=1}^{9} x_{o_i}$,

$\bar{y}_v = (1/n) \sum_{i=1}^{9} y_{v_i}$,

and x_o and y_v are oblique and vertical incidence data respectively. That is to say, we will test a null hypothesis H_o, (that $\bar{d}_i$ or $\delta_i = 0$) vs. H_1 . If we accept the hypothesis, this would, of course, indicate that the difference between OI - VI = 0 for each time slot at $\alpha = .01$. If we assume the alternate hypothesis H_1 to be true, then the OI and VI data would be significantly different.

The critical region of these tests are:[2]

(14) $\quad \dfrac{\bar{d}_i \sqrt{n}}{S_{d_i}} > t\,[(n-1);\, \alpha/2] \quad$ for E_1, which can be written as:

$$t[(n-1);\ \alpha/2] > \frac{\bar{d}_i \sqrt{n}}{S_{d_i}} > t[(n-1);\ (1-\alpha/2)]$$ for a two-tailed test, and

(15) $$\frac{\bar{x}_o - \bar{y}_v}{S_p \sqrt{2/n}} > t[(n^*);\ \alpha/2]$$ for E_2 which can be written as:

$$t[(n');\ \alpha/2] > \frac{\bar{x}_o - \bar{y}_v}{S_p\sqrt{2/n}} > t[(n');\ (1-\alpha/2)]\ ,$$

where $n=9$, $n^*=(2n-2)=16$ degrees of freedom, and with $n=9$, $(n-1)=8$ degrees of freedom. This indicates that if (14) is satisfied by $\bar{d}_i$ and S_{d_i}, and (15) is satisfied by δ_i and S_p, the tests are rejected under the null hypothesis H_o and that $\bar{d}_i$ and δ_i does differ significantly from "0" in the critical region (region of rejection).

The critical regions can be explained by the following probabilities:(5)

(16) $$Pr\left[\frac{|\bar{d}_i \sqrt{n}|}{S_{d_i}} > t[(n-1);\ \alpha/2]\right] = \alpha$$ for E_1 which can be written:

$$Pr\left[t[(n-1);\ \alpha/2] > \frac{\bar{d}_i \sqrt{n}}{S_{d_i}} > t[(n-1);\ (1-\alpha/2)]\right] = \alpha\ ,$$

and

(17) $$Pr\left[\frac{|\bar{x}_o - \bar{y}_v|}{S_p\sqrt{2/n}} > t[n^*;\alpha/2]\right] = \alpha$$ for E_2 which can be written:

$$\Pr\left[t[n^*;\ \alpha/2] > \frac{\bar{x}_o - \bar{y}_v}{S_p\sqrt{2/n}} > t[n^*;\ (1-\alpha/2)]\right] = \alpha ,$$

where $\alpha = .01$ is the pre-determined critical region, or region of rejection. Regarding the comparison of means and of variances, if the null hypothesis H_o is found to be false, then the power function, π, would be used to find the probability that the alternate hypothesis H_1 will fall completely in the critical region. Let β = region of acceptance of the alternate hypothesis H_1 . Normally $\pi = 1-\beta$ should be very large or β very small. To illustrate the concept of a rejection region, suppose we have the following hypothetical probability density function of a variable X :

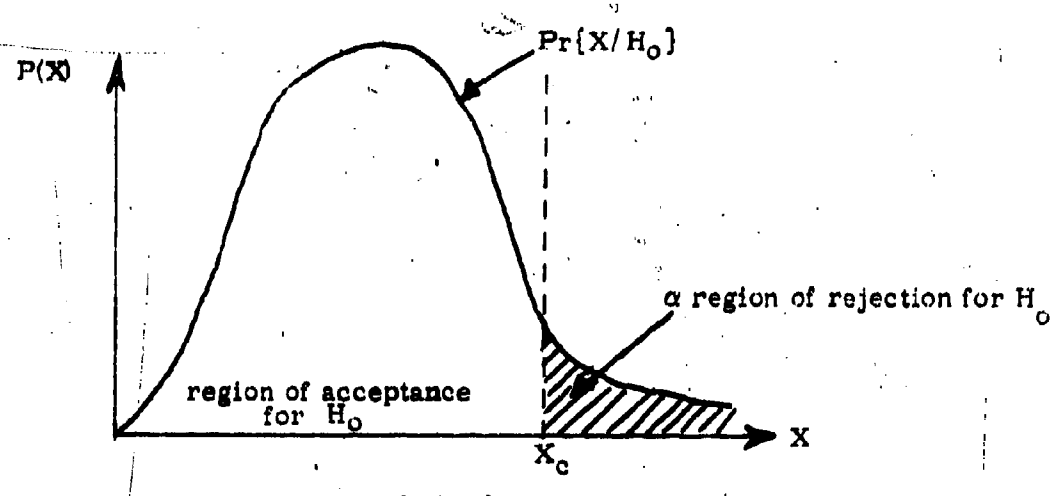


Illustration (a), $\Pr\{X/H_o\}$ vs. X

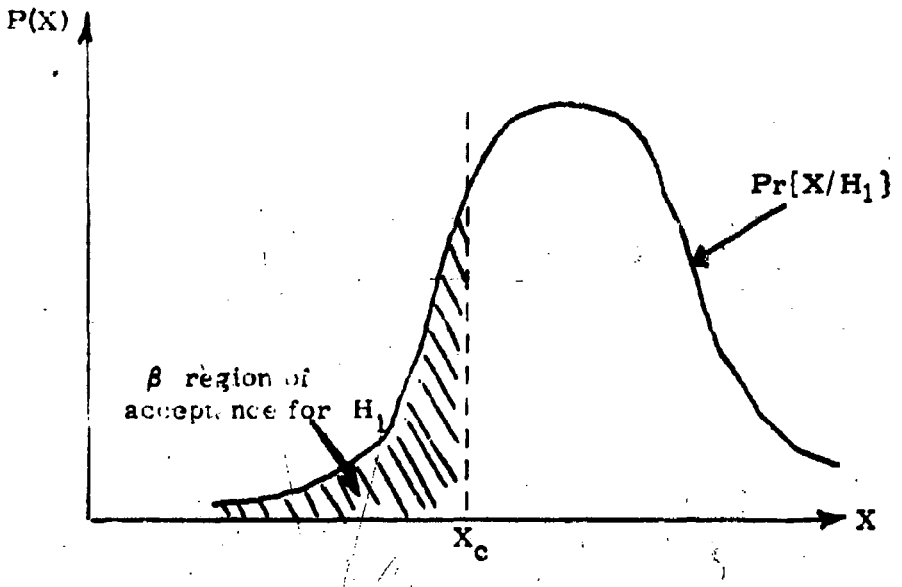


Illustration (b), $Pr\{X/H_1\}$ vs. X

In illustration (a):

$$(18) \qquad \int_{X_c}^{\infty} Pr\{X/H_o\}\, dX = \alpha$$

Therefore, if H_o is true (so that X has the probability distribution $Pr\{X/H_o\}$), the probability of a random observation falling in the critical or rejection region, $X > X_c$ is α,[4] that is: X_c satisfies illustration (a). Now consider H_1 true and X having the density function $Pr\{X/H_1\}$. The probability of a random observation falling in the acceptance region (illustration (b)), $X < X_c$ is β, that is:

(19) $$\int_{-\infty}^{X_c} \Pr\{X/H_1\}\, dX = \beta$$

The probability of correctly rejecting H_o, is called the power of test where:

(20) $$\pi = 1 - \beta = \int_{X_c}^{\infty} \Pr\{X/H_1\}\, dX$$

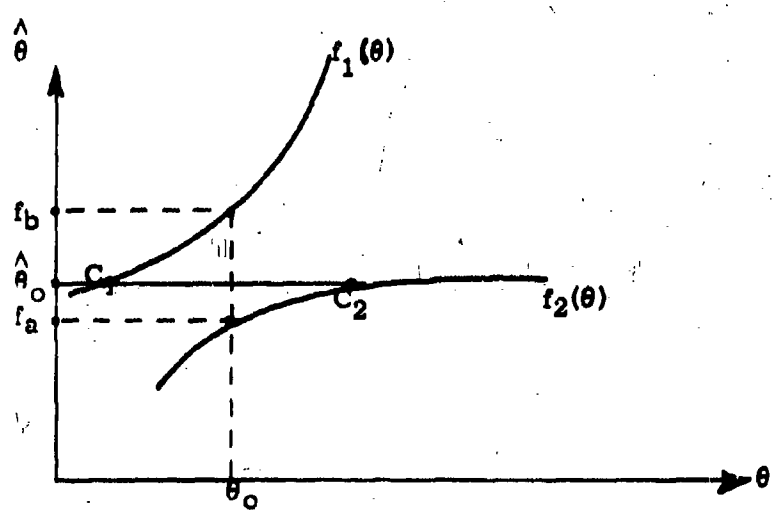
In addition to hypothesis testing, the analysis estimates intervals I and I' for which we can expect, with 99% confidence, that $\bar{d}_i \in I$, and $\frac{S_o^2}{S_v^2} \in I'$.

That is, we utilize the information at each time slot using the t tests described in equations 12, and place a 99% confidence bound on the true state of nature at these points, i.e. $d_i = (x_{o_i} - y_{v_i})$ for E_1 and $\delta_i = (\bar{x}_{o_i} - \bar{y}_{v_i})$ for E_2 respectively. This means that if experiments E_1 and E_2 were to be performed say, 100 times, we could be confident that 99% of such intervals will contain the true state of nature at each time slot. Thus by putting confidence bounds on each set of data points, we would have 85 upper and lower bounds which would generate an envelope. From this envelope we can conclude that for the spectrum of data generated in this experiment, we are 99% confident that the data will be contained with the envelope.

Interval estimation aids in obtaining limits C_1 and C_2 which are functions of the sample values $\{f_c\}$ or functions of the sample values and known population parameters $\{\bar{d}_i\}$, $\{\delta_i\}$ and $\left\{\frac{S_o^2}{S_v^2}\right\}$. The limits are determined so that the probability:[5]

$$\Pr(c_1 < \theta < c_2) \geq 1 - \alpha \tag{21}$$

where θ is the parameter being estimated and $(1-\alpha)$ is the confidence probability. Consider the problem graphically, where $f_1(\theta)$ and $f_2(\theta)$ are drawn so that $c_1 \in f_1(\theta)$, $c_2 \in f_2(\theta)$, $\Pr[f_1(\theta) < \theta < f_2(\theta)] = 1$ and $\hat{\theta}$ is a sufficient estimator of θ obtained from the data.



The line segment (c_1, c_2) will intersect $\theta = \theta_o$ (true value of parameter) if and only if $f_a \leq \hat{\theta}_o \leq f_b$. This is to say that $\Pr(f_a \leq \hat{\theta} \leq f_b) = 1 - \alpha$; this is also the probability that (c_1, c_2) includes θ_o.

It will be shown subsequently that we can be 99% confident, $(1-\alpha = .99)$, that F^* and t^* will be between the calculated upper and lower limits of the confidence interval. In the paired difference test, the probability of accepting H_o:

(22a) $$Pr\left[t[(n-1);\alpha/2] < t^*_1 < t[(n-1);(1-\alpha/2)]\right] = 1-\alpha$$

where $t^*_1 = \dfrac{\bar{d}_i\sqrt{n}}{S_{d_i}}$. From this equation and that of (16) we obtain the confidence interval:

(23a) $$\left[\bar{d}_i \pm t[(n-1);\alpha/2]\left(\frac{S_{\bar{d}_i}}{\sqrt{n}}\right)\right]$$

This means that we can be $100(1-\alpha)\%$ confident that this interval contains $\bar{d}_i = 0$ under H_o. (The critical region is: $t^*[(n-1);\alpha/2] > t^* > t^*[(n-1);(1-\alpha/2)]$). Likewise, for the paired comparison test, the probability of accepting H_o is:

(22b) $$Pr\left[t[n';\alpha/2] < t^*_2 < t[n';(1-\alpha/2)]\right] = 1-\alpha$$

where $$t^*_2 = \frac{\bar{x}_o - \bar{y}_v}{S_p\sqrt{2/n}}$$

This equation and equation (17) lead to the confidence interval:[2]

(23b) $\left[(\bar{x}_o - \bar{y}_v) \pm t[n', \alpha/2]\, S_p\sqrt{2/n}\right]$

The F test for the variances as given in equation (7) can now be rewritten:

$$(24)\qquad \Pr\left[\left(\frac{S_o^2}{S_v^2}\right) F[(n-1),\ (n-1);\ \alpha/2] < \frac{\sigma_o^2}{\sigma_v^2} < \left(\frac{S_o^2}{S_v^2}\right) F[(n-1), (n-1); (1-\alpha/2)]\right] = 1 - \alpha$$

where the confidence interval is:[1]

$$(25)\qquad \left[\left(\frac{S_o^2}{S_v^2}\right)F[(n-1), (n-1); \alpha/2]; \left(\frac{S_o^2}{S_v^2}\right)F[(n-1), (n-1), (1-\alpha/2)]\right]$$

The probability that $\left(\frac{S_o^2}{S_v^2}\right)$ will be contained within this interval, under H_o is $(1-\alpha)$.

CONCLUSIONS:

For E_2, the paired comparison test, the computed value of t used to test the hypothesis $\mu_o = \mu_1$ at 10:00 is 0.0174, (see Figure (16)). The corresponding confidence interval for the same time slot is: I=[-0.9405, 0.9293]. Note that the interval is quite wide considering the small difference between the sample means (for 10:00 hours). Examination of the data, Figures 8, 9, 10, 11, show a marked consistency with this conclusion. The VI measurements (fixed station) are generally smaller than the corresponding value for the OI (mobile station) measurements. Their differences are recorded as: $d_i = f_o - f_v = x_o - y_v$.

The paired comparison test, E_2, Figures 16, 17, 18, 19, requires that two random samples be independent. The data shows, however, that a pair of measurements for VI and OI for any particular time are of approximately the same magnitude. In other words, the variance within the blocks is small compared to the variance between the blocks.

The following data from Figure (9) were taken at 11:30 hours:

Day	x_o	Δx_o	y_v	Δy_v	$x_o - y_v$
1	10.80		10.80		0
2	10.00	0.80	9.80	1.00	0.20
3	9.00	1.00	9.20	0.60	0.20
4	9.50	0.50	9.50	0.30	0
5	10.00	0.50	10.00	0.50	0
6	10.50	0.50	10.50	0.50	0
7	9.50	1.00	9.60	0.90	0.10
8	9.50	0	9.50	0.10	0
9	10.50	1.00	10.40	0.90	0.10

In other words $\Delta x_o, \Delta y_v > x_o - y_v$; as a result of the homogeneity within the blocks, the new experimental design, that of Paired Differences, utilizes the nine difference measurements, di, per time slot to test the hypothesis:

$$H_o: \quad \mu_v = \mu_I \quad \text{vs.} \qquad H_1: \quad \mu_v \neq \mu_I$$

with the t test and confidence intervals as shown in equations 12a, 13a, 14, 23a. This statistical design is a simple example of a randomized block design. The test is commonly called a paired difference test. It is emphasized that the pairing was part of the planning of the experiments, and was not done after the data were collected. Each of the blocks consists of the two observations x_o and y_v for the same day at a specific time. (See Figures 8, 9, 10, and 11.)

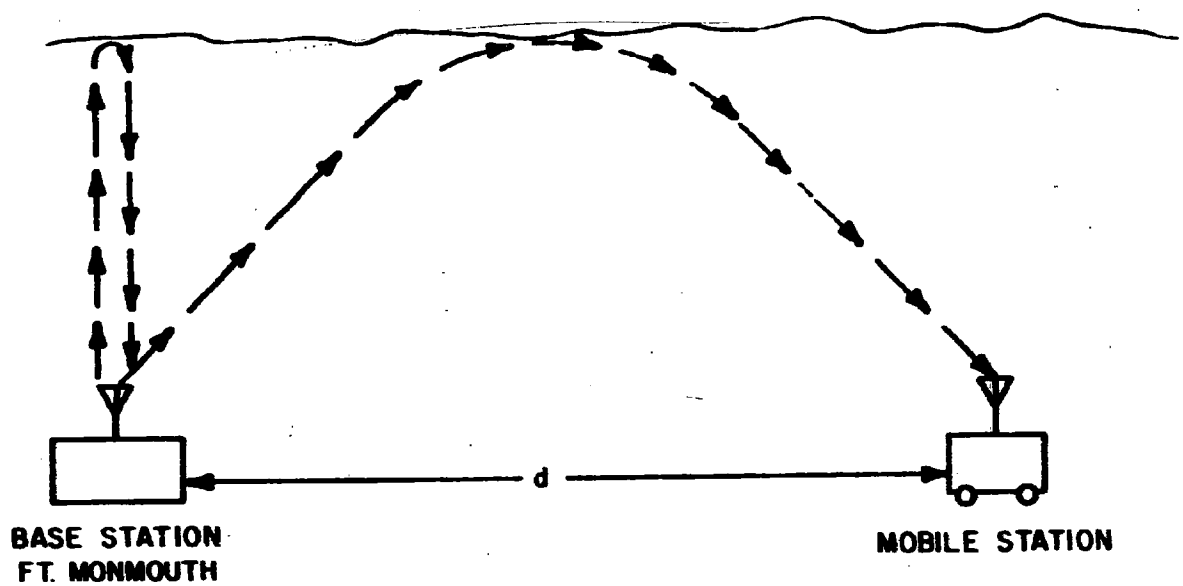
By comparing the computed confidence intervals for the 85 time slots for the paired difference model with those of the unpaired model, see Figures 14, 15, 18, 19 we see a decided gain in information favoring the randomized block design. The gain of information is reflected in the difference in the width of the confidence intervals. Again using data at 10:00 hours, in Figure 20, the interval for the paired comparison test $I_{pc}=(-.9405, .9293)$. The interval for the paired difference test $I_{pd}=(-.1859, .1961)$, and $I_{pd} < I_{pc}$.

The I_{pd_i}, $i=1, \ldots, 85$ are much narrower as a result of blocking in this experiment. Figures 20, 21, 22, and 23 show the comparison of the confidence limits for both methods, as well as a large reduction in the standard deviation S_d as compared to the pooled standard deviation, S_p, of the unpaired observations. Variances are presented graphically in Figures 24, 25.

No appreciable difference exists between the data of the fixed and mobile ionosondes for a distance of 60 Km. The data of the fixed terminal is very nearly identical to those of the mobile terminal for this distance. This means that only one terminal is needed at this distance to provide useful ionospheric data under these given conditions. The result bears out the expectation. Experimentation is planned for investigating critical frequencies at distances beyond 60 Km, (up to 500 Km), to determine the distance within 500 km where the conclusion becomes invalid. This would provide insight as to an extreme distance limit for the usefulness of vertical incidence ionospheric soundings with respect to critical frequency.

Bibliography

1. Ehrenfeld, S. and Littauer, S., (1964), Introduction to Statistical Method, McGraw-Hill.
2. Hald, A., (1952), Statistical Theory with Engineering Applications, J. Wiley.
3. Anderson, R. L. and Bancroft, T. A., (1952), Statistical Theory in Research, McGraw-Hill.
4. Brownlee, K. A., (1960), Statistical Theory and Methodology.
5. Mood, A. M. (1950), Introduction to the Theory of Statistics, McGraw-Hill.



FIELD TESTS FOR A NEAR-REAL TIME
IONOSPHERIC FORCASTING SCHEME

FIG. 1

RAW IONOSONDE DATA

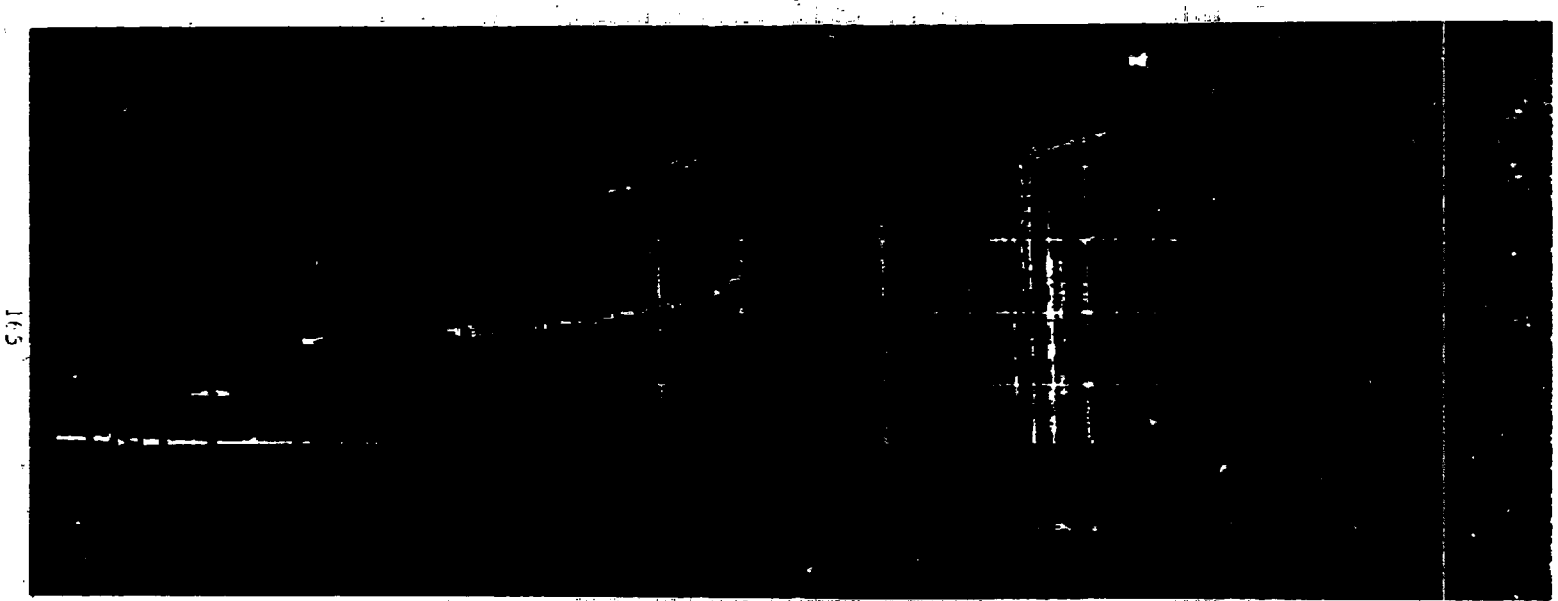


FIG. 2

RAW IONOSONDE DATA

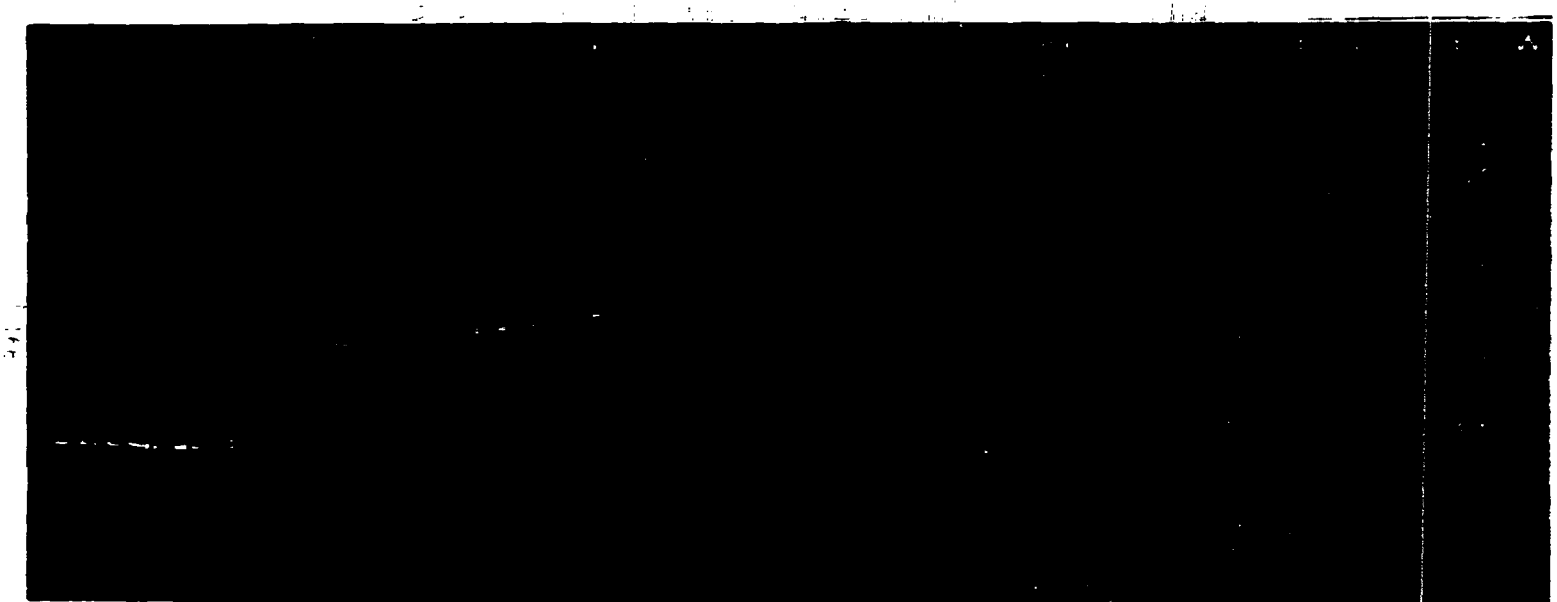


FIG. 3

RAW IONOSCOPE DATA



FIG. 4

RAW IONOSONDE DATA



FIG. 5

TIME OF DAY	ORDERED DAYS OF EXPERIMENTATION									EXPERIMENTAL MEAN	STANDARD DEVIATION
	1	2	3	4	5	6	7	8	9		
0530	$f_{1,1}$	$f_{1,2}$	$f_{1,3}$	$f_{1,4}$	$f_{1,5}$	$f_{1,6}$	$f_{1,7}$	$f_{1,8}$	$f_{1,9}$	$\overline{f}_1$	S_1
0540	$f_{2,1}$	$f_{2,2}$	$f_{2,3}$	$f_{2,4}$	$f_{2,5}$	$f_{2,6}$	$f_{2,7}$	$f_{2,8}$	$f_{2,9}$	$\overline{f}_2$	S_2
0550	$f_{3,1}$	$f_{3,2}$	$f_{3,3}$	$f_{3,4}$	$f_{3,5}$	$f_{3,6}$	$f_{3,7}$	$f_{3,8}$	$f_{3,9}$	$\overline{f}_3$	S_3
0600	$f_{4,1}$	$f_{4,2}$	$f_{4,3}$	$f_{4,4}$	$f_{4,5}$	$f_{4,6}$	$f_{4,7}$	$f_{4,8}$	$f_{4,9}$	$\overline{f}_4$	S_4
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1910	$f_{83,1}$	$f_{83,2}$	$f_{83,3}$	$f_{83,4}$	$f_{83,5}$	$f_{83,6}$	$f_{83,7}$	$f_{83,8}$	$f_{83,9}$	$\overline{f}_{83}$	S_{83}
1920	$f_{84,1}$	$f_{84,2}$	$f_{84,3}$	$f_{84,4}$	$f_{84,5}$	$f_{84,6}$	$f_{84,7}$	$f_{84,8}$	$f_{84,9}$	$\overline{f}_{84}$	S_{84}
1930	$f_{85,1}$	$f_{85,2}$	$f_{85,3}$	$f_{85,4}$	$f_{85,5}$	$f_{85,6}$	$f_{85,7}$	$f_{85,8}$	$f_{85,9}$	$\overline{f}_{85}$	S_{85}

1. $\overline{f}_i = \frac{1}{n} \sum_{j=1}^{9} f_{ij} \qquad i = 1, 2, \quad 85 \; ; \; j = 1, 2, \quad , 9$

2. $S_i^2 = \frac{1}{n-1} \sum_{j=1}^{9} (f_{ij} - \overline{f}_i)^2 \qquad i = 1, 2, \quad 85 \; ; \; j = 1, 2, \quad , 9$

FIG. 6 PRELIMINARY DATA SCHEME – PAIRED COMARISON

TIME OF DAY	ORDERED DAYS OF EXPERIMENTATION									EXPERIMENTAL MEAN	STANDARD DEVIATION
	1	2	3	4	5	6	7	8	9		
0530	$d_{1,1}$	$d_{1,2}$	$d_{1,3}$	$d_{1,4}$	$d_{1,5}$	$d_{1,6}$	$d_{1,7}$	$d_{1,8}$	$d_{1,9}$	$\bar{d}_1$	S_{d1}
0540	$d_{2,1}$	$d_{2,2}$	$d_{2,3}$	$d_{2,4}$	$d_{2,5}$	$d_{2,6}$	$d_{2,7}$	$d_{2,8}$	$d_{2,9}$	$\bar{d}_2$	S_{d2}
0550	$d_{3,1}$	$d_{3,2}$	$d_{3,3}$	$d_{3,4}$	$d_{3,5}$	$d_{3,6}$	$d_{3,7}$	$d_{3,8}$	$d_{3,9}$	$\bar{d}_3$	S_{d3}
0600	$d_{4,1}$	$d_{4,2}$	$d_{4,3}$	$d_{4,4}$	$d_{4,5}$	$d_{4,6}$	$d_{4,7}$	$d_{4,8}$	$d_{4,9}$	$\bar{d}_4$	S_{d4}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1910	$d_{83,1}$	$d_{83,2}$	$d_{83,3}$	$d_{83,4}$	$d_{83,5}$	$d_{83,6}$	$d_{83,7}$	$d_{83,8}$	$d_{83,9}$	$\bar{d}_{83}$	S_{d83}
1920	$d_{84,1}$	$d_{84,2}$	$d_{84,3}$	$d_{84,4}$	$d_{84,5}$	$d_{84,6}$	$d_{84,7}$	$d_{84,8}$	$d_{84,9}$	$\bar{d}_{84}$	S_{d84}
1930	$d_{85,1}$	$d_{85,2}$	$d_{85,3}$	$d_{85,4}$	$d_{85,5}$	$d_{85,6}$	$d_{85,7}$	$d_{85,8}$	$d_{85,9}$	$\bar{d}_{85}$	S_{d85}

1. $d_{ij} = (f_{o1} - f_{v1})_{ij}$ $\quad i = 1,2,3, \cdots 85$; $j = 1,2, \cdots 9$

2. $\bar{d}_i = \frac{1}{n} \sum_{j=1}^{9} d_{ij}$ $\quad i = 1,2,3 \cdots 85$

3. $S^2_{di} = \frac{1}{n-1} \sum_{j=1}^{9} (d_{ij} - \bar{d}_i)^2$ $\quad i = 1,2,3 \cdots 85$; $i = 1,2, \cdots 9$

FIG. 7 PRELIMINARY DATA SCHEME-PAIRED DIFFERENCE

REDUCED IONOSONDE DATA

TIME OF DAY (24-HOUR CLOCK)	CRITICAL FREQUENCY (VERTICAL INCIDENCE)	CRITICAL FREQUENCY (OBLIQUE INCIDENCE)	DIFFERENCE BETWEEN (VERTICAL AND OBLIQUE
430.	3.850	3.950	-0.100
430.	3.500	3.600	-0.100
430.	3.700	3.700	0.000
430.	4.700	4.700	0.000
430.	3.600	3.800	-0.200
430.	3.850	3.950	-0.100
430.	3.850	3.950	-0.100
430.	3.850	3.950	-0.100
430.	3.850	3.950	-0.100
440.	3.400	3.500	-0.100
440.	3.500	3.600	-0.100
440.	3.900	4.600	-1.600
440.	4.700	3.600	+1.100
440.	3.400	3.500	-0.100
440.	3.400	2.600	+0.800
440.	2.500	3.500	-1.000
440.	3.200	3.500	-0.300
440.	3.350	3.200	+0.150
550.	3.500	3.500	0.000
550.	3.700	3.500	+0.200
550.	3.500	3.400	+0.100
550.	4.400	4.400	0.000
550.	3.600	3.600	0.000
550.	3.500	3.500	0.000
550.	2.550	2.600	-0.050
550.	3.550	3.500	+0.050
550.	3.100	3.200	-0.100
400.	3.300	3.500	-0.200
400.	2.800	3.500	-0.700
400.	3.400	3.400	0.000
400.	4.300	4.300	0.000
400.	3.600	3.600	0.000
400.	3.300	3.500	-0.200
400.	2.500	3.100	-0.600
400.	3.300	3.500	-0.200
400.	3.100	3.200	-0.100
410.	3.400	3.300	+0.100
510.	3.000	2.800	+0.200
410.	3.400	3.400	0.000
410.	4.200	4.000	+0.200
410.	3.700	3.700	0.000
410.	3.400	3.300	+0.100
410.	3.000	2.600	+0.400
410.	3.400	3.300	+0.100
410.	3.200	3.100	+0.100
420.	3.500	3.100	+0.400
420.	3.000	2.700	+0.300
420.	3.300	3.300	0.000
420.	4.300	3.100	+1.200
420.	3.400	3.700	-0.300
420.	3.500	3.100	+0.400

FIG. 8

REDUCED IONOSONDE DATA

TIME OF DAY (GREENWICH CLOCK)	CRITICAL FREQUENCY VERTICAL INCIDENCE	CRITICAL FREQUENCY OBLIQUE INCIDENCE	DIFFERENCE BETWEEN VERTICAL AND OBLIQUE
1110.	10.400	10.600	0.200
1110.	9.800	9.600	-0.200
1110.	9.000	8.600	-0.400
1110.	10.500	10.200	-0.300
1110.	10.700	10.800	0.100
1110.	11.200	11.200	0.000
1110.	9.500	9.500	0.000
1110.	9.200	9.000	-0.200
1110.	10.200	10.400	0.200
1120.	10.800	10.800	0.000
1120.	10.200	9.200	-1.000
1120.	9.600	9.600	0.000
1120.	10.000	10.200	0.200
1120.	10.400	10.500	0.100
1120.	10.800	11.000	0.200
1120.	9.700	9.500	-0.200
1120.	9.500	9.500	0.000
1120.	10.200	10.000	-0.200
1130.	10.800	10.800	0.000
1130.	9.800	10.000	0.200
1130.	9.200	9.000	-0.200
1130.	9.500	9.500	0.000
1130.	10.000	10.000	0.000
1130.	10.500	10.500	0.000
1130.	9.600	9.500	-0.100
1130.	9.500	9.500	0.000
1130.	10.400	10.500	0.100
1140.	10.600	10.800	0.200
1140.	9.800	9.800	0.000
1140.	8.800	8.800	0.000
1140.	9.300	9.300	0.000
1140.	9.850	10.200	0.350
1140.	10.800	10.800	0.000
1140.	9.500	9.600	0.100
1140.	9.600	9.700	0.100
1140.	10.300	10.400	0.100
1150.	10.200	10.200	0.000
1150.	9.700	9.800	0.100
1150.	9.200	9.000	-0.200
1150.	9.300	9.500	0.200
1150.	10.500	10.500	0.000
1150.	10.800	10.700	-0.100
1150.	9.500	9.300	-0.200
1150.	9.500	9.500	0.000
1150.	10.200	10.200	0.000
1200.	10.000	8.900	-1.100
1200.	9.200	9.200	0.000
1200.	8.800	9.200	0.400
1200.	10.200	9.400	-0.800
1200.	10.600	10.500	-0.100
1200.	10.800	10.800	0.000

FIG. 9



REDUCED IONOSONDE DATA

TIME OF DAY (24 HOUR CLOCK)	CRITICAL FREQUENCY VERTICAL INCIDENCE	CRITICAL FREQUENCY OBLIQUE INCIDENCE	DIFFERENCE BETWEEN VERTICAL AND OBLIQUE
1450.	9.200	9.000	-0.200
1490.	9.400	9.500	0.100
1450.	10.400	10.800	0.400
1500.	10.900	10.200	0.300
1500.	9.000	9.000	0.000
1500.	9.900	10.000	0.100
1500.	9.000	9.000	0.000
1500.	11.000	11.400	0.400
1500.	11.000	11.000	0.000
1500.	9.400	9.400	-0.200
1500.	9.900	9.600	0.600
1500.	10.400	10.800	0.000
1510.	9.400	10.400	0.400
1510.	9.400	9.400	-0.200
1510.	10.900	10.850	0.000
1510.	9.000	9.000	0.000
1510.	10.800	11.000	0.200
1510.	11.900	11.200	0.000
1510.	9.400	9.000	-0.600
1510.	9.900	9.800	0.000
1510.	10.400	10.400	0.000
1520.	10.000	10.000	0.000
1520.	9.000	9.600	0.600
1520.	10.000	10.000	0.000
1520.	9.600	9.000	-0.600
1520.	11.400	10.400	-1.000
1520.	11.900	11.200	0.000
1520.	9.200	9.500	0.300
1520.	10.000	10.000	0.000
1520.	10.900	10.400	0.200
1530.	9.900	10.000	0.200
1530.	9.000	9.600	0.600
1530.	9.950	9.950	0.000
1530.	9.400	9.000	-0.400
1530.	10.400	10.600	0.200
1530.	11.400	11.400	0.200
1530.	9.400	9.100	-0.300
1530.	10.000	10.000	0.000
1530.	9.800	10.000	0.200
1540.	9.600	9.600	0.000
1540.	9.800	9.000	-0.800
1540.	10.200	10.000	-0.200
1540.	9.400	9.400	-0.200
1540.	10.800	10.600	-0.200
1540.	11.400	11.200	-0.200
1540.	10.200	10.000	-0.200
1540.	10.400	9.800	-0.200
1540.	10.200	10.200	0.000
1550.	9.800	9.800	0.000
1550.	9.400	9.600	0.000
1550.	9.900	9.900	0.000

FIG. 10



REDUCED IONOSONDE DATA

TYPE OF DAY (24 - 08 CLOCK)	CRITICAL FREQUENCY VERTICAL INCIDENCE	CRITICAL FREQUENCY OBLIQUE INCIDENCE	DIFFERENCE BETWEEN VERTICAL AND OBLIQUE
1940.	6.400	6.500	-0.100
1940.	8.100	8.300	-0.200
1940.	7.600	7.800	-0.100
1940.	7.700	7.500	-0.200
1940.	6.800	6.800	0.000
1940.	7.400	7.200	-0.200
1950.	7.100	6.900	-0.200
1950.	6.400	6.200	-0.200
1950.	7.000	6.400	-0.600
1950.	6.800	6.800	0.000
1950.	8.600	8.600	0.000
1950.	7.400	6.400	-1.000
1950.	7.300	7.400	0.100
1950.	6.400	6.400	0.000
1950.	7.200	7.200	0.000
1900.	6.800	6.800	0.000
1900.	5.700	5.700	0.000
1900.	6.500	6.600	0.100
1900.	6.600	6.800	0.200
1900.	8.200	8.400	0.200
1900.	7.500	7.400	-0.100
1900.	6.500	6.600	0.100
1900.	6.200	6.000	-0.200
1900.	7.000	7.000	0.000
1910.	6.450	6.450	0.000
1910.	5.300	5.400	0.100
1910.	5.600	5.600	0.000
1910.	6.600	6.600	0.000
1910.	7.600	7.800	0.200
1910.	7.400	7.400	0.000
1910.	6.500	6.000	-0.500
1910.	6.200	6.200	0.000
1910.	6.450	6.450	0.000
1920.	6.100	6.050	-0.050
1920.	5.000	5.200	0.200
1920.	5.500	5.400	-0.100
1920.	6.400	6.400	0.000
1920.	7.000	6.800	-0.200
1920.	6.600	6.600	-0.200
1920.	6.000	6.000	0.000
1920.	5.900	5.900	0.000
1920.	6.100	6.050	-0.050
1930.	5.600	5.900	0.300
1930.	4.800	4.800	0.000
1930.	5.400	5.000	-0.400
1930.	5.800	5.900	0.100
1930.	6.200	6.500	0.300
1930.	5.400	7.500	0.900
1930.	5.800	5.800	0.000
1930.	5.600	5.900	0.300
1930.	5.600	5.900	0.300

FIG. 11

PRELIMINARY CALCULATIONS

TIME OF DAY	MEAN CRITICAL FREQUENCY VI=VERTICAL INCIDENCE	MEAN CRITICAL FREQUENCY OI=OBLIQUE INCIDENCE	MEAN OF DIFFERENCES BETWEEN OI AND VI CRITICALS	ESTIMATE OF VARIANCE VI	ESTIMATE OF VARIANCE OI	ESTIMATE OF VARIANCE OF DIFFERENCES	STANDARD DEVIATION VI	STANDARD DEVIATION OI	STANDARD DEVIATION OF DIFFERENCES
530.	3.441	3.350	-0.0889	0.116	0.096	0.004	0.341	0.310	0.060
540.	3.406	3.511	0.1056	0.344	0.266	0.682	0.587	0.516	0.826
550.	3.511	3.489	-0.0222	0.290	0.296	0.008	0.539	0.544	0.087
600.	3.240	3.511	0.2222	0.258	0.114	0.087	0.508	0.337	0.259
610.	3.411	3.278	-0.1333	0.136	0.179	0.015	0.369	0.424	0.122
620.	3.489	3.389	-0.0000	0.129	0.116	0.197	0.359	0.341	0.444
630.	3.111	3.189	0.0778	0.237	0.297	0.097	0.487	0.545	0.311
640.	3.417	3.378	-0.0389	0.050	0.102	0.059	0.224	0.319	0.229
650.	3.933	3.456	-0.1833	0.118	0.165	0.032	0.344	0.324	0.180
700.	4.328	4.211	-0.1167	0.062	0.104	0.104	0.249	0.322	0.322
710.	5.300	5.300	-0.0000	0.165	0.072	0.047	0.301	0.269	0.218
720.	5.422	5.533	0.0111	0.154	0.168	0.271	0.393	0.384	0.521
730.	5.878	5.922	0.0444	0.268	0.094	0.104	0.518	0.307	0.324
740.	6.511	6.478	-0.0333	0.096	0.097	0.013	0.310	0.295	0.112
750.	6.289	6.456	0.1389	0.046	0.117	0.072	0.215	0.342	0.268
800.	7.289	7.111	-0.1778	0.109	0.109	0.099	0.330	0.330	0.315
810.	7.411	7.433	0.0222	0.079	0.068	0.018	0.280	0.260	0.135
820.	7.444	7.389	-0.0556	0.170	0.169	0.027	0.412	0.411	0.165
830.	7.378	7.511	0.1333	0.284	0.116	0.148	0.533	0.341	0.384
840.	7.700	7.700	-0.0000	0.925	0.273	0.028	0.474	0.520	0.159
850.	8.156	8.267	0.0111	0.225	0.145	0.026	0.475	0.381	0.162
900.	8.100	8.178	0.0778	0.183	0.349	0.049	0.427	0.591	0.208
910.	8.389	8.433	0.0444	0.479	0.487	0.169	0.692	0.698	0.410
920.	8.544	8.633	-0.0111	0.388	0.515	0.026	0.623	0.719	0.169
930.	8.933	8.956	0.0222	0.369	0.333	0.014	0.604	0.577	0.120
940.	9.122	9.156	0.1333	0.149	0.320	0.070	0.387	0.566	0.265
950.	9.256	9.211	-0.0389	0.300	0.374	0.027	0.548	0.611	0.165
1000.	9.328	9.333	0.0056	0.397	0.525	0.029	0.630	0.725	0.170
1010.	9.539	9.522	-0.0056	0.269	0.459	0.040	0.519	0.678	0.201
1020.	9.578	9.650	0.0722	0.212	0.207	0.027	0.460	0.456	0.164
1030.	9.600	9.522	-0.0778	0.128	0.112	0.012	0.357	0.335	0.109
1040.	9.789	9.650	-0.1389	0.264	0.342	0.115	0.513	0.602	0.339
1050.	9.933	9.922	-0.0111	0.535	0.487	0.019	0.731	0.669	0.139
1100.	9.989	9.967	-0.0222	0.634	0.776	0.029	0.796	0.877	0.173
1110.	10.167	10.011	-0.0556	0.585	0.786	0.055	0.765	0.887	0.235
1120.	10.133	10.033	-0.1000	0.233	0.403	0.135	0.482	0.634	0.367
1130.	9.922	9.922	0.0000	0.992	0.354	0.012	0.566	0.595	0.112
1140.	9.939	9.933	-0.0056	0.407	0.459	0.014	0.638	0.676	0.118
1150.	9.478	9.856	0.0722	0.319	0.333	0.017	0.565	0.577	0.133
1200.	9.467	9.522	-0.1444	0.498	0.472	0.235	0.705	0.687	0.485
1210.	9.422	9.600	0.0722	0.432	0.460	0.017	0.657	0.678	0.131
1220.	9.456	9.567	0.0889	0.895	0.948	0.054	0.946	0.974	0.237
1230.	9.567	9.511	-0.0444	0.725	0.719	0.013	0.851	0.848	0.113
1240.	9.489	9.756	0.0667	0.811	0.928	0.180	0.901	0.963	0.400
1250.	9.900	9.733	-0.1333	1.100	1.032	0.050	1.049	1.015	0.224

FIG. 12

PRELIMINARY CALCULATIONS

TIME OF DAY	[illegible]	[illegible]	MEAN OF DIFFERENCES BETWEEN ST AND AT CRITICALS	ESTIMATE OF VARIANCE VI	OI	OF DIFFERENCES	STANDARD DEVIATION VI	OI	OF DIFFERENCES
130	9.722	9.689	-0.0333	0.979	0.741	0.020	0.938	0.861	.141
131	9.178	9.211	0.0667	0.934	1.156	0.030	0.967	1.003	.173
132	9.244	9.300	0.0556	1.218	1.230	0.028	1.104	1.109	0.167
133	10.122	9.989	-0.1333	0.904	1.031	0.060	0.951	1.015	0.245
134	9.344	10.122	0.0333	1.101	0.904	0.235	1.049	0.951	0.485
135	10.17	10.122	-0.0778	1.030	1.188	0.139	1.015	1.032	0.373
140	10.189	10.422	0.0500	1.104	0.874	0.098	1.051	0.935	0.312
141	10.189	10.378	0.0556	0.779	0.694	0.038	0.883	0.946	0.194
142	10.17	10.356	0.0556	0.870	0.918	0.038	0.933	0.958	0.194
143	10.[illegible]	10.156	0.0444	0.603	0.717	0.089	0.776	0.847	0.298
144	10.172	10.156	-0.0222	0.564	0.678	0.054	0.751	0.823	0.233
145	10.[illegible]	10.[illegible]	-0.1000	0.418	0.657	0.030	0.646	0.811	0.173
150	9.722	10.[illegible]29	0.3000	0.689	0.734	0.060	0.830	0.857	0.245
151	10.[illegible]	10.[illegible]5	0.0000	0.437	0.617	0.100	0.661	0.786	0.316
152	10.[illegible]7	10.[illegible]11	-0.0556	0.650	0.396	0.228	0.806	0.629	0.477
153	9.[illegible]	9.989	0.0111	0.487	0.535	0.104	0.698	0.731	0.326
154	10.322	10.[illegible]00	-0.2222	0.314	0.410	0.054	0.561	0.640	0.233
155	9.[illegible]	9.911	0.0111	0.276	0.378	0.051	0.526	0.611	0.226
160	9.722	9.65[illegible]	-0.0722	0.299	0.348	0.029	0.547	0.589	0.172
161	9.633	9.65[illegible]	-0.0778	0.274	0.327	0.054	0.524	0.572	0.233
162	9.388	9.411	0.0667	0.268	0.276	0.023	0.517	0.525	0.150
163	9.156	9.100	-0.0556	0.628	0.525	0.098	0.792	0.725	0.313
164	8.922	[illegible].933	-0.0889	0.789	0.525	0.124	0.877	0.725	0.352
165	8.756	8.678	-0.0722	0.908	0.807	0.042	0.899	0.947	0.205
170	[illegible].556	[illegible].489	-0.0611	0.898	0.879	0.093	0.947	0.937	0.304
171	8.344	[illegible].32[illegible]	-0.0389	0.494	0.812	0.049	0.703	0.782	0.204
172	8.200	[illegible].25[illegible]	-0.0500	0.418	0.370	0.060	0.646	0.608	0.245
173	8.[illegible]11	[illegible].039	-0.0278	0.296	0.370	0.017	0.544	0.608	0.130
174	7.[illegible]	7.489	-0.0111	0.526	0.489	0.011	0.726	0.699	0.105
175	7.522	7.[illegible]79	-0.0500	0.752	0.934	0.045	0.867	0.967	0.212
180	7.[illegible]	7.75[illegible]	-0.1000	1.093	0.852	0.110	1.045	0.808	0.332
181	7.741	7.761	-0.0000	0.737	0.852	0.017	0.859	0.923	0.132
182	7.[illegible]	7.70[illegible]	0.0611	0.675	0.893	0.045	0.821	0.945	0.212
183	7.322	7.37[illegible]	-0.0556	0.564	0.572	0.008	0.751	0.756	0.088
184	7.[illegible]	7.189	-0.1000	0.601	0.509	0.007	0.775	0.713	0.087
185	7.[illegible]	8.[illegible]99	-0.1667	0.524	0.559	0.15[illegible]	0.724	0.748	0.387
190	8.[illegible]	[illegible].411	0.0333	0.534	0.611	0.014	0.731	0.782	0.132
191	8.[illegible]	8.433	-0.0222	0.585	0.604	0.037	0.738	0.777	0.192
192	8.[illegible]	[illegible].[illegible]	-0.0889	0.376	0.270	0.019	0.613	0.519	0.121
193	5.[illegible]	5.611	0.3111	0.140	0.619	[illegible]	0.374	0.787	0.039

FIG. 13

PAIRED DIFFERENCE TEST

TIME OF DAY	CONFIDENCE INTERVALS OF PAIRED DIFFERENCES LOWER LIMIT	MEAN VALUES OF DIFFERENCES	UPPER LIMIT	MEAN VALUES OF DIFFERENCES BETWEEN VI AND CI	T-TEST	STANDARD DEVIATION OF DIFFERENCES BETWEEN VI AND CI
530.	-0.0217	-0.[illegible]	0.1561	0.069	0.183[illegible]	.0[illegible]0252
540.	-0.0127	-0.[illegible]54	1.0288	0.106	0.3016	.82554[illegible]
550.	-0.1195	-0.[illegible]222	0.0751	-0.022	0.7224	.0[illegible]0255
600.	-0.0471	0.2222	0.5118	0.222	2.4293	.25873424
610.	-0.2703	-0.1333	0.0836	-0.133	3.0792	.12247849
620.	-0.[illegible]07[illegible]	-0.[illegible]0[illegible]	0.[illegible]073	-0.[illegible]00	2.[illegible]458	.[illegible]444[illegible]972
630.	-0.2704	0.[illegible]778	0.4260	0.[illegible]78	0.7645	.31135[illegible]3
640.	-0.2048	-0.[illegible]389	0.2170	-0.[illegible]39	0.4807	.27002[illegible]5[illegible]
650.	-0.3669	-0.1[illegible]33	0.[illegible]1[illegible]3	-0.1[illegible]3	2.[illegible]744	.1[illegible]627756
700.	-0.4769	-0.1167	0.2436	-0.117	1.9285	.3721[illegible]247
710.	-0.2437	-0.[illegible]00[illegible]	0.2437	-0.[illegible]00	0.0000	.2179[illegible]49[illegible]
720.	-0.5712	0.[illegible]111	0.5934	0.[illegible]11	0.[illegible]696	.52[illegible]68331
730.	-0.3227	0.[illegible]446	0.4116	0.[illegible]44	0.3829	.32[illegible]29[illegible]2[illegible]
740.	-0.1584	-0.[illegible]333	0.[illegible]917	-0.[illegible]33	0.9433	.11[illegible]80[illegible]80
750.	-0.4384	-0.1389	0.1606	-0.139	1.4667	.26783[illegible]9[illegible]
800.	-0.5308	-0.177[illegible]	0.1749	-0.[illegible]76	1.5985	.31534[illegible]13
810.	-0.1122	0.[illegible]222	0.1566	0.[illegible]22	0.5230	.12[illegible]8[illegible]4
820.	-0.2398	-0.[illegible]55[illegible]	0.12[illegible]7	-0.[illegible]56	0.9536	.[illegible]447[illegible]1[illegible]
830.	-0.2962	0.1333	0.5628	0.133	0.9819	.3[illegible]8[illegible]5729
840.	-0.1768	-0.[illegible]00[illegible]	0.1768	-0.[illegible]00	0.[illegible]000	.1[illegible]5[illegible]13[illegible]
850.	-0.1494	0.[illegible]111	0.1918	0.[illegible]11	0.1045	.10[illegible]5[illegible]933
900.	-0.1913	0.[illegible]77[illegible]	0.3068	0.[illegible]7[illegible]	1.0741	.2[illegible]8[illegible]343
910.	-0.4134	0.[illegible]44[illegible]	0.5[illegible]25	0.[illegible]44[illegible]	0.3659	.4[illegible]96[illegible]4[illegible]6
920.	-0.1919	-0.[illegible]111	0.1896	-0.[illegible]11	0.1985	.10[illegible]5[illegible]933
930.	-0.1122	0.[illegible]222	0.1566	0.[illegible]22	0.5230	.12[illegible]8[illegible]4
940.	-0.1625	0.1333	0.[illegible]292	0.133	1.[illegible]254	.2[illegible]65[illegible]13
950.	-0.2239	-0.[illegible]38[illegible]	0.1461	-0.[illegible]39	0.[illegible]465[illegible]	.1[illegible]5[illegible]1[illegible]4
1000.	-0.1[illegible]50	0.[illegible]5[illegible]	0.1961	0.[illegible]06[illegible]	0.[illegible]972	.17[illegible]375[illegible]
1010.	-0.2380	-0.[illegible]5[illegible]	0.2[illegible]9	-0.[illegible]06[illegible]	0.[illegible]753	.2[illegible]0[illegible]69328
1020.	-0.1113	0.[illegible]722	0.2554	0.[illegible]72	1.[illegible]2445	.1[illegible]4[illegible]1[illegible]43
1030.	-0.2000	-0.[illegible]77[illegible]	0.[illegible]444	-0.[illegible]72	2.[illegible]129	.1[illegible]926[illegible]4[illegible]
1040.	-0.5179	-0.1389	0.2461	-0.139	1.1591	.3389[illegible]1[illegible]
1050.	-0.1437	-0.[illegible]11[illegible]	0.1415	-0.[illegible]11	0.2304	.1384[illegible]2[illegible]
1100.	-0.2141	-0.[illegible]22[illegible]	0.1697	-0.[illegible]22	0.3643	.17[illegible]5[illegible]3[illegible]
1110.	-0.3185	-0.[illegible]55[illegible]	0.2078	-0.[illegible]56	0.6643	.2351[illegible]727
1120.	-0.[illegible]09	-0.[illegible]70[illegible]	0.3129	-0.[illegible]1[illegible]	0.[illegible]7[illegible]8	.3[illegible]7[illegible]2[illegible]4
1130.	-0.1250	0.[illegible]000	0.1250	0.[illegible]00	0.[illegible]000	.1[illegible]1[illegible]36[illegible]
1140.	-0.[illegible]80	-0.[illegible]44[illegible]	0.[illegible]24[illegible]	0.[illegible]6[illegible]	2.[illegible]556	.1[illegible]0[illegible]3[illegible]0
1150.	-0.1478	-0.[illegible]22[illegible]	0.[illegible]123[illegible]	-0.[illegible]22	0.[illegible]48[illegible]	.1[illegible]6[illegible]7[illegible]3
1200.	-0.[illegible]469	-0.[illegible]46[illegible]	0.[illegible]38[illegible]	-0.[illegible]46[illegible]	0.[illegible]8[illegible]3	.[illegible]45[illegible]5[illegible]1
1210.	-0.1478	-0.[illegible]22[illegible]	0.[illegible]123[illegible]	-0.[illegible]22	0.[illegible]48[illegible]	.13[illegible]7[illegible]3
1220.	-0.[illegible]53[illegible]	-0.[illegible]88[illegible]	0.[illegible]176[illegible]	-0.[illegible]89	1.[illegible]414	.23[illegible]7[illegible]4
1230.	-0.[illegible]20[illegible]	0.[illegible]44[illegible]	0.[illegible]76[illegible]	0.[illegible]44	1.[illegible]121	.113[illegible]3
1240.	-0.3407	0.[illegible]467	0.[illegible]5140	0.[illegible]67	0.[illegible]871[illegible]	.[illegible]0[illegible]0[illegible]0
1250.	-0.1167	0.1133	0.[illegible]3834	0.133	1.[illegible]455	.[illegible]234[illegible]4[illegible]

FIG. 14

PAIRED DIFFERENCE TEST

PAIRED COMPARISON TEST

TIME OF DAY	ESTIMATED VARIANCES OF CRITICAL FREQUENCIES: VERTICAL INCIDENCE	OBLIQUE INCIDENCE	PAIRED VARIANCE	MEAN CRITICAL FREQUENCIES: OBLIQUE INCIDENCE	VERTICAL INCIDENCE	DIFFERENCE	F RATIO OF VARIANCES	T-TEST
530.	0.1141	0.1952	0.1062	3.761	3.650	0.109	1.7043	0.5787
540.	0.3481	0.2661	0.3051	3.406	3.511	-0.106	1.2929	-0.6054
550.	0.2905	0.2661	0.2783	3.511	3.389	0.122	1.0918	0.6894
600.	0.2536	0.1136	0.1836	3.289	3.511	-0.222	2.2323	-1.1001
610.	0.1361	0.1794	0.1478	3.411	3.278	0.133	0.7585	0.7121
620.	0.1286	0.1141	0.1228	3.489	3.089	0.400	1.1077	2.4257
630.	0.2367	0.2967	0.2667	3.111	3.189	-0.078	0.7978	-0.3195
640.	0.0500	0.1019	0.0760	3.417	3.378	0.039	0.4905	0.2993
650.	0.1181	0.1059	0.1116	3.833	3.650	0.183	1.1250	1.1644
700.	0.0619	0.1036	0.0828	4.328	4.211	0.117	0.5979	0.8602
710.	0.1451	0.0725	0.1087	5.000	5.000	0.000	2.0000	0.0000
720.	0.1544	0.1475	0.1519	5.022	5.033	-0.011	1.0471	-0.0867
730.	0.2844	0.0948	0.1704	5.878	5.922	-0.044	2.0000	-0.2220
740.	0.0961	0.0869	0.0915	6.511	6.478	0.033	1.1058	0.2337
750.	0.0461	0.1169	0.0815	6.989	6.850	0.139	0.3945	1.0321
800.	0.1085	0.1086	0.1084	7.289	7.111	0.178	1.1000	1.1443
810.	0.0784	0.0675	0.0731	7.411	7.433	-0.022	1.1615	-0.1744
820.	0.1697	0.1686	0.1691	7.444	7.389	0.056	1.0042	0.2866
830.	0.2844	0.1161	0.2003	7.378	7.511	-0.133	2.4498	-0.8390
840.	0.2251	0.2700	0.2475	7.700	7.700	0.000	0.8333	0.0000
850.	0.2253	0.1453	0.1851	8.056	8.167	-0.011	1.5536	-0.0548
900.	0.1825	0.3494	0.2660	8.100	8.178	-0.078	0.5223	-0.3199
910.	0.4785	0.4895	0.4831	8.389	8.433	-0.044	0.9818	-0.1357
920.	0.3878	0.5150	0.4514	8.444	8.433	0.011	0.7530	0.0351
930.	0.3651	0.3324	0.3489	8.933	8.956	-0.022	1.0949	-0.0798
940.	0.1494	0.3203	0.2349	9.022	9.156	-0.133	0.4646	-0.5936
950.	0.3003	0.3736	0.3369	9.250	9.211	0.039	1.0630	0.1421
1000.	0.3963	0.5253	0.4610	9.328	9.333	-0.006	0.7551	-0.0174
1010.	0.2694	0.4598	0.3644	9.528	9.522	0.006	0.5865	0.0195
1020.	0.2119	0.2075	0.2097	9.578	9.550	0.028	1.0214	0.1345
1030.	0.1275	0.1119	0.1197	9.600	9.522	0.078	1.1390	0.4749
1040.	0.2636	0.3625	0.3131	9.789	9.650	0.139	0.7272	0.5246
1050.	0.5351	0.4469	0.4910	9.933	9.922	0.011	1.1970	0.0336
1100.	0.4334	0.7704	0.7019	9.980	9.957	0.022	1.0229	0.0543
1110.	0.485	0.7961	0.6656	10.167	10.011	0.156	1.0693	0.1655
1120.	0.2325	0.4025	0.3175	10.133	10.133	0.000	1.5776	0.0000
1130.	0.2811	0.3584	0.3232	9.922	9.922	0.000	1.3237	0.0000
1140.	0.4074	0.3573	0.4324	9.939	9.933	0.006	1.4	0.0157
1150.	0.3194	0.3328	0.3261	9.878	9.956	-0.078	1.0439	-0.2825
1200.	0.4875	0.4719	0.4847	9.867	9.522	0.344	1.0531	0.4871
1210.	0.4314	0.4607	0.4460	9.622	9.600	0.022	1.0334	0.0740
1220.	0.4953	0.4601	0.4778	9.656	9.567	0.089	1.0747	0.2776
1230.	0.725	0.7194	0.7221	9.767	9.711	0.056	1.0039	0.111
1240.	0.9111	0.9279	0.8694	9.689	9.756	-0.067	1.0743	-0.1517
1250.	1.1077	1.1397	1.0452	9.800	9.[illegible]	[illegible]	[illegible]	0.2781

FIG. 16

PAIRED COMPARISON TEST

1				9.722			1.1847	
131								
132	1.7174			9.758	9.500	*0.048	0.8801	
133				10.022	9.480		0.8772	
134				9.989	10.022	*0.033	1.2174	
135				10.100	10.022		0.8678	
				10.372	10.422	*0.050	1.2830	
141				10.322	10.378	*0.056	0.8714	
142				10.300	10.350	*0.050	0.9482	
143				10.056	10.150	*0.094	0.8431	
144				10.172	10.150	0.022	0.9331	
145				10.150	10.050	0.100	0.8392	
1				9.922	10.022	*0.100	0.8387	
151				10.050	10.050	0.000	0.7845	
152				10.067	10.056	0.011	1.0813	
153				9.950	9.961	*0.011	0.8115	
154				10.022	10.003	0.022	0.7869	
155				9.900	9.911	*0.011	0.7227	
156				9.722	9.450	0.072	0.8617	
161				9.422	9.450	*0.028	0.8350	
162				9.344	9.411	*0.067	0.9608	
163				9.156	9.100	0.056	1.1958	
164				8.922	8.833	0.089	1.4636	
165				8.750	8.678	0.072	0.9003	
171				8.550	8.489	0.061	1.1215	
172				8.289	8.328	*0.039	0.8066	
173				8.900	8.950	*0.050	1.1284	
174				8.011	8.039	*0.028	0.8006	
175				7.900	7.889	0.011	1.2796	
176				7.922	7.872	0.050	0.8047	
177				7.850	7.750	0.100	1.0783	
181				7.761	7.761	0.000	0.8691	
182				7.639	7.700	*0.061	0.7561	
183				7.322	7.378	*0.056	0.9869	
184				7.269	7.169	0.100	1.1819	
185				7.089	6.922	0.167	0.9359	
1				6.778	6.811	*0.033	0.8745	
191				6.456	6.433	0.022	0.9012	
192				6.089	6.144	*0.056	1.1988	
193				5.900	5.911	*0.011	0.2243	

FIG. 17

PAIRED COMPARISON TEST

TIME	CONFIDENCE INTERVALS FOR (MEANS OF DISTRIBUTION) LOWER CONFIDENCE LIMIT	DIFFERENCE BETWEEN THE CI AND VI MEANS	UPPER CONFIDENCE LIMIT	CONFIDENCE INTERVALS FOR VARIANCES OF DISTRIBUTION LOWER CONFIDENCE LIMIT	RATIO OF VARIANCES	UPPER CONFIDENCE LIMIT
530.	-0.5374	-0.0889	0.3598	0.1363	1.9063	9.0674
540.	-0.4441	-0.1056	0.2329	0.1461	1.2928	8.6983
550.	-0.7047	0.0222	0.7491	0.1238	1.0916	8.1870
600.	-0.4123	-0.2222	0.3678	0.2522	2.9323	10.7421
610.	-0.4136	0.1333	0.6803	0.0897	0.7585	5.6888
620.	-0.0817	0.4000	0.8817	0.1292	1.1077	8.3079
630.	-0.7889	-0.0778	0.6334	0.0902	0.7978	5.9835
640.	-0.3804	0.0389	0.4582	0.0552	0.4905	3.6795
650.	-0.2756	0.1833	0.6433	0.1271	1.1250	8.4375
700.	-0.2795	0.1847	0.5128	0.0878	0.9979	6.4838
710.	-0.4941	0.0000	0.4941	0.2260	2.0000	15.0000
720.	-0.5461	-0.0111	0.5239	0.1183	1.0871	7.8531
730.	-0.8277	-0.0444	0.7389	0.3154	2.8000	21.0000
740.	-0.3817	0.0333	0.4483	0.1249	1.1058	8.2907
750.	-0.2542	0.1389	0.5320	0.0446	0.3945	2.9590
800.	-0.2740	0.1778	0.6318	0.1130	1.0000	7.5000
810.	-0.3944	-0.0222	0.3500	0.1318	1.1666	8.7384
820.	-0.5117	0.0556	0.6218	0.1137	1.0062	7.5461
830.	-0.7804	-0.1333	0.4829	0.2768	2.4498	18.3739
840.	-0.6450	0.0000	0.6450	0.0941	0.8333	6.2500
850.	-0.6414	-0.0111	0.5814	0.1754	1.5536	11.6521
900.	-0.7878	-0.0778	0.6328	0.0590	0.5223	3.9169
910.	-1.0015	-0.0444	0.9128	0.1109	0.9818	7.3632
920.	-0.6111	0.0111	0.6334	0.0851	0.7530	5.6472
930.	-0.8356	-0.0222	0.7912	0.1239	1.0968	8.2260
940.	-0.8014	-0.1333	0.5348	0.0527	0.4666	3.4994
950.	-0.7452	-0.0389	0.6674	0.0987	0.8030	6.0221
1000.	-0.9315	-0.0084	0.9147	0.0854	0.7561	5.6707
1010.	-0.8957	0.0000	0.8957	0.0643	0.5693	4.2698
1020.	-0.7078	-0.0722	0.5634	0.1154	1.0214	7.6607
1030.	-0.3987	0.0778	0.5543	0.1287	1.1390	8.5429
1040.	-0.4315	0.1389	0.7093	0.0722	0.6392	4.7940
1050.	-0.3747	0.1111	0.5969	0.0913	0.8078	6.0585
1100.	-0.4311	0.0000	0.4311	0.0613	0.5426	4.0695
1110.	-0.4778	0.0556	0.5889	0.0793	0.7019	5.2643
1120.	-0.4756	0.0111	0.4979	0.0693	0.6134	4.6005
1130.	-0.3778	0.0778	0.5334	0.0931	0.8237	6.1777
1140.	-0.2560	0.1444	0.5447	0.1016	0.8988	6.7416
1150.	-0.2782	0.0222	0.3226	0.1088	0.9599	7.1994
1200.	-0.1902	0.1444	0.4789	0.1148	1.0161	7.6208
1210.	-0.2431	0.0778	0.3987	0.0441	0.3900	2.9251
1220.	-0.2438	0.0778	0.3994	0.0863	0.7637	5.7276
1230.	-0.1721	0.1444	0.4609	0.1244	1.1008	8.2560
1240.	-0.1561	0.1444	0.4449	0.0829	0.7337	5.5027
1250.	-0.1561	0.1333	0.4228	0.0743	0.6576	4.9320

FIG. 18

PAIRED COMPARISON TEST

TEST	CONFIDENCE LIMIT	DIFFERENCE BETWEEN TWO MEANS	UPPER CONFIDENCE LIMIT	LOWER CONFIDENCE LIMIT	VARIANCE RATIO	UPPER CONFIDENCE LIMIT
1300.	-1.2560	0.0333	1.2728	0.1341	0.1987	5.9259
1310.	-1.2867	0.0667	1.4233	0.1150	0.0298	6.7398
1320.	-1.5474	-0.0444	1.4789	0.1119	0.0951	7.4255
1330.	-1.2911	0.1333	1.4879	0.0991	0.0772	8.5787
1340.	-1.6122	-0.0333	1.3456	0.1376	0.2174	9.1308
1350.	-1.3811	0.0778	1.4868	0.1193	0.0476	7.2573
1400.	-1.6197	-0.0556	1.3187	0.1427	0.2830	9.6727
1410.	-1.3151	-0.0556	1.2042	0.0985	0.0716	6.5357
1420.	-1.3818	-0.0500	1.2518	0.1071	0.0482	7.1117
1430.	-1.2132	-0.0944	1.0243	0.0949	0.0401	6.3008
1440.	-1.4429	0.0222	1.4073	0.0941	0.0331	6.2485
1450.	-1.0004	0.1000	1.1994	0.1061	0.0392	7.0437
1500.	-1.2418	-0.1000	1.0618	0.1041	0.0387	7.0405
1510.	-1.0001	0.0000	1.0001	0.0801	0.0085	5.3138
1520.	-1.0983	0.0556	1.0514	0.1854	1.8810	12.3072
1530.	-1.0054	-0.0111	0.9734	0.1030	0.0115	6.8359
1540.	-1.0049	0.2222	1.0510	0.0867	0.7869	5.7520
1550.	-1.7022	-0.0111	0.7700	0.0817	0.7227	5.4201
1600.	-1.7119	0.0722	0.8556	0.0974	0.4617	6.4828
1610.	-1.7339	-0.0278	0.7274	0.0947	0.4380	6.2850
1620.	-1.7847	-0.0647	0.6514	0.1096	0.9698	7.2734
1630.	-1.0898	0.0556	1.1010	0.1351	1.1958	8.9683
1640.	-1.0188	0.0889	1.1967	0.1656	1.4656	10.9921
1650.	-1.1948	0.0722	1.3834	0.1017	0.9003	6.7521
1700.	-1.2365	0.0611	1.3587	0.1154	1.0215	7.6612
1710.	-1.0427	-0.0389	0.9849	0.0911	0.8066	6.0497
1720.	-1.0140	-0.0560	0.8140	0.1275	1.1284	8.4628
1730.	-1.0224	-0.0278	0.7668	0.0905	0.8006	6.0045
1740.	-1.0704	0.0111	0.9926	0.1220	1.0796	8.0969
1750.	-1.2144	0.0500	1.3144	0.0909	0.8047	6.0358
1800.	-1.1962	0.1000	1.3962	0.1892	1.6743	12.5575
1810.	-1.2274	0.0000	1.2274	0.0978	0.8651	6.4881
1820.	-1.2961	-0.0811	1.1579	0.0854	0.7561	5.6711
1830.	-1.0935	-0.0556	0.9824	0.1115	0.9869	7.4017
1840.	-1.0957	0.1000	1.1957	0.1336	1.1819	8.8640
1850.	-1.0666	0.1667	1.1800	0.1158	1.0244	7.6818
1900.	-1.0755	-0.0333	1.0088	0.0988	0.8745	6.5591
1910.	-1.0219	0.0222	1.0659	0.1018	0.9012	6.7589
1920.	-1.7380	0.0444	0.8269	0.1576	1.3948	10.4610
1930.	-1.1592	-0.0111	0.9369	0.0256	0.2263	1.6974

FIG. 19

GAIN OF INFORMATION

TIME OF DAY	PAIRED COMPARISON TEST: CONFIDENCE INTERVALS: LOWER LIMIT	DIFFERENCE BETWEEN THE MEANS	UPPER LIMIT	POOLED STANDARD DEVIATION	PAIRED DIFFERENCE TEST: CONFIDENCE INTERVALS: LOWER LIMIT	MEAN VALUES OF DIFFERENCES BETWEEN OI AND VI	UPPER LIMIT	STANDARD DEVIATION OF DIFFERENCES
530.	-0.5374	-0.0849	0.3676	0.3259	-0.0217	0.0809	0.1551	0.0741
540.	-0.8861	-0.1858	0.0659	0.5523	-0.0177	0.1854	1.0295	0.0955
550.	-0.7682	0.0222	0.7685	0.5275	-0.1195	-0.0222	0.0751	0.0870
600.	-0.8123	-0.2222	0.3679	0.4285	-0.0671	0.2222	0.5114	0.2687
610.	-0.4135	0.1333	0.6803	0.3972	-0.2703	-0.1333	0.0038	0.1925
620.	-0.0817	0.4000	0.8817	0.3498	-0.8970	-0.4000	0.097	0.4444
630.	-0.7059	-0.0778	0.4338	0.5165	-0.2704	0.0778	0.4261	0.3114
640.	-0.3609	0.0349	0.4106	0.2756	-0.2948	-0.0388	0.2171	0.2298
650.	-0.2784	0.1833	0.6433	0.3360	-0.3849	-0.1833	0.0183	0.1903
700.	-0.2795	0.1167	0.5129	0.2877	-0.0769	-0.1167	0.2435	0.3721
710.	-0.4541	0.0000	0.4541	0.3398	-0.2437	-0.0000	0.2437	0.2170
720.	-0.5461	-0.0111	0.5239	0.3884	-0.5719	0.0111	0.5934	0.5907
730.	-0.4277	-0.0444	0.5389	0.4234	-0.3227	0.0444	0.8115	0.3383
740.	-0.3832	0.0333	0.4499	0.3025	-0.1584	-0.0333	0.0917	0.1118
750.	-0.2582	0.1389	0.5359	0.2855	-0.4384	-0.1389	0.1606	0.2478
800.	-0.2781	0.1778	0.6336	0.3296	-0.5304	-0.1778	0.1749	0.3153
810.	-0.3944	-0.0222	0.3500	0.2703	-0.1122	0.0222	0.1566	0.1202
820.	-0.5107	0.1556	0.8218	0.4113	-0.2394	-0.0556	0.1282	0.1448
830.	-0.7496	-0.1333	0.4829	0.4475	-0.2902	0.1333	0.5569	0.3841
840.	-0.6851	0.0000	0.6851	0.4075	-0.1748	-0.0000	0.1748	0.1581
850.	-0.6035	-0.0111	0.5813	0.4303	-0.1696	0.0111	0.1918	0.1415
900.	-0.7879	-0.0778	0.6324	0.5157	-0.1513	0.0778	0.3068	0.2048
910.	-1.0015	-0.0444	0.9126	0.6050	-0.4136	0.0444	0.5025	0.4094
920.	-0.9143	0.0111	0.9365	0.6718	-0.1918	-0.0111	0.1696	0.1418
930.	-0.9354	-0.0222	0.7811	0.5907	-0.1122	0.0222	0.1566	0.1202
940.	-0.8806	-0.1333	0.5339	0.4986	-0.1625	0.1333	0.4291	0.2448
55.	-0.7802	0.1389	0.8579	0.5903	-0.2230	-0.0389	0.1451	0.1454
1 00.	-0.9805	-0.0556	1.0293	0.6786	-0.1950	0.0056	0.1961	0.1178
1 10.	-0.8257	0.0056	0.8368	0.6037	-0.2300	-0.0056	0.2188	0.2217
1020.	-0.7024	-0.0722	0.5580	0.4585	-0.1113	0.0722	0.2558	0.1481
1 3.	-0.3887	0.2778	0.5583	0.3463	-0.2700	-0.0778	0.1144	0.1193
1 0.	-0.4315	0.1389	0.9093	0.5995	-0.5178	-0.1389	0.2411	0.3389
1 0.	-0.9937	0.0111	0.9758	0.7207	-0.1437	0.0111	0.1661	0.1688
1 0.	-1.1313	0.0222	1.1758	0.8377	-0.2144	-0.0222	0.1700	0.2401
1110.	-1.0677	0.0556	1.1788	0.8188	-0.3565	-0.0556	0.2453	0.2565
1120.	-0.6783	0.1000	1.8756	0.5835	-0.5139	-0.1000	0.3139	0.4431
1130.	-0.7822	0.0389	0.7822	0.5884	-0.1750	0.0389	0.1528	0.1918
1140.	-0.8892	-0.0444	0.8115	0.6575	-0.0968	0.0444	0.2856	0.1484
115 .	-0.7641	0.0222	0.8084	0.5711	-0.1478	-0.0222	0.1033	0.13
12 .	-0.8149	0.1844	1.1831	0.8062	-0.6489	-0.1044	0.3	0.6551
1210.	-0.8874	0.0222	0.9414	0.6677	-0.1478	-0.0222	0.1033	0.1302
122 .	-0.8821	0.1889	1.0855	0.6611	-0.3534	-0.0889	0.1756	0.2369
123 .	-1.2143	-0.0444	1.1256	0.8455	-0.0826	0.0444	0.1715	0.1335
1240.	-1.3547	-0.0667	1.2179	0.9326	-0.3417	0.0000	0.3417	0.2568
125 .	-1.5561	-0.1333	1.2877	1.0326	-0.1167	0.1333	0.3833	0.2939

FIG. 20



GAIN OF INFORMATION

13 .	-1.2[illegible]	[illegible]	1.2[illegible]	0.[illegible]	-0.1[illegible]	-0.[illegible]	0.[illegible]	[illegible]
131 .	-1.2[illegible]	0.[illegible]	1.4[illegible]	0.[illegible]	-0.2004	-0.0667	0.1270	0.1732
132 .	-1.[illegible]	-0.[illegible]	1.[illegible]	1.[illegible]	-0.1419	0.[illegible]	0.2318	0.1467
133 .	-1.2[illegible]	0.[illegible]	1.[illegible]	0.[illegible]	-0.6073	-0.1333	0.[illegible]	0.2449
134 .	-1.4[illegible]	-0.[illegible]	1.3[illegible]	1.[illegible]	-0.5088	0.0333	0.5755	0.4448
135 .	-1.[illegible]	-0.[illegible]	1.[illegible]	1.[illegible]	-0.[illegible]	-0.0778	0.3398	0.3734
140 .	-1.[illegible]	-0.[illegible]	1.3[illegible]	0.[illegible]	-0.2092	0.0500	0.3992	0.3122
141 .	-1.3153	-0.[illegible]	1.2[illegible]	0.[illegible]	-0.1618	0.0556	0.3729	0.1944
142 .	-1.3514	-0.[illegible]	1.2[illegible]	0.[illegible]	-0.1666	0.0500	0.2666	0.1836
143 .	-1.2132	-0.[illegible]	1.0[illegible]	0.[illegible]	-0.2392	0.0944	0.4251	0.2984
144 .	-1.[illegible]	0.[illegible]	1.[illegible]	0.[illegible]	-0.2532	-0.0222	0.2387	0.2333
145 .	-0.[illegible]	0.[illegible]	1.[illegible]	0.[illegible]	-0.2037	-0.1000	0.0937	0.1732
150 .	-1.2614	-0.[illegible]	1.0[illegible]	0.[illegible]	-0.1738	0.1000	0.3738	0.2449
151 .	-1.[illegible]	0.[illegible]	1.0[illegible]	0.[illegible]	-0.3536	-0.0000	0.3536	0.3162
152 .	-0.[illegible]	0.[illegible]	1.[illegible]	0.[illegible]	-0.5893	-0.0556	0.4782	0.4773
153 .	-0.[illegible]	-0.[illegible]	0.[illegible]	0.[illegible]	-0.3532	0.0111	0.3754	0.3257
154 .	-0.[illegible]	0.2222	1.[illegible]	0.[illegible]	-0.4832	-0.2222	0.0387	0.2333
155 .	-0.7922	-0.[illegible]	0.7700	0.[illegible]	-0.2417	0.0111	0.2639	0.2241
160 .	-0.7133	0.[illegible]	0.[illegible]	0.[illegible]	-0.2041	-0.0722	0.1197	0.1716
161 .	-0.7832	-0.[illegible]	0.7276	0.[illegible]	-0.2332	0.0278	0.2887	0.2333
162 .	-0.7847	-1.[illegible]	0.[illegible]	0.[illegible]	-0.1911	0.0667	0.2344	0.1900
163 .	-0.[illegible]	0.0556	1.1010	0.7592	-0.4053	-0.0556	0.2941	0.3127
164 .	-1.[illegible]	0.[illegible]	1.1947	0.[illegible]	-0.4821	-0.0889	0.3043	0.3416
165 .	-1.1943	0.0722	1.3034	0.9232	-0.3013	-0.0722	0.1368	0.2448
170 .	-1.2365	0.[illegible]	1.3597	0.9424	-0.4010	-0.0611	0.2788	0.3039
171 .	-1.0627	-0.[illegible]	0.[illegible]	0.7435	-0.1913	0.0389	0.2491	0.2458
172 .	-0.814[illegible]	-0.[illegible]	0.[illegible]	0.6275	-0.2239	0.0500	0.3239	0.2449
173 .	-0.8226	-0.[illegible]	0.7844	0.5770	-0.1178	0.0278	0.1734	0.1302
174 .	-0.9738	0.[illegible]	0.[illegible]	0.7125	-0.1290	-0.0111	0.1068	0.1454
175 .	-1.2144	0.0500	1.3144	0.9183	-0.2872	-0.0500	0.1872	0.2121
180 .	-1.1842	0.[illegible]	1.3862	0.9341	-0.4709	-0.1000	0.2709	0.3317
181 .	-1.2275	0.0000	1.2275	0.8815	-0.1479	-0.0000	0.1479	0.1323
182 .	-1.[illegible]	-0.[illegible]	1.[illegible]	0.[illegible]	-0.1758	0.0611	0.2980	0.2118
183 .	-1.0935	-0.0556	0.[illegible]	0.7539	-0.0431	0.0556	0.1542	0.0882
184 .	-0.9267	0.[illegible]	1.1257	0.7848	-0.1969	-0.[illegible]	-0.0031	0.[illegible]
185 .	-0.9465	0.[illegible]	1.[illegible]	0.7359	-0.5958	-0.1667	0.2485	0.3873
191 .	-1.0755	-0.[illegible]	1.[illegible]	0.[illegible]	-0.1146	0.[illegible]	0.[illegible]	0.1323
191 .	-1.0215	0.[illegible]	1.[illegible]	0.[illegible]	-0.2372	-0.0222	0.1927	0.1922
192 .	-0.738[illegible]	0.[illegible]	1.[illegible]	0.[illegible]	-0.1798	-0.[illegible]	0.[illegible]	0.1910
193 .	-1.1[illegible]	-0.[illegible]	0.[illegible]	0.[illegible]	-0.4035	0.[illegible]	1.[illegible]	0.[illegible]

FIG. 21

GAIN OF INFORMATION

TIME OF DAY	POOLED VARIANCE	VARIANCE OF DI-VI DIFFERENCES	DIFFERENCE	POOLED STANDARD DEVIATION	STANDARD DEVIATION OF DI-VI DIFFERENCES	DIFFERENCE
530.	0.1362	0.0038	0.1024	0.3259	0.1951	0.2658
540.	0.3491	0.8815	-0.3765	0.5573	0.9255	-0.2732
550.	0.2783	0.0078	0.2707	0.5275	0.1870	0.4405
600.	0.1836	0.0869	0.1167	0.4285	0.2587	0.1698
610.	0.1578	0.0150	0.1428	0.3972	0.1225	0.2747
620.	0.1224	0.1975	-0.0751	0.3498	0.4444	-0.0946
630.	0.2667	0.0969	0.1698	0.5165	0.3114	0.2051
640.	0.0760	0.0524	0.0236	0.2756	0.2288	0.0468
650.	0.1116	0.0325	0.0791	0.3340	0.1803	0.1537
700.	0.0829	0.1037	-0.0210	0.2877	0.3221	-0.0344
710.	0.1087	0.0474	0.0613	0.3298	0.2179	0.1118
720.	0.1312	0.2711	-0.1201	0.3806	0.5207	-0.1321
730.	0.1796	0.1078	0.0717	0.4238	0.3283	0.0953
740.	0.0915	0.0125	0.0790	0.3025	0.1118	0.1907
750.	0.0815	0.0717	0.0098	0.2855	0.2678	0.0176
800.	0.1094	0.0994	0.0092	0.3294	0.3153	0.0142
810.	0.0731	0.0146	0.0586	0.2703	0.1202	0.1501
820.	0.1881	0.3672	-0.1792	0.4113	0.1648	0.2465
830.	0.2003	0.1475	0.0528	0.4475	0.3841	0.0635
840.	0.2475	0.0250	0.2225	0.4975	0.1581	0.3394
850.	0.1451	0.0241	0.1210	0.4303	0.1616	0.2687
900.	0.2841	0.0419	0.2240	0.5157	0.2048	0.3109
910.	0.4831	0.1878	0.3153	0.6950	0.4096	0.2854
920.	0.4514	0.0261	0.4253	0.6718	0.1616	0.5103
930.	0.3689	0.0144	0.3544	0.5907	0.1202	0.4705
940.	0.2349	0.0150	0.1849	0.4846	0.2648	0.2200
950.	0.3364	0.0274	0.3094	0.5803	0.1654	0.4149
1000.	0.4611	0.0324	0.4319	0.6790	0.1704	0.5086
1010.	0.3644	0.0402	0.3242	0.6037	0.2007	0.4030
1020.	0.2287	0.0266	0.1982	0.4580	0.1641	0.2938
1030.	0.1197	0.0319	0.1078	0.3460	0.1093	0.2367
1040.	0.3131	0.1189	0.1942	0.5599	0.3389	0.2210
1050.	0.4911	0.0186	0.4725	0.7007	0.1364	0.5643
1100.	0.7018	0.0294	0.6724	0.8377	0.1716	0.6661
1110.	0.6654	0.0551	0.6103	0.8158	0.2351	0.5807
1120.	0.3179	0.1354	0.1825	0.5638	0.3679	0.1959
1130.	0.3232	0.1125	0.2107	0.5685	0.3355	0.2330
1140.	0.0324	0.0206	0.0118	0.1800	0.1436	0.0364
1150.	0.3251	0.1149	0.2102	0.5701	0.3391	0.2310
1200.	0.4867	0.2359	0.2508	0.6978	0.4856	0.2122
1210.	0.4465	0.1597	0.2868	0.6682	0.3997	0.2685
1220.	0.0774	0.0211	0.0563	0.2782	0.1453	0.1329
1230.	0.7119	0.0223	0.6896	0.8439	0.1493	0.7946
1240.	0.3694	0.1588	0.2106	0.6078	0.3985	0.2093
1250.	1.0485	0.1613	1.0872	1.0239	0.4016	1.0823

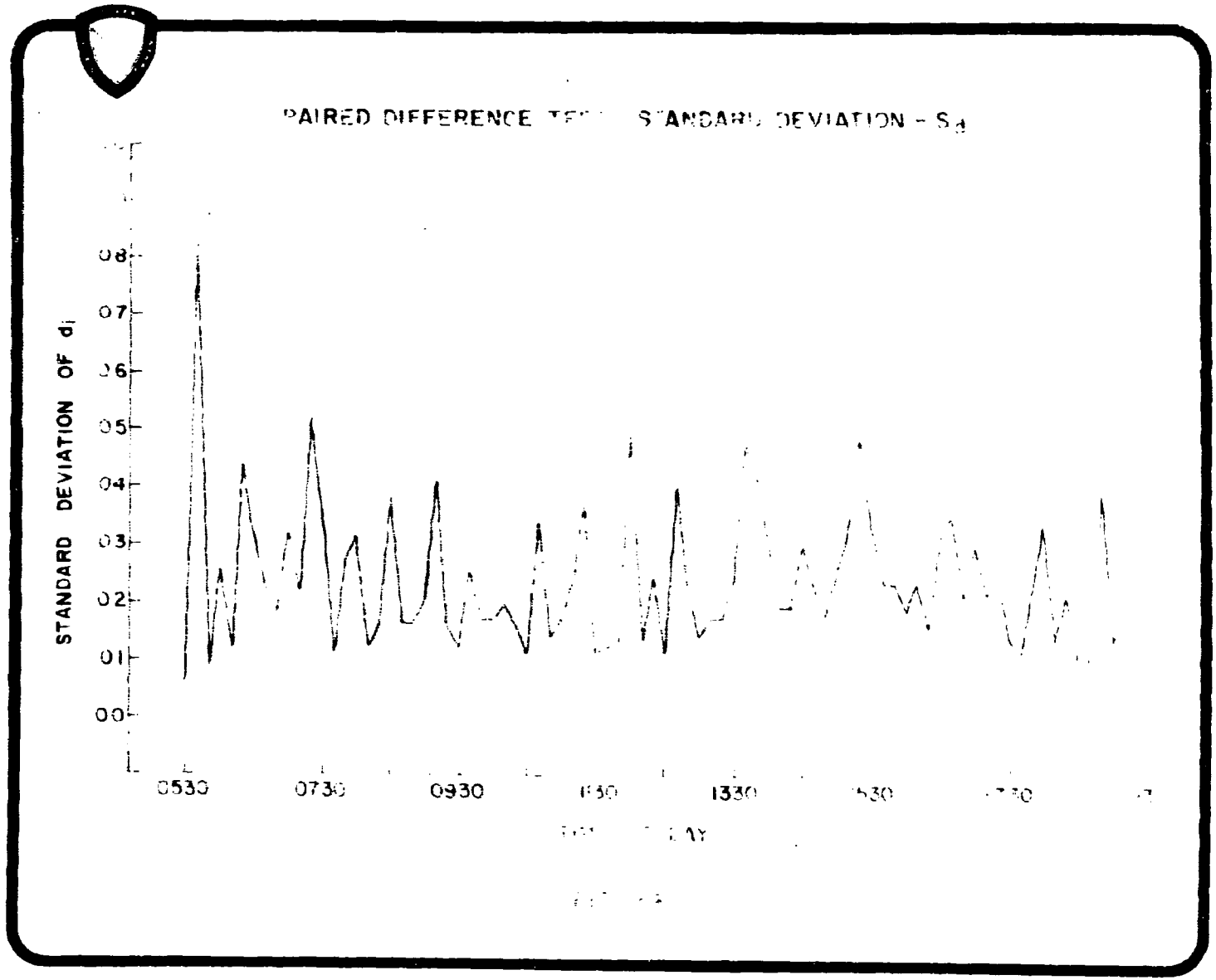
FIG. 22



GAIN OF INFORMATION

			DIFFERENCE			DIFFERENCE
130	[illegible]	[illegible]	[illegible]	.8958	[illegible]	[illegible]
131	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
132	1.[illegible]	[illegible]	[illegible]	1.1963	[illegible]	[illegible]
133	[illegible]	[illegible]	.9876	0.9838	0.2440	0.7398
134	1.[illegible]	[illegible]	0.7878	1.0316	[illegible]	[illegible]
135	1.[illegible]	[illegible]	0.9078	1.0233	0.3734	0.6499
136	[illegible]	[illegible]	[illegible]	0.9947	[illegible]	0.6825
141	[illegible]	[illegible]	[illegible]	0.9148	[illegible]	[illegible]
142	[illegible]	[illegible]	.8562	0.9454	0.1936	0.7517
143	[illegible]	[illegible]	0.5711	0.8125	0.2984	0.5141
144	[illegible]	[illegible]	0.5865	0.7880	0.2333	0.5547
145	[illegible]	[illegible]	0.6075	0.7986	0.1732	0.6252
15	[illegible]	[illegible]	0.6519	0.8458	0.2449	0.5988
151	[illegible]	[illegible]	0.4275	0.7263	0.3162	0.4101
152	[illegible]	[illegible]	.2953	0.7232	0.4773	0.2460
153	[illegible]	[illegible]	0.4051	0.7150	0.3257	0.3892
154	[illegible]	[illegible]	0.3079	0.6018	0.2333	0.3685
155	[illegible]	[illegible]	0.2707	0.5873	0.2261	0.3612
16	[illegible]	[illegible]	0.2949	0.5607	0.1716	0.3892
161	[illegible]	[illegible]	0.2465	0.5486	0.2333	0.3153
162	[illegible]	[illegible]	.2494	0.5215	0.1500	0.3715
163	[illegible]	[illegible]	0.4784	0.7592	0.3127	0.4465
164	[illegible]	[illegible]	.5234	0.8005	0.3516	0.4529
165	[illegible]	[illegible]	0.8103	0.9232	0.2048	0.7184
17	[illegible]	[illegible]	0.7957	0.9424	0.3039	0.6385
171	[illegible]	[illegible]	0.5134	0.7435	0.2058	0.5377
172	[illegible]	[illegible]	0.3337	0.6275	0.2449	0.3825
173	[illegible]	[illegible]	0.3180	0.5778	0.1302	0.4469
174	[illegible]	[illegible]	.6989	0.7178	0.1054	0.6074
175	[illegible]	[illegible]	0.7952	0.9183	0.2121	0.7061
18	[illegible]	[illegible]	0.7025	0.9341	0.3317	0.6024
181	[illegible]	[illegible]	.7774	0.8915	0.1323	0.7593
182	[illegible]	[illegible]	0.7388	0.8853	0.2118	0.6735
183	[illegible]	[illegible]	0.5604	0.7538	0.0982	0.6656
184	[illegible]	[illegible]	0.5474	0.7449	0.0866	0.6583
185	[illegible]	[illegible]	0.3915	0.7359	0.3873	0.3486
19	[illegible]	[illegible]	.5553	0.7559	0.1323	0.6245
191	[illegible]	[illegible]	0.5274	0.7580	0.1922	0.5658
192	[illegible]	[illegible]	.3089	0.5642	[illegible]	0.4472

FIG. 23


PAIRED DIFFERENCE TEST STANDARD DEVIATION - S_d
STANDARD DEVIATION OF d_i
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0530
0730
0930
1330

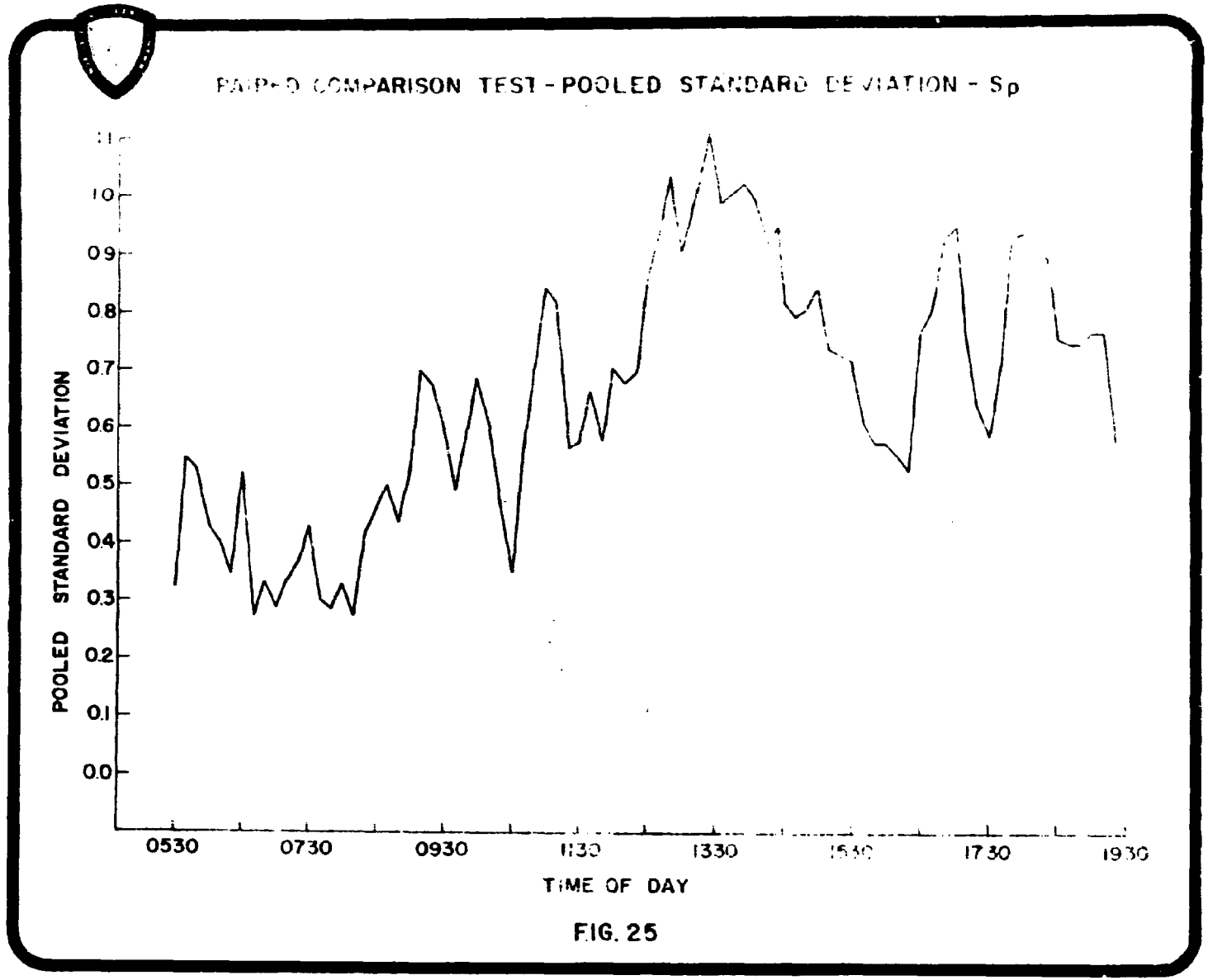

PAIRED COMPARISON TEST - POOLED STANDARD DEVIATION - Sp
POOLED STANDARD DEVIATION
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0530
0730
0930
1130
1330
1530
1730
1930
TIME OF DAY

FIG. 25

POSITION LOCATION VIA MULTIPLE TRIANGULATION*

Glenn A. Stoops and Edward I. Spitznagel, Jr.
Litton Scientific Support Laboratory
Fort Ord, California

1. INTRODUCTION. Classical triangulation in the plane involves locating an unknown position by measuring its direction from two known points and finding the intersection point of the two location lines. If, more generally, there are n known points reporting directions--and there are errors in the observed directions--then the n lines cannot be expected to intersect in a common point. Two different methods of obtaining a closed form estimate of the true position, with variations on each, will be derived and discussed, along with an error analysis of each method.

2. ESTIMATION METHODS.

a. Least Squares (LSQ). The n known positions are denoted by $P_i\ (x_i, y_i)$, $i = 1, 2, \ldots, n$, and the observed directions by the respective angles ϕ_i. This yields an equation for the i th direction line L_i:

$$y - y_i = \tan \phi_i\ (x - x_i).$$

The perpendicular distance d from an arbitrary point P (x, y) to the line L_i is given by:

$$d\ (P, L_i) = |\ (x - x_i) \sin \phi_i - (y - y_i) \cos \phi_i\ |.$$

The LSQ method, roughly, determines a point that is close to all the lines L_i, in the least squares sense. Specifically, define the function

$$\begin{aligned} f\ (P) &= f\ (P\ (x, y)) \\ &= \sum_{i=1}^{n} (d\ (P, L_i))^2 \\ &= \Sigma\ d^2\ (P, L_i) \\ &= \Sigma[(x - x_i)^2 \sin^2\phi_i + (y - y_i)^2 \cos^2\phi_i \\ &\qquad -2\ (x - x_i)\ (y - y_i) \sin\phi_i \cos\phi_i]. \end{aligned}$$

The (unweighted) LSQ estimate is the point P that minimizes this function. A slightly more general function is

$$g\ (P) = \Sigma\lambda_i d^2\ (P, L_i),$$

where $\{\lambda_i\}$ is a set of fixed, but arbitrary, nonnegative numbers. Physically, the minimization of this function corresponds to weighting some of the

*This article has been reproduced photographically from the author's manuscript.

information more heavily than the rest. Setting each w_i equal to 1 reduces the problem to the unweighted case.

Minimization of $g(P)$ is straightforward: compute $\partial g/\partial x$ and $\partial g/\partial y$, set both equal to 0, and solve the pair of equations. The estimates derived are:

$$x_{LSQ} = \frac{\begin{vmatrix} \Sigma(w_i x_i \sin^2\theta_i - w_i y_i \sin\theta_i \cos\theta_i) & -\Sigma w_i \sin\theta_i \cos\theta_i \\ \Sigma(w_i x_i \sin\theta_i \cos\theta_i - w_i y_i \cos^2\theta_i) & -\Sigma w_i \cos^2\theta_i \end{vmatrix}}{\begin{vmatrix} \Sigma w_i \sin^2\theta_i & -\Sigma w_i \sin\theta_i \cos\theta_i \\ \Sigma w_i \sin\theta_i \cos\theta_i & -\Sigma w_i \cos^2\theta_i \end{vmatrix}}$$

$$y_{LSQ} = \frac{\begin{vmatrix} \Sigma w_i \sin^2\theta_i & \Sigma(w_i x_i \sin^2\theta_i - w_i y_i \sin\theta_i \cos\theta_i) \\ \Sigma w_i \sin\theta_i \cos\theta_i & \Sigma(w_i x_i \sin\theta_i \cos\theta_i - w_i y_i \cos^2\theta_i) \end{vmatrix}}{\begin{vmatrix} \Sigma w_i \sin^2\theta_i & -\Sigma w_i \sin\theta_i \cos\theta_i \\ \Sigma w_i \sin\theta_i \cos\theta_i & -\Sigma w_i \cos^2\theta_i \end{vmatrix}}$$

b. Point of intersection (POI). Each pair of direction lines intersects in a unique point, and there are $\binom{n}{2} = n(n-1)/2$ such points. The intersection point of lines L_i and L_j, $P_{ij}(x_{ij}, y_{ij})$, is derived easily from their equations and is given by:

$$x_{ij} = \frac{\sin\theta_i (y_j \cos\theta_j - x_j \sin\theta_j) - \cos\theta_j (y_i \cos\theta_i - x_i \sin\theta_i)}{\sin(\theta_i - \theta_j)}$$

$$y_{ij} = \frac{\sin\theta_i (y_j \cos\theta_j - x_j \sin\theta_j) - \sin\theta_j (y_i \cos\theta_i - x_i \sin\theta_i)}{\sin(\theta_i - \theta_j)}$$

The POI estimation simply involves computing a weighted or unweighted average of the $\binom{n}{2}$ points P_{ij}. That is, a set of nonnegative numbers $\{w_{ij}\}$, $1 \le i < j \le n$, is selected, and the coordinates of the estimated point are:

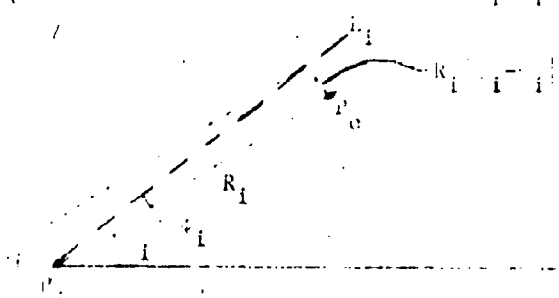
$$x_{POI} = \sum_i \sum_j w_{ij} x_{ij} / \sum_i \sum_j w_{ij}$$

$$y_{POI} = \sum_i \sum_j w_{ij} y_{ij} / \sum_i \sum_j w_{ij}.$$

Easiest of course is to choose the weights $\{w_{ij}\}$ so that they sum to 1.

3. APPROXIMATIONS USED IN ERROR ANALYSES. If one assumes that the errors in measuring the true angles from the respective known points P_i are small, then the following approximation is a very sound one and aids considerably in the error analyses. Note that the error in locating an unknown position is in general a function of the position, so that in this sense the errors derived are conditional errors, conditioned by the actual (unknown) position.

Let (R_i, θ_i) denote the true distance and direction of the unknown point P_o from the i^{th} known point P_i. The line L_i goes through P_i and misses P_o by a distance of approximately $R_i \sin \varepsilon_i \approx \varepsilon_i$ (see figure). The



geometry of the error analyses is eased considerably if line L_i is replaced by line Λ_i, parallel to the line from P_i to P_o and displaced by an amount $R_i \sin \varepsilon_i \approx \varepsilon_i$ in the appropriate direction. Lines L_i and Λ_i are virtually indistinguishable, and both miss P_o by the same amount.

4. ERROR ANALYSES. If P_o (x_o, y_o) denotes the true point to be located and P (x, y) denotes the computed estimate (via either LSQ or POI), then the basic criterion used to assess the quality of the respective estimates is $E[(x - x_o)^2 + (y - y_o)^2] = E[|P - P_o|^2]$, the mean squared radial error. The only assumptions made about the random variables ε_i are that they be independent, unbiased, and have common variance σ^2.

a. LSQ Method. Using the counterclockwise convention for positive angles, one can show that the distance from an arbitrary point P (x, y) to Λ_i is given by:

$$d(P, \Lambda_i) = (x - x_o) \sin \theta_i - (y - y_o) \cos \theta_i + R_i (\varepsilon_i - \varepsilon_i).$$

Given P_o--and thus R_i and θ_i --the point P (x_{LSQ}, y_{LSQ}), because of the way it was computed, also minimizes the function $h(P) = \sum_i d^2 (P, \Lambda_i)$.

Therefore, minimizing E (P) expresses x_{LSQ} and y_{LSQ} (conditionally) in terms of the ϕ_i's and θ_i's; the expressions are:

$$x_{LSQ} = \frac{\Sigma \ell_i R_i (\phi_i - \theta_i)\cos\theta_i)(\Sigma \ell_i \sin\theta_i \cos\theta_i) - (\Sigma \ell_i R_i (\phi_i - \theta_i)\sin\theta_i)(\Sigma \ell_i \cos^2\theta_i)}{(\Sigma \ell_i \sin^2\theta_i)(\Sigma \ell_i \cos^2\theta_i) - (\Sigma \ell_i \sin\theta_i \cos\theta_i)^2} + x_o$$

and similarly for y_{LSQ}. Note immediately that x_{LSQ} and y_{LSQ} are linear functions of $(\phi_i - \theta_i)$, so that the mean squared radial error is in fact a function only of the variance of $(\phi_i - \theta_i)$ (and the geometry).

The error expression, $E[(x_{LSQ}-x_o)^2 + (y_{LSQ}-y_o)^2]$, is still relatively [illegible] by the fact that the random variables ϕ_i are independent [illegible] given by:

$$\frac{\left\{[\Sigma \ell_i \cos^2\theta_i)^2 + (\Sigma \ell_i \cos\theta_i \sin\theta_i)^2](\Sigma \ell_i^2 R_i^2 \sin^2\theta_i) + [(\Sigma \ell_i \sin^2\theta_i)^2 + (\Sigma \ell_i \cos\theta_i \sin\theta_i)^2](\Sigma \ell_i^2 R_i^2 \cos^2\theta_i) - 2(\Sigma \ell_i \cos\theta_i \sin\theta_i)(\Sigma \ell_i)(\Sigma \ell_i^2 R_i^2 \sin\theta_i \cos\theta_i)\right\}}{[(\Sigma \ell_i \sin^2\theta_i)(\Sigma \ell_i \cos^2\theta_i) - (\Sigma \ell_i \sin\theta_i \cos\theta_i)^2]^2} \sigma^2.$$

b. POI Method. The coordinates of the point P_{ij} are determined approximately by finding the intersection of Λ_i and Λ_j. This yields:

$$x_{ij} = x_o + \frac{(R_i(\phi_i - \theta_i))\cos\theta_j - (R_j(\phi_j - \theta_j))\cos\theta_i}{\sin(\theta_i - \theta_j)}$$

$$y_{ij} = y_o + \frac{(R_i(\phi_i - \theta_i))\sin\theta_j - (R_j(\phi_j - \theta_j))\sin\theta_i}{\sin(\theta_i - \theta_j)}$$

Since the estimate (x_{POI}, y_{POI}) is given by

$$\left.\begin{aligned} x_{POI} &= \Sigma_{i<j} \lambda_{ij} x_{ij} \\ y_{POI} &= \Sigma_{i<j} \lambda_{ij} y_{ij} \end{aligned}\right\} \quad \Sigma_{i<j} \lambda_{ij} = 1,$$

the error function $E[(x_{POI}-x_o)^2 + (y_{POI}-y_o)^2]$ is derived by combining the x_1 terms, x_2 terms, etc., squaring, and careful bookkeeping, noting again the independence and common variance σ^2 of the random variables. For convenience of notation, define $\ell_{ij} = \ell_{ji}$, if $i>j$, and define $\lambda_{ij} = \ell_{ij}/\sin(\theta_i - \theta_j)$, for all $i \neq j$. Then the mean squared radial error is given by:

$$\left\{ \sum_{i=1}^{n} R_i^2 \left[\left(\sum_{j \neq i} \lambda_{ij} \cos \theta_j \right)^2 + \left(\sum_{j \neq i} \lambda_{ij} \sin \theta_j \right)^2 \right] \right\}^{-2}.$$

5. IMPROVEMENTS BY ADJUSTING THE WEIGHTS. The derived error expressions are too complex to permit many general observations to be made. Extensive study of examples indicates that the LSQ method leads to smaller error than does the POI method. In particular, it is conjectured that unweighted LSQ is always better than unweighted POI. However, either method can be considerably improved through the use of even imperfect information about the unknown location. In the following subsections idealized weights are derived for each method, weights that minimize the respective error expressions but are unattainable because they require perfect information about the unknown locations. In later sections these idealized weights are interpreted as yielding lower bounds on the error expressions, bounds that cannot be attained but can be approached by various iterative schemes.

a. LSQ Method. The intention here is to find the set of nonnegative weights $\{\lambda_i\}$ that minimizes the (conditional) error expression, E, for a particular unknown location P_o. Of course the set is different for each P_o and thus cannot be derived, even in theory, without perfect knowledge of P_o itself. However, the mere existence of such a minimal set indicates a lower limit on how much improvement can be expected even with partial information about P_o.

Note first that E is homogeneous in the λ_i's, that is, multiplying the λ_i's by a common factor leaves E unchanged. Note also that a minimum could not occur along a boundary (one or more λ_i's equal to 0), since this means ignoring some of the data. Thus, a necessary condition for a local minimum to occur is that all the partials, $\partial E/\partial \lambda_i$, be equal to 0 at some point (or any multiple thereof). One solution (and, it is conjectured, the unique one) is:

$$\lambda_i = R_i^{-2}, \text{ for all } i.$$

Again, this solution was suggested through study of numerous examples, and it can be checked, through straightforward but tedious computation, that

it does indeed satisfy $\partial E/\partial \ell_i = 0$, for all i. The simplicity and plausibility of this solution, once attained, make it a most likely candidate for unique global minimum. In particular, the data from more remote points P_i should obviously be weighted less heavily. Interestingly, the minimal weights do not depend on the angles $\{\theta_i\}$. The error expression, E_{MIN}, for $\ell_i = R_i^{-2}$, is given by:

$$\Sigma R_i^{-2}/[(\Sigma R_i^{-2}\cos^2\theta_i)(\Sigma R_i^{-2}\sin^2\theta_i) - (\Sigma R_i^{-2}\cos\theta_i\sin\theta_i)^2],$$

a relatively simple expression.

b. POI Method. Since E is a simpler expression in the POI case, it would be expected that minimization is also easier and this is true. In fact, since E is a quadratic function of the ℓ_{ij}'s (or λ_{ij}'s) and the constraint, $\sum_{i<j}\ell_{ij} = 1$, is linear, the minimization problem--via partial differentiation and Lagrange multipliers--reduces to solving a set of simultaneous linear equations in ℓ_{ij}. Once more, study of examples suggested a solution, hence <u>the</u> solution, which is given by:

$$\ell_{ij} = R_i^{-2}R_j^{-2}\sin^2(\theta_i-\theta_j)/\sum_{i<j}R_i^{-2}R_j^{-2}\sin^2(\theta_i-\theta_j),$$

for all i<j. As noted, the optimal weights for POI involve both the R_i's and the θ_i's, perhaps because the points of intersection are so intimately tied to both. The most striking fact is that the error expression, E_{MIN}, is precisely the same as that derived in the previous subsection for LSQ. (In fact, the "optimal" estimates themselves are identical as well.) Certainly this fact is more than coincidence, and some of its unifying implications will be discussed in a later section.

6. INTERPRETATION OF "OPTIMAL" WEIGHTINGS. In Section 5, best possible weightings were derived for each estimation method, given a particular P_0. As stated, these weights are unattainable, requiring omniscience, but they indicate directions for improvement of the respective unweighted methods. Simply stated, some information about the location of P_0 is better than none at all. This suggests that an initial (unweighted) estimate, $P_{(1)}$, be computed, via either method, and the distances R_{i1} (and angles θ_{i1}) from each P_i to $P_{(1)}$ be determined. Then these R_{i1}'s (and θ_{i1}'s, if applicable) can be used to compute a second, weighted estimate, $P_{(2)}$. Since $P_{(2)}$ is

presumed better than $P_{(1)}$, its associated R_{i2}'s and γ_{i2}'s can be used to compute a third, weighted estimate, $P_{(3)}$. Naturally this iterative scheme can be continued as long as desired. Unfortunately, very little work has been done at this time to investigate convergence, and rate of convergence, of the iterations to any type of best estimate, but this is certainly an area for continued research. It is felt that the iterations for the LSQ method probably converage rather well, whereas the POI iterations, because they involve angles as well as distances and the data is in terms of angles, allow the possibility of circularity and instability.

7. CONCLUSIONS. Two methods have been discussed for estimating the location of an unknown point, given direction data from n known points. Closed form formulas for the estimates were derived, as well as general expressions for the mean squared radial error of the two methods, least squares (LSQ) and point of intersection (POI). In addition, idealized "optimal" weights were derived for each method, weights that reduce the respective errors to their smallest possible values, conditioned by the true location of the unknown point.

It was noted in Section 6 that the optimally weighted solutions for the two methods are identical. To understand the underlying reason for this, consider the following estimation method: for an arbitrary point P (x, y), let γ_i be the angle it makes with P_i and let $F(P) = \sum_{i=1}^{n} (\cdot_i - \gamma_i)^2$. Find, if possible, the point P that minimizes F (P). To the degree of approximation used throughout the paper, this P is identical to the optimal LSQ or POI point. To see this, rewrite F (P) as:

$$F(P) = \sum_{i=1}^{n} R_i^{-2}[R_i^2(\cdot_i - \gamma_i)^2] \qquad \cdot w_i d^2(P, \cdot_i),$$

where $\{w_i\}$ is the optimal set of weights for LSQ. This method, finding a least squares fit to the raw data itself, $\{\cdot_i\}$, was not used until late in the investigation and still is not preferable to the others merely because it does not lend itself to a closed form solution. Conceivably, with sufficient computational facilities, this method may be preferable to either LSQ or POI, particularly if the latter require a large number of iterations. Note that this estimate is in fact the maximum likelihood estimate in the case where the $\cdot_i$'s are normally distributed.

In summary, the two methods, LSQ and POI, were presented separately as parallel methods, although the preceding paragraph does tend to unify the two under one broad theory. Emphasized was the simplicity of both methods, especially in the absence of unequal weights. The POI method is easier to visualize graphically, while the LSQ method has fewer computations and smaller associated error. Both methods are systematic and easy to apply in many practical situations.

A General Computational Algorithm for Bayesian Confidence Bounds

by Richard W. Clarke

WATERVLIET ARSENAL
WATERVLIET, NEW YORK

INTRODUCTION

For anyone unfamiliar with Bayesian analysis this paper should serve as an introduction to this very useful confidencing technique. The aspect of this paper which might be interesting to those already familiar with this subject is simply that I have outlined a computational algorithm which will eliminate the very messy mapping which arises in applying the Bayes formulation.

I found it convenient in what follows to work with a specific example in order to make a few basic points. A more general treatment may be found in a Watervliet Arsenal technical report by the same title.[1]

A BAYESIAN CONFIDENCE BOUND ON RELIABILITY

The basic contention of the Bayesian analysis is that any physical parameter about which we have less than precise knowledge may be treated as a random variable. For instance, the shape and location Weibull parameters might be treated as random variables if we are using the Weibull density to represent a set of failure data. If from the data we can construct the joint density of those parameters, then the reliability density for a given safe life or the safe life density for a given reliability will follow directly.

To arrive at that joint parameter density we must first specify some prior knowledge of those parameters. This consists of stating that, from prior testing of similar items, these parameters are likely to be within certain bounds. If very little information is available we might say only that a certain parameter can take on any value between two limits, and that each value between those bounds is equally likely before testing.

Suppose that five components have been cycled, under actual conditions of field use, to failure and that these failures (cycles or hours, etc.) are $X_1, X_2 \ldots X_5$. From prior testing of similar mechanical components we deduce that the population from which these failures come can be reasonably approximated by the two parameter Weibull density. Then the density of X is given by:

$$f(x/\beta,\tau) = \frac{\beta}{\tau}\left[\frac{x}{\tau}\right]^{\beta-1} \exp\left[-\left(\frac{x}{\tau}\right)^{\beta}\right] \qquad (1)$$

This article has been reproduced photographically from the author's manuscript.

in which X is a random variable drawn from $W(\beta, T)$. For any set of Weibull parameters we have that the joint density of the above five independent observations is:

$$g(X_1, X_2 \cdots X_5/\beta, T) = \prod_{i=1}^{5} f(X_i/\beta, T) \qquad (2)$$

or:

$$g(\mathbf{X}/\Theta) = \prod_{i=1}^{5} f(X_i/\Theta) \qquad (3)$$

Now let our prior knowledge of the parameters be represented by:

$$g_1(\Theta) = g_1(\beta, T) \qquad (4)$$

In this case we might say, for instance, that:

$$g_1(\beta, T) = \frac{1}{Q_1 Q_2} \qquad \begin{matrix} \beta_0 < \beta < \beta_0 + Q_1 \\ T_0 < T < T_0 + Q_2 \end{matrix} \qquad (5)$$

$$= 0 \qquad \text{ELSEWHERE}$$

Bayes theorem then states that the posterior knowledge available on the parameter space for these five observations is:

$$g_2(\hat{\Theta}/\mathbf{X}) = \frac{g(\mathbf{X}/\Theta)\, g_1(\Theta)}{\int_{\Theta} g(\mathbf{X}/\Theta)\, g_1(\Theta)\, d\Theta} \qquad (6)$$

In the example we've been following, noting that g_1 is a constant:

$$g_2(\hat{\Theta}/\underline{X}) = \frac{\prod_{i=1}^{5} f(x_i/\Theta)}{\int_0^{Q_1}\int_0^{Q_2}\prod_{i=1}^{5} f(x_i/\Theta)\,d\tau\,d\beta} \qquad (7)$$

We now have an expression which assigns a probability density to each point in the parameter space.

At any point in that space we can related reliability (R) to safe life (X_S).

$$R = 1 - \int_0^{X_S} f(x/\Theta)\,dx \qquad (8)$$

or for the Weibull example:

$$R = \exp\left[-\left(\frac{X_S}{T}\right)^{\beta}\right] \qquad (9)$$

Then for a given safe life the density of the reliability estimator may be found by mapping the parameter joint density (g_2) onto reliability (R) through Eqn. (8). Analytically:

$$r(\hat{R}) = \int_0^{\infty} g_2(\hat{\beta},\hat{T})\left|\frac{\partial(\hat{\beta},\hat{T})}{\partial(\hat{R},\hat{T})}\right| d\hat{T} \qquad (10)$$

In which the Jacobian is evaluated from Eqn. (8). A one sided, $(1-\alpha)$ 100% lower confidence bound on reliability is then the 100αth percentile in the reliability estimator density or:

$$\alpha = \int_0^{\hat{R}_c} r(\hat{R})\,d\hat{R} \qquad (11)$$

Expression (10) is not particularly simple to evaluate. If we happened to be working with a three parameter density the Jacobian would contain three terms instead of two and two of the variables would have to be eliminated by integration instead of one. In general, that is for most two and three parameter densities, the integrations could not be carried out in closed form and some numerical or computer solution would be required.

A SUBSTITUTE FOR THE ANALYTICAL MAPPING

Instead of the usual analytical mapping as defined by Expression (10), we can start directly with the posterior joint parameter density and do a numerical mapping onto reliability as follows.

In specifying the prior parameter density (g_1) choose a rectangular region of definition such as in Eqn. (5) above. Then divide that region into small subregions by dividing the parameter axes into equal intervals. In the Weibull example we have been following the midpoint of a specific subregion would be represented by:

$$\beta_i = \beta_0 + \frac{Q_1}{N_\beta}(i - 1/2) \qquad i = 1,2,\cdots N_\beta \qquad (12)$$

$$T_j = T_0 + \frac{Q_2}{N_T}(j - 1/2) \qquad j = 1,2,\cdots N_T \qquad (13)$$

In which Q_1 and Q_2 represent the ranges over which the parameters β and T are defined (See Eqn. (5)), and N_β and N_T are the number of intervals into which those ranges have been partitioned. (See Figure I)

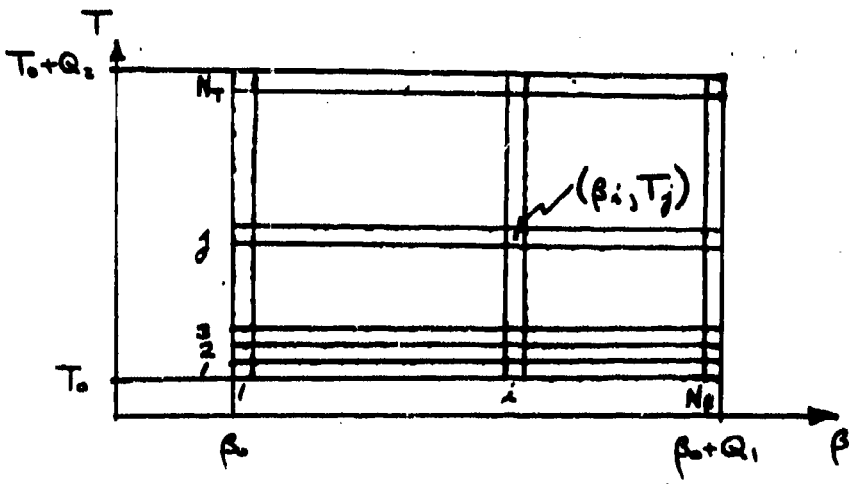


FIGURE I

If these subregions are "small" enough (how small will be discussed later), the joint posterior density (g_2) at the subregion midpoint will be a reasonable approximation throughout the subregion. Then the probability that any subregion contains the actual population parameters can be represented by

$$P_{ij} = g_2(\beta_i, T_j)\, V \tag{14}$$

$$V = \frac{Q_1}{N_\beta}\,\frac{Q_2}{N_T} \tag{15}$$

This probability (Pij) can then be associated with an interval on the range of possible reliabilities by calculating the reliability for the parameters β i and Tj:

$$R_{ij} = \exp\left[-\left(\frac{X_s}{T_j}\right)^{\beta_i}\right] \tag{16}$$

An actual mapping of the ijth subregion onto reliability might look like Figure II below.

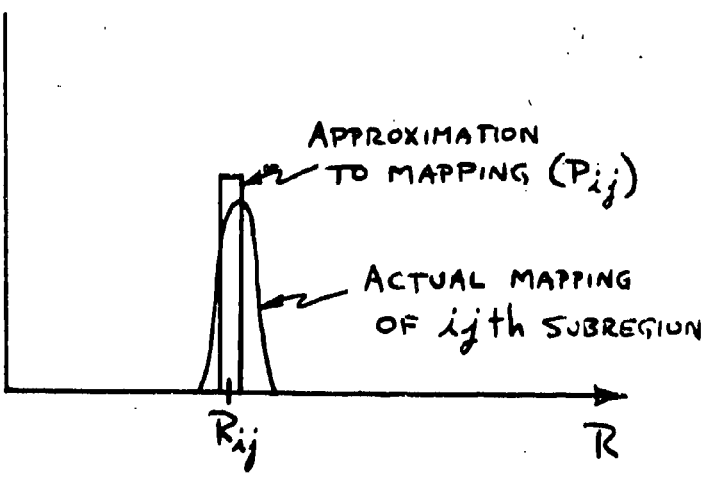


FIGURE II

We will approximate this mapping by dividing the reliability axis into intervals and assigning the entire Pij to the interval in which Rih falls. Mathematically, calculate:

$$\Delta R = \frac{R_{MAX} - R_{MIN}}{M} \qquad (17)$$

with R_{MAX} = Maximum reliability possible

R_{MIN} = Minimum reliability possible

M = Number of intervals on the reliability axis

then:

$$I_{ij} = \frac{R_{ij} - R_{MIN}}{\Delta R} + 1 \qquad (18)$$

Truncating I_{ij} to an integer value then defines the interval number to which Pij is to be assigned.

By running through all the (βi, Tj) combinations and assigning each Pij to an interval on the reliability axis we are constructing a histogram which approximates the reliability estimator density. (See Figure III)

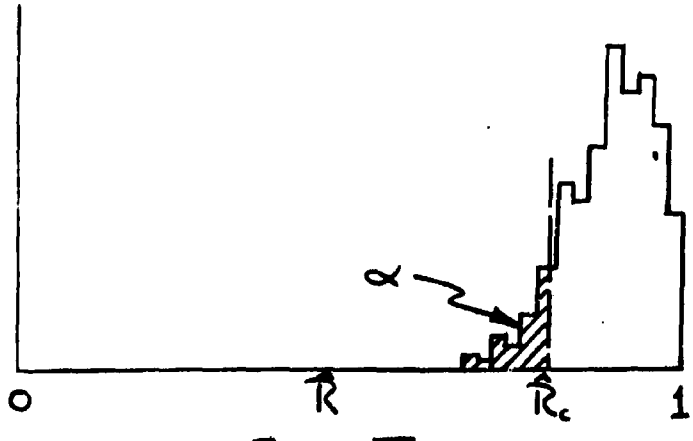


FIGURE III

The accuracy of the process depends only on the interval sizes chosen. We have simply replaced the integral evaluation of the mapping process (Eqn. 10), by a much more straightforward numerical evaluation. Confidenced reliability follows from the histogram by replacing Eqn. (11) with a summation.

Looking back on the process we can note that certain simplications are possible. The evaluation of the posterior parameter joint density could be written:

$$g_2(\hat{\theta}/\bar{X}) = K\, g(\bar{X}/\theta)\, g_1(\theta) \qquad (19)$$

in which the constant is:

$$K = \frac{1}{\int_{\theta} g(\bar{X}/\theta)\, g_1(\theta)\, d\theta} \qquad (20)$$

Then:

$$P_{ij} = KV\, g(\bar{X}/\beta_i, T_j)\, g_1(\beta_i, T_j) \qquad (21)$$

But the sum of all Pij should be unity so that:

$$KV = \frac{1}{\sum_{i=1}^{N_\beta} \sum_{j=1}^{N_T} g(\bar{X}/\beta_i, T_j)\, g_1(\beta_i, T_j)} \qquad (22)$$

In other words we do not have to evaluate the integral in Eqn. (20).

GRID SIZE

The remaining problem then is to determine in any case what interval size is sufficiently "small." No satisfactory solution to this problem is presently available. In applying the technique to actual data sets, however, the following points were noted.

1. One specific application to the lognormal density with a uniform prior parameter density yielded the following confidence bounds on safe life for a given reliability:

Grid Size (N_μ X N_σ)	Bound (Cycles)
15 x 15	800
40 x 40	1150
70 x 70	1150

In this case 15 x 15 was too coarse, but 40 x 40 was agood as 70 x 70.

2. An application of the three parameter Weibull seemed to converge with a grid size of 5 x 5 x 5. Finer grids resulted in a negligible change in the confidence bound.

3. For very high reliabilities (.999, .9999, etc.) the lower confidence bound on safe life seems to increase as the grid is made finer. This would indicate that this method yields, for given grid size, a confidence bound which is on the conservative side of the "exact" Bayesian confidence bound.

CONCLUSION

The point which makes this computational mapping extremely interesting is that it can be extended to any distributional form; it can be extended to system reliability work in which the joint posterior parameter space for all components is mapped onto system reliability, and so on. It's drawback, of course, is that it is completely computer dependent and for large parameter spaces the computations can be expensive.

References:

1. Clarke, R. W., "A General Computational Algorithm for Bayesian Confidence Bounds," Watervliet Arsenal Report WVT-6911

2. Clarke, R. W., "Statistical Determination of Confidenced Safe Fatigue Life for the 175mm M113E1 Gun Tube," Watervliet Arsenal Report WVT-6909

Exact Lower Confidence Limits on Normal and Lognormal Reliability

by Royce W. Soanes, Jr.

WATERVLIET ARSENAL
WATERVLIET, NEW YORK

This paper is a synopsis of References (8) and (9) which were written in order to document more fully the solution to the problem of concern:

Given a population having a normal or lognormal life distribution, and a representative sample of failures drawn from this population, calculate an exact 100 C % lower confidence limit on population reliability (R) for a given mission life (or calculate the mission life (x) corresponding to a given lower confidence limit on reliability.)

The normal reliability estimator* is given by

$$\hat{R} = 1 - \Phi\left(\frac{x - \hat{\mu}}{\hat{\sigma}}\right) \tag{1}$$

By performing a bivariate change of variable, the joint density of $\hat{R}$ and $\hat{\sigma}$ may be obtained in terms of the joint density of $\hat{\mu}$ and $\hat{\sigma}$.

$$h(\hat{R}, \hat{\sigma}) = f(\hat{\mu}, \hat{\sigma}) \sqrt{2\pi}\, \hat{\sigma}\, e^{\frac{1}{2} z_{\hat{R}}^2} \tag{2}$$

The joint density of $\hat{\mu}$ and $\hat{\sigma}$ may be determined from the fact that:

(1) $\hat{\mu}$ and $\hat{\sigma}$ are independent random variables

(2) $\frac{n\hat{\sigma}^2}{\sigma^2}$ has a chi-square distribution with n-1 degrees of freedom and

(3) $\hat{\mu}$ is normally distributed with mean μ and standard deviation $\frac{\sigma}{\sqrt{n}}$

---

*Estimates are maximum likelihood

This article has been reproduced photographically from author's manuscript.

The joint density of $\hat{\mu}$ and $\hat{\sigma}$ is therefore:

$$g(\hat{\mu},\hat{\sigma}) = \frac{2}{\sigma^2}\sqrt{\frac{n}{2\pi}}\frac{\left(\frac{n}{2}\right)^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{\hat{\sigma}}{\sigma}\right)^{n-2} e^{-\frac{n}{2}\left[\left(\frac{\hat{\sigma}}{\sigma}\right)^2+\left(\frac{\hat{\mu}-\mu}{\sigma}\right)^2\right]} \tag{3}$$

The joint density of $\hat{R}$ and $\hat{\sigma}$ is therefore:

$$h(\hat{R},\hat{\sigma}) = \frac{2\sqrt{n}}{\sigma}\frac{\left(\frac{n}{2}\right)^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)}\left(\frac{\hat{\sigma}}{\sigma}\right)^{n-1} e^{-\frac{n}{2}\left[\left(\frac{\hat{\sigma}}{\sigma}\right)^2+\left(\frac{\hat{\mu}-\mu}{\sigma}\right)\right]+\frac{1}{2}\zeta_{\hat{R}}^2} \tag{4}$$

but by definition,

$$\hat{\mu} = \bar{x} + \hat{\sigma}\zeta_{\hat{R}}$$

$$\mu = \bar{x} + \sigma\zeta_R$$

letting $K = \dfrac{2\sqrt{n}\left(\frac{n}{2}\right)^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)}$

$$h(\hat{R},\hat{\sigma}) = \frac{K}{\sigma}\left(\frac{\hat{\sigma}}{\sigma}\right)^{n-1} e^{-\frac{n}{2}\left[\left(\frac{\hat{\sigma}}{\sigma}\right)^2+\left(\frac{\hat{\sigma}}{\sigma}\zeta_{\hat{R}}-\zeta_R\right)^2\right]+\frac{1}{2}\zeta_{\hat{R}}^2} \tag{5}$$

Now $\hat{\sigma}$ is integrated out to obtain the density of $\hat{R}$:

$$h(\hat{R}) = \int_0^\infty \frac{K}{\sigma}\left(\frac{\hat{\sigma}}{\sigma}\right)^{n-1} e^{-\frac{n}{2}\left[\left(\frac{\hat{\sigma}}{\sigma}\right)^2 + \left(\frac{\hat{\sigma}}{\sigma}z_{\hat{R}} - z_R\right)^2\right] + \frac{1}{2}z_{\hat{R}}^2} d\hat{\sigma}$$

letting $$s = \frac{\hat{\sigma}}{\sigma}$$

$$h(\hat{R}) = K\int_0^\infty s^{n-1} e^{-\frac{n}{2}\left[s^2 + \left(sz_{\hat{R}} - z_R\right)^2\right] + \frac{1}{2}z_{\hat{R}}^2} ds \qquad (6)$$

Since s is a dummy variable of integration and $\hat{R}$ is the argument of h, the only numbers upon which the form of h is dependent are R and n. The density of the reliability estimator is therefore a one parameter (R) density which is independent of the life density population parameters and mission life.

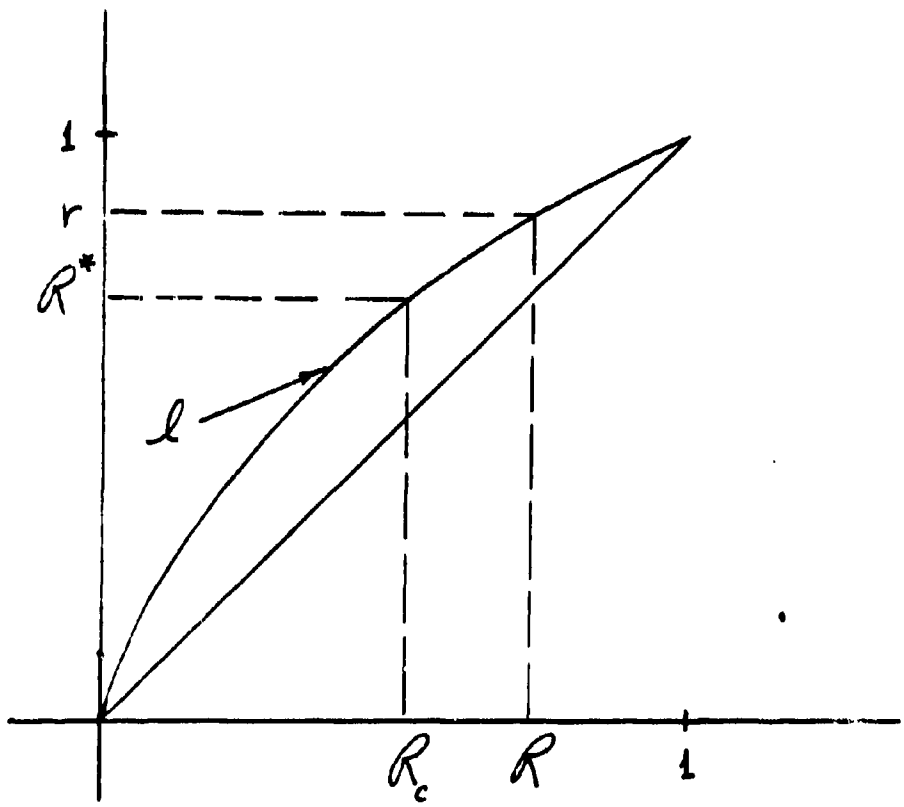
Changing the argument of n to avoid confusion and adding the subscript R to n to indicate its dependence on the population reliability R, the density of $\hat{R}$ is:

$$h_R(v) = K\int_0^\infty s^{n-1} e^{-\frac{n}{2}\left[s^2 + \left(z_R - sz_v\right)^2\right] + \frac{1}{2}z_v^2} ds \qquad (7)$$

The distribution function of $\hat{R}$ is therefore given by:

$$H_R(r) = K \int_0^r \int_0^\infty s^{m-1} e^{-\frac{m}{2}\left[s^2 + (z_R - s z_v)^2\right] + \frac{1}{2} z_v^2} \, ds \, dv \qquad (8)$$

The meaning of the Neyman method of finding a one sided confidence interval for R may be explained through the following diagram:



The curve 1 is determined by:

$$P(\hat{R} < r ; R) = C$$

or

$$H_R(r) = C \tag{9}$$

i.e., 1 is determined such that for any population reliability R, the reliability estimator $\hat{R}$ falls below 1 100 C % of the time. Suppose now that the true value of population reliability is R as shown in the figure. We don't know R or r but we do know 1. If the experiment is now performed and the reliability estimate R* is calculated, the 100 C % lower confidence limit on R is R_c from the diagram. This is so because if the experiment is performed many times, R* will be below r 100 C % of the time and hence R_c will be below R 100 C % of the time.

Confidenced reliability R_c is therefore determined by solving for R_c:

$$H_{R_c}(R^*) = C \tag{10}$$

Before this is done, however, the distribution function of $\hat{R}$ should be simplified. Changing the order of integration in Eqn. (8) and making some appropriate changes of variable, one has:

$$H_R(r) = \frac{K}{\sqrt{n}} \int_0^{\infty} s^{n-2} e^{-\frac{ns^2}{2}} \left\{ 1 - \Phi\left[\sqrt{n}\left(z_R - s z_r\right)\right] \right\} ds \tag{11}$$

Multiplying out in Eqn. (11) and evaluating an integral, one has:

$$H_R(r) = 1 - \frac{2\left(\frac{n}{2}\right)^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} \int_0^{\infty} s^{n-2} e^{-\frac{ns^2}{2}} \Phi\left[\sqrt{n}\left(z_R - s z_r\right)\right] ds \qquad (12)$$

Using Eqn. (12), Eqn. (10) now becomes:

$$\frac{(1-C)\Gamma\left(\frac{n-1}{2}\right)}{2\left(\frac{n}{2}\right)^{\frac{n-1}{2}}} = \int_0^{\infty} s^{n-2} e^{-\frac{ns^2}{2}} \Phi\left[\sqrt{n}\left(z_{R_c} - s z_{R^*}\right)\right] ds \qquad (13)$$

If R^* were calculated using the desired mission life and the sample parameter estimates, Eqn. (13) could be solved numerically for R_c, but for purposes of calculating tables, it is better to stipulate R_c and solve Eqn. (13) for z_{R^*} instead. This was done for confidence levels of 90% and 95%, confidenced reliabilities of .999, .995, .99, .975, .95, .925, .90, .875, .85 and sample sizes of 2—10, 15, 20, 25, 30.

The equations used with the tables to calculate mission life for the normal and lognormal models are:

$$x = \mu^* - \sigma^* z_{R^*} \qquad (14)$$

$$x = e^{\mu_e^* - \sigma_e^* z_{R^*}} \qquad (15)$$

The lognormal case is only trivially different from the normal case because the lognormal reliability estimator is

$$\hat{R}_L = 1 - \Phi\left(\frac{\ln x - \hat{\mu}_L}{\hat{\sigma}_L}\right) \tag{16}$$

and the logs of the data are by definition normally distributed.

P_c	n												
	C=90%												
	2	3	4	5	6	7	8	9	10	15	20	25	30
.999	34.764	11.820	8.232	6.833	6.086	5.619	5.297	5.060	4.879	4.363	4.113	3.962	3.858
.995	28.972	9.911	6.914	5.742	5.115	4.722	4.451	4.252	4.099	3.663	3.451	3.322	3.233
.99	26.163	8.990	6.280	5.217	4.647	4.290	4.044	3.862	3.723	3.325	3.131	3.013	2.933
.975	22.042	7.647	5.354	4.451	3.966	3.661	3.450	3.294	3.174	2.831	2.664	2.562	2.492
.95	18.512	6.505	4.569	3.801	3.387	3.126	2.944	2.811	2.707	2.411	2.265	2.176	2.115
.925	16.232	5.772	4.064	3.383	3.015	2.781	2.619	2.499	2.407	2.139	2.008	1.927	1.872
.90	14.500	5.215	3.681	3.066	2.732	2.520	2.372	2.262	2.177	1.932	1.811	1.737	1.685
.875	13.083	4.760	3.367	2.806	2.500	2.305	2.168	2.067	1.989	1.762	1.649	1.579	1.532
.85	11.877	4.371	3.099	2.583	2.300	2.120	1.994	1.900	1.827	1.615	1.509	1.444	1.399

R_c	n												
	C=95%												
	2	3	4	5	6	7	8	9	10	15	20	25	30
.999	69.687	16.971	10.64	8.387	7.243	6.548	6.080	5.742	5.485	4.769	4.430	4.228	4.091
.995	58.084	14.241	8.947	7.058	6.097	5.512	5.118	4.832	4.615	4.010	3.722	3.550	3.434
.99	52.458	12.924	8.132	6.419	5.545	5.014	4.654	4.394	4.198	3.644	3.381	3.223	3.116
.975	44.205	11.006	6.945	5.488	4.742	4.288	3.980	3.757	3.588	3.110	2.883	2.746	2.653
.95	37.137	9.377	5.940	4.699	4.062	3.672	3.407	3.215	3.068	2.656	2.458	2.339	2.259
.925	32.573	8.331	5.295	4.193	3.625	3.276	3.040	2.867	2.735	2.364	2.185	2.076	2.003
.90	29.107	7.539	4.806	3.809	3.293	2.976	2.760	2.603	2.482	2.141	1.976	1.876	1.808
.875	26.273	6.891	4.406	3.495	3.022	2.730	2.531	2.386	2.274	1.958	1.804	1.711	1.647
.85	23.860	6.338	4.064	3.226	2.789	2.519	2.335	2.199	2.096	1.800	1.657	1.569	1.509

References

1. Shewhart, W. A. : Statistical Method from the Viewpoint of Quality Control

   Graduate School, U. S. Department of Agriculture, 1939

2. Johnson, N. L., and Welch, B. L. : "Applications of the Non-central t-Distribution" Biometrika, 1940

3. Wald, A., and Wolfowitz, J. : "Tolerance Limits for a Normal Distribution" Annals of Mathematical Statistics, 1946

4. Eisenhart, C., Hastay, M. W., and Wallis, W. A. : Techniques of Statistical Analysis, McGraw Hill, 1947

5. Resnikoff, G. J., and Lieberman, G. J. : Tables of the Non-central t-Distribution

   Stanford University Press, 1957

6. Bowker, A. H., and Lieberman, G. J. : Engineering Statistics, Prentice-Hall, 1959

7. Lloyd, D., and Lipow, M., : Reliability: Management, Methods, and Mathematics, Prentice-Hall, 1962

8. Soanes, R. W., "Confidenced Normal and Lognormal Reliability for Any Sample Size"

   Watervliet Arsenal Report
   WVT-6910 1969

9. -------: "Tables Facilitating Confidenced Reliability Calculations for the Normal or Lognormal Distribution"

   Watervliet Arsenal Report
   WVT-6937 1969

REAL-TIME SIMULATION TECHNIQUE FOR EVALUATING A GYRO-SEEKER ASSEMBLY

Elwood D. Baas
Systems Analysis Directorate
White Sands Missile Range, New Mexico

ABSTRACT. Missile simulations of systems including a gyro-seeker guidance assembly have often excluded the gyro-seeker representation by assuming that some ideal proportional tracking ratio will be achieved. Thus, some of the basic characteristics of the guidance loop are omitted or approximated. This paper develops a real-time simulation technique so as to include the basic functions of the gyro-seeker assembly such as precession, nutation, drift, gain, noise, etc.

In the first section of the paper, the model equations are derived and are used in a discussion of the system parameters and system dynamics.

The second section of the paper presents the analog computer mechanization and results of the simulation, some of which have been verified by system experiments, and some predicted by analytical theory.

SYSTEM DESCRIPTION. The basic gyro-seeker unit consists of a gyroscope, rotating gyro magnet, and stationary induction coils about the gyro. (See Figure 1.) Target source energy is collected and focused to produce a spot image on a reticle centered on the spin axis. When the image spot is off center the reticle pattern produces an error signal which is modulated at spin frequency. The amplitude of the error signal is a function of the radial displacement of the image from the reticle center, while the phase corresponds to azimuthal position about the seeker axis. After being amplified and filtered, the signal is fed to the precession coil which torques the gyro magnet so as to precess the gyro toward a null position with respect to the line of sight. The processed a.c. signal can also be demodulated into orthogonal components using reference coils. The demodulated d.c. components can be used for tracking or guidance signals.

ACKNOWLEDGMENT. The author wishes to thank George E. Hoffman, Karl G. Goodloe, and Nancy M. Wade for their assistance in the preparation of this paper.

The remainder of this article has been reproduced photographically from the author's manuscript.

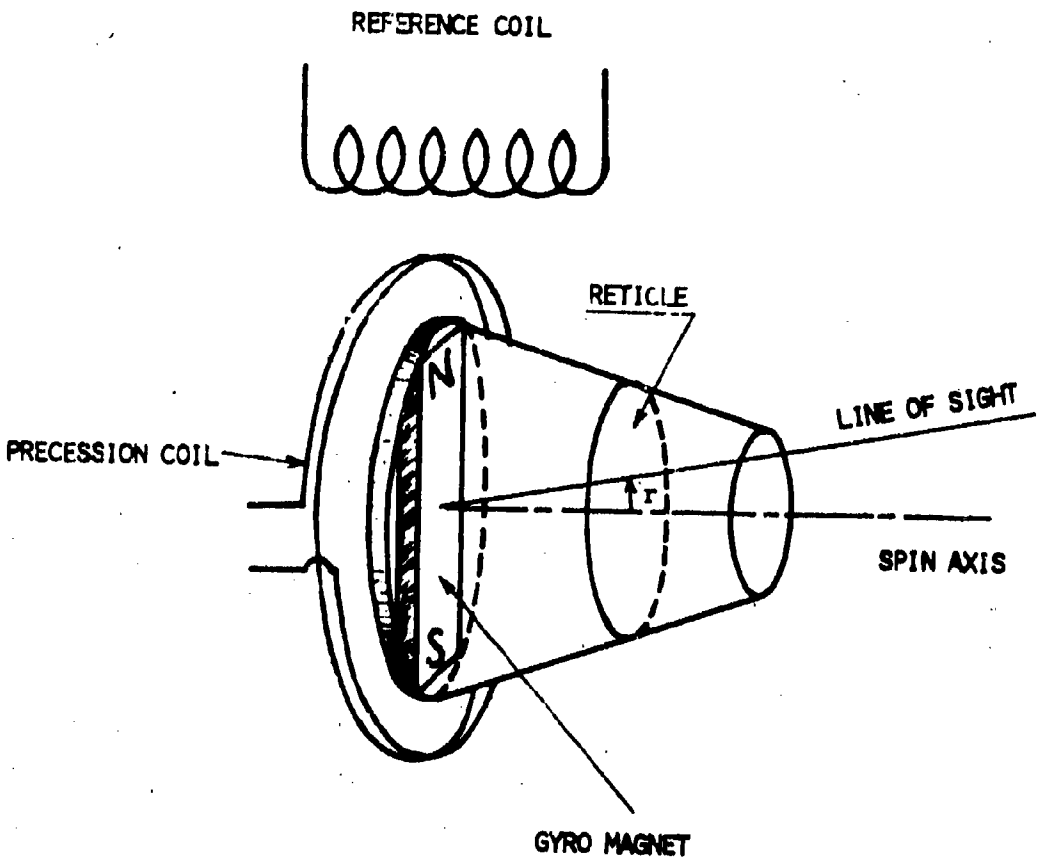

REFERENCE COIL
RETICLE
LINE OF SIGHT
PRECESSION COIL
N
S
SPIN AXIS
GYRO MAGNET

FIGURE 1

Mathematical Model

Two coordinate systems will be referenced in the model derivation. The g> system, or ground system, will be fixed at the initial gyro position with $\bar{g}_1$ axis horizontal and pointed at the initial target ground position. The s> coordinate system, or seeker system, will be fixed to the center of gravity of the gyro assembly. The $\bar{s}_1$ axis is along the gyro spin axis and $\bar{s}_2$ is along the North-South axis of the gyro magnet. (See Fig. 2).

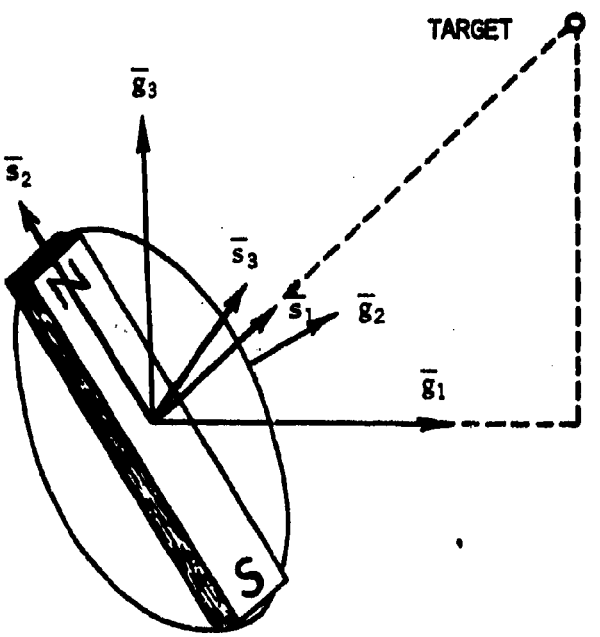


FIGURE 2

The yaw, pitch, roll sequence for the Euler transformation between coordinate systems is given by,

$s\rangle = M(\phi_1, \phi_2, \phi_3)\, g\rangle$, where the subscript denotes rotation about the respective axis. In detail then, the model Euler equations are:

$$\begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} = \begin{bmatrix} C\phi_3\, C\phi_2 & S\phi_3\, C\phi_2 & -S\phi_2 \\ C\phi_3\, S\phi_2\, S\phi_1 - S\phi_3\, C\phi_1 & S\phi_3\, S\phi_2\, S\phi_1 + C\phi_3\, C\phi_1 & C\phi_2\, S\phi_1 \\ C\phi_3\, S\phi_2\, C\phi_1 + S\phi_3\, S\phi_1 & S\phi_3\, S\phi_2\, C\phi_1 - C\phi_3\, S\phi_1 & C\phi_2\, C\phi_1 \end{bmatrix} \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix}$$

$$\dot{\phi}_3 = \sec\phi_2\,(\omega_2 \sin\phi_1 + \omega_3 \cos\phi_1)$$

$$\dot{\phi}_2 = \omega_2 \cos\phi_1 - \omega_3 \sin\phi_1$$

$$\dot{\phi}_1 = \omega_1 + \sin\phi_2\, \dot{\phi}_3$$

$$\phi_3 = (\phi_3)_o + \int \dot{\phi}_3\, dt$$

$$\phi_2 = (\phi_2)_o + \int \dot{\phi}_2\, dt$$

$$\phi_1 = (\phi_1)_o + \int \dot{\phi}_1\, dt$$

To derive an equation for the seeker output signal we assume that the reticle is parallel to the $\bar{s}_2 - \bar{s}_3$ plane and centered on the $\bar{s}_1$ axis. The optics produce a target image point on the reticle, as shown in Fig. 3, whenever the $\bar{s}_1$ axis deviates from the line of sight. The radial displacement d of the image point from the reticle center is proportional to the angle ϵ between the $\bar{s}_1$ axis and the line of sight vector to the target. The distance d is now modeled by deriving a distance r in the $\bar{s}_2 - \bar{s}_3$ plane which is proportional to d. Let $\bar{\sigma}$ be the unit

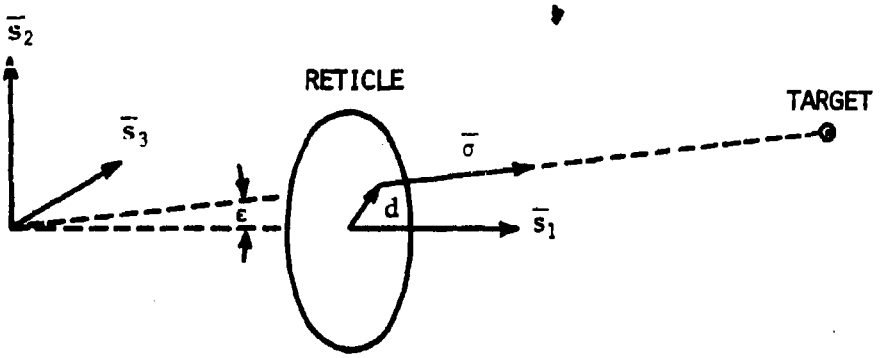


FIGURE 3

vector along line of sight vector in the ground coordinate system. Then,

$$\bar{\sigma} = \frac{X}{R}\bar{g}_1 + \frac{Y}{R}\bar{g}_2 + \frac{Z}{R}\bar{g}_3, \text{ where}$$

$$R = \sqrt{X^2 + Y^2 + Z^2} .$$

Let σ_{s_2} and σ_{s_3} denote the components of $\bar{\sigma}$ in the seeker coordinate system. (See Fig.4). Then $r = \sqrt{\sigma_{s_2}^2 + \sigma_{s_3}^2}$ depends on ϵ and

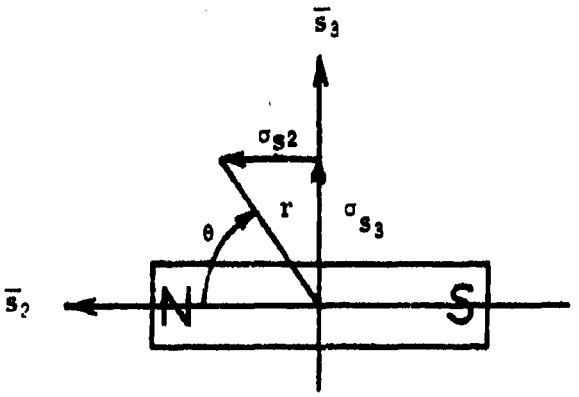


FIGURE 4

not on the length of $\bar{\sigma}$. Also the length r is proportional to the actual displacement d which depends on the optics and other physical parameters. The azimuthal phase determined by the angle θ measured from the $\bar{s}_2$ axis to the image point line is given by $\sin\theta = \frac{\sigma_{s_3}}{r}$. (See Fig. 4). This derivation assumes that reticle modulated output has the form of a sine wave. Other wave shapes could be generated by using various reticle patterns and electronic processing. For a discussion of other wave forms and their effect on gyro precession see reference 1. Then $e_o = Kr \sin\theta = K\sigma_{s_3}$ is an equation which represents the seeker output signal. To compute σ_{s_3} we recall that,

$\bar{\sigma}_s = M(\phi_1, \phi_2, \phi_3)\,\bar{\sigma}g$, so that

$$\sigma_{s_3} = \frac{X}{R}\left[c\phi_3\, s\phi_2\, c\phi_1 + s\phi_3\, s\phi_1\right] + \frac{Y}{R}\left[s\phi_3\, s\phi_2\, c\phi_1 - c\phi_3\, s\phi_1\right] + \frac{Z}{R}\, c\phi_2\, c\phi_1 .$$

The signal $e_o = Kr\sin\theta$ is amplified and applied to the coil about the gyro whose field acts on the permanent gyro magnet to precess the gyro. The variation of the magnetic field is thus proportional to r sin θ and its direction is perpendicular to the plane of the coil as shown in Fig. 5 . Suppose that the magnetic field $\bar{B}_c$ of the coil makes an angle λ with respect to $\bar{s}_1$, then, $|\bar{B}_c \cdot \bar{s}_1|\,\bar{s}_1 = |\bar{B}_c|\cos\lambda\,\bar{s}_1$. The magnetic field of the permanent magnet can be written simply

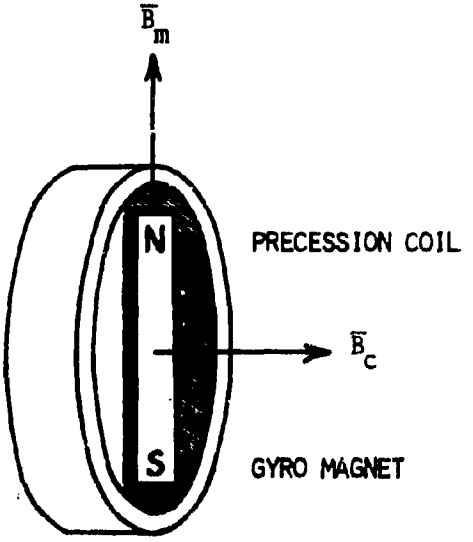


FIGURE 5

as $\bar{B}_m = |\bar{B}_m|\bar{s}_2$. The interaction of the two magnetic fields produces a torque which tends to align the two fields. This torque is given by the vector cross product, $T_3\bar{s}_3 = |\bar{B}_m|\ \bar{s}_2 \times |\bar{B}_c|\cos\lambda\ \bar{s}_1$ where $|\bar{B}_c| = Kr\sin\theta$ and $|\bar{B}_m|$ is constant. Note that the torque T_3 will be a maximum when the spot line is along the $\bar{s}_3$ axis, then the rotation of the magnet will be in the $\bar{s}_2$ - $\bar{s}_1$ plane and the precession in the $\bar{s}_3$ - $\bar{s}_1$ plane. Let $\bar{T} = T_1\bar{s}_1 + T_2s_2 + T_3\bar{s}_3$ be the total external torque acting on the gyro. Then the angular accelerations are given by,

$$\dot{\omega}_3 = \frac{T_3}{I_3} - \omega_1\omega_2\ \frac{[I_2 - I_1]}{I_3}$$

$$\dot{\omega}_2 = \frac{T_2}{I_2} - \omega_1\omega_3\ \frac{[I_1 - I_3]}{I_2}$$

$$\dot{\omega}_1 = \frac{T_1}{I_1} - \omega_3\omega_2\ \frac{[I_3 - I_2]}{I_1}$$

If we assume that the only torque acting is T_3 as derived above, and that the transverse moments are equal, then the equations reduce to:

$$\dot{\omega}_3 = \frac{T_3}{I_3} - \frac{[I_2 - I_1]}{I_3}$$

$$\dot{\omega}_2 = - \omega_1 \omega_3 \frac{[I_1 - I_3]}{I_2}$$

Integrating the components of $\dot{\bar{\omega}}$ yields the angular velocity components of $\bar{\omega}$ of the seeker relative to the seeker coordinate system. Using the seeker to ground transformation, $\bar{\omega}$ can be transformed to $\bar{\omega}_g = \dot{\phi}_1 + \dot{\phi}_2 + \dot{\phi}_3$ by,

$$\dot{\phi}_3 = \sec \phi_2 \, (\omega_2 \sin \phi_1 + \omega_3 \cos \phi_1)$$

$$\dot{\phi}_2 = \omega_2 \cos \phi_1 - \omega_3 \sin \phi_1$$

$$\dot{\phi}_1 = \omega_1 + \sin \phi_2 \, \dot{\phi}_3.$$

Integrating $\dot{\phi}_>$ yields the angles $\phi_>$ of the ground coordinate system which are needed to compute $\dot{\theta}_{s_3}$. Thus the loop is closed.

Orthagonal components of the seeker signal are obtained by demodulating the signal using two reference signals. Reference signals can be obtained by mounting coils about the gyro 90° apart. (See Fig. 6). A sinusoidal voltage is produced by each coil as the flux lines of the gyromagnet cut the windings of the coil. These signals can simply be modeled as $K\sin(\phi_1+\gamma)$ and $K\cos(\phi_1+\gamma)$ where γ is the angle the coils are rotated

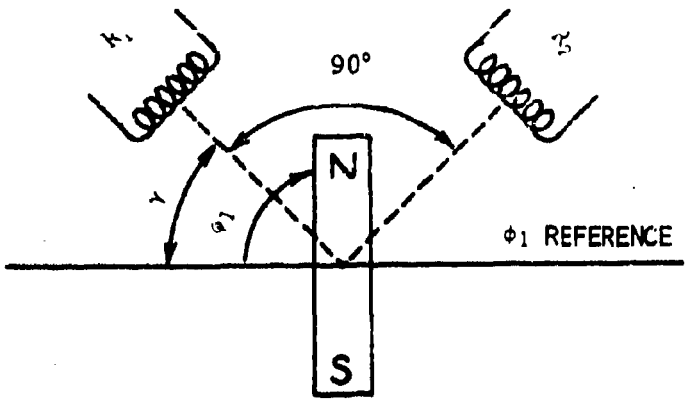


FIGURE 6

from the ϕ_1 reference. The inputs to each demodulator are the seeker signal and one of the reference signals. The filtered output is a d.c. level which is proportional to the error amplitude component in the respective plane or direction determined by the angle γ. The mathematical representation for the phase demodulator is not given since the actual electronic network is easily adapted to the analog computer components. Fig.7 shows a typical phase demodulator bridge network used for one plane.

The basic gyro-seeker model which has been developed can easily be expanded or modified to include hardware changes or known parameter variations. For example, the actual seeker output signal could be a function of source intensity, target range, noise, filtering and other phase and amplitude

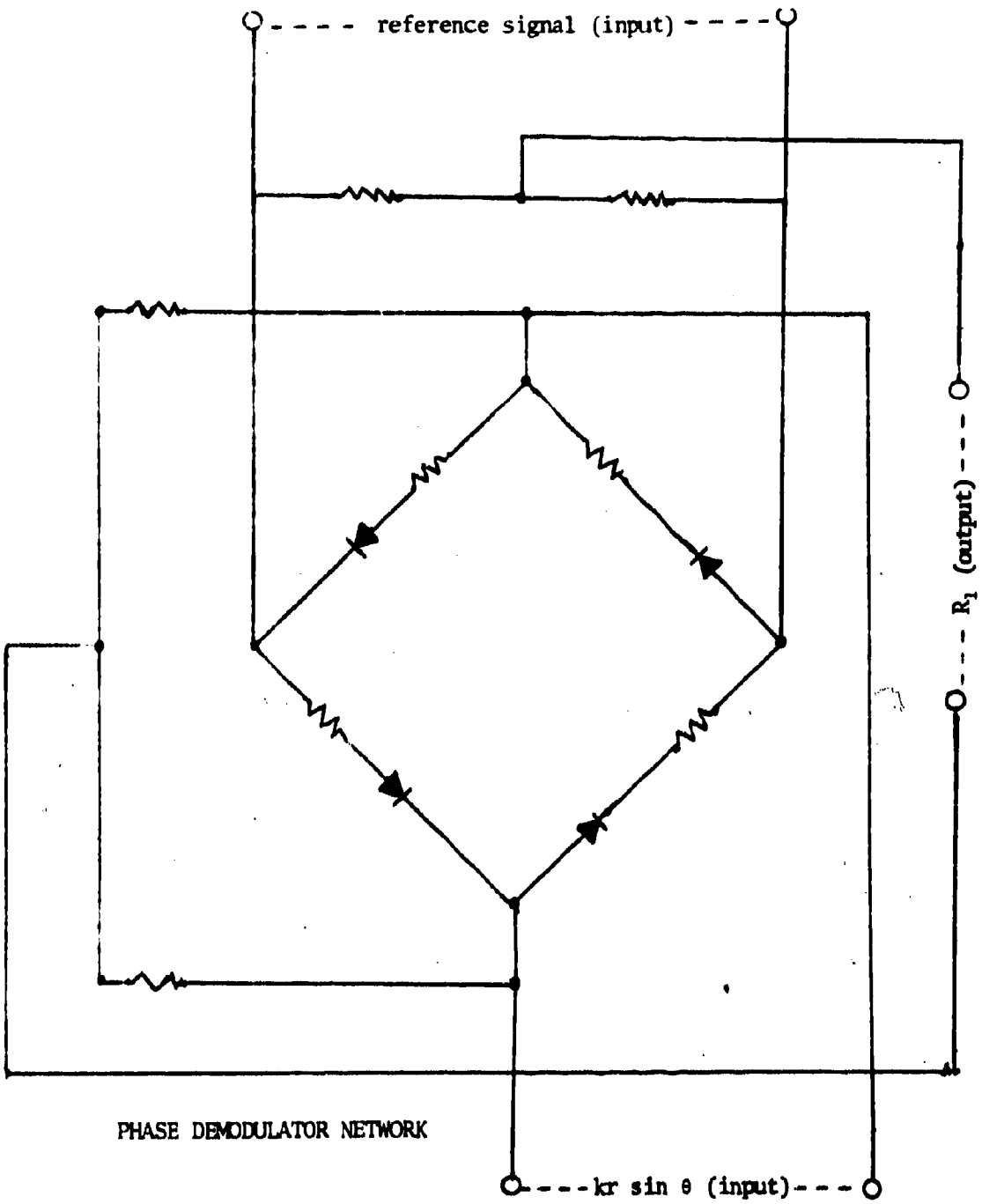

reference signal (input)
R_1 (output)
PHASE DEMODULATOR NETWORK
kr sin θ (input)

FIGURE 7

variations. Most of these variations can be modeled as transfer function additions or multiples of the signal $e_o = Kr \sin \theta$.

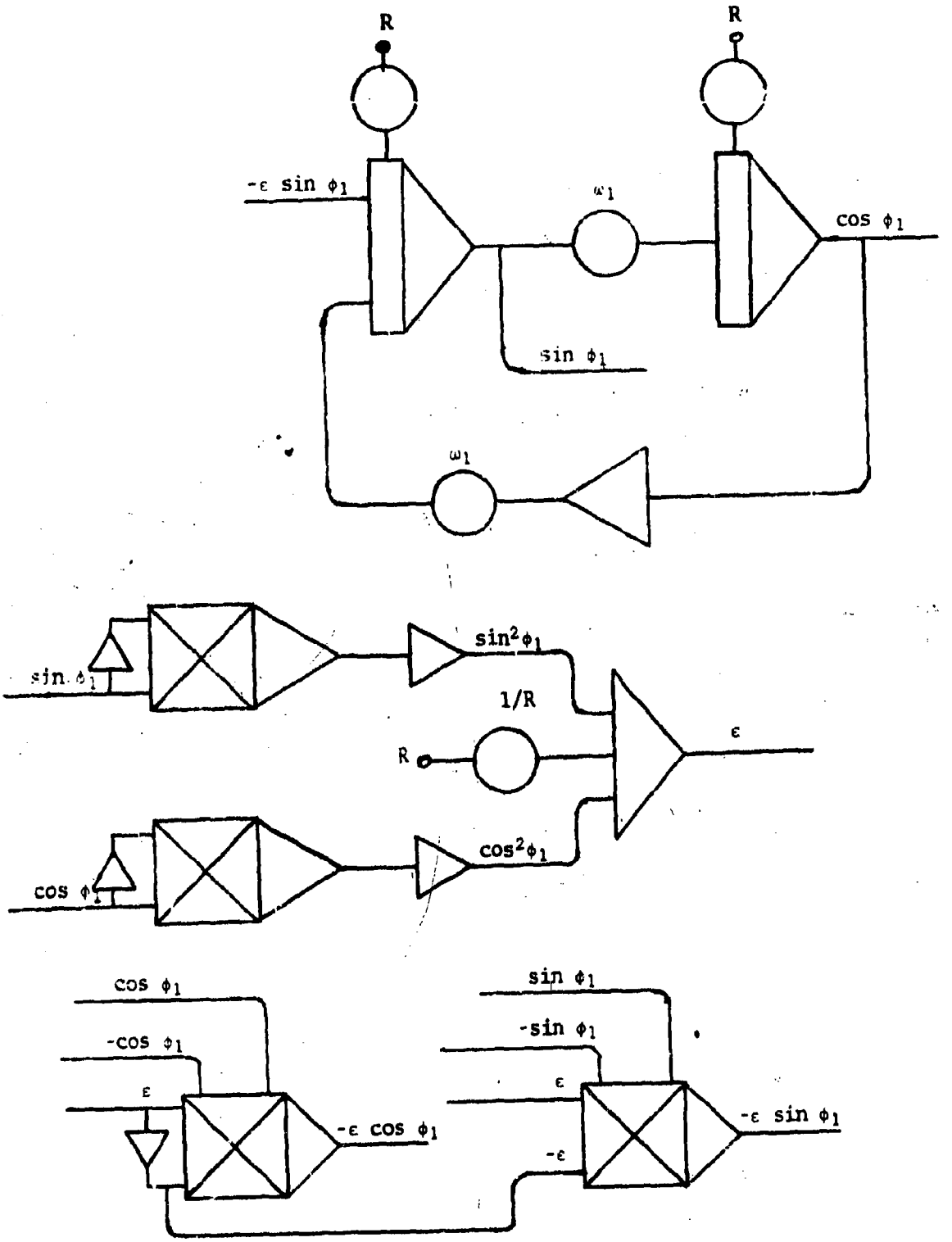
Simulation

The simulation of the gyro-seeker employs the mathematical equations as derived in the math model with the following exceptions:

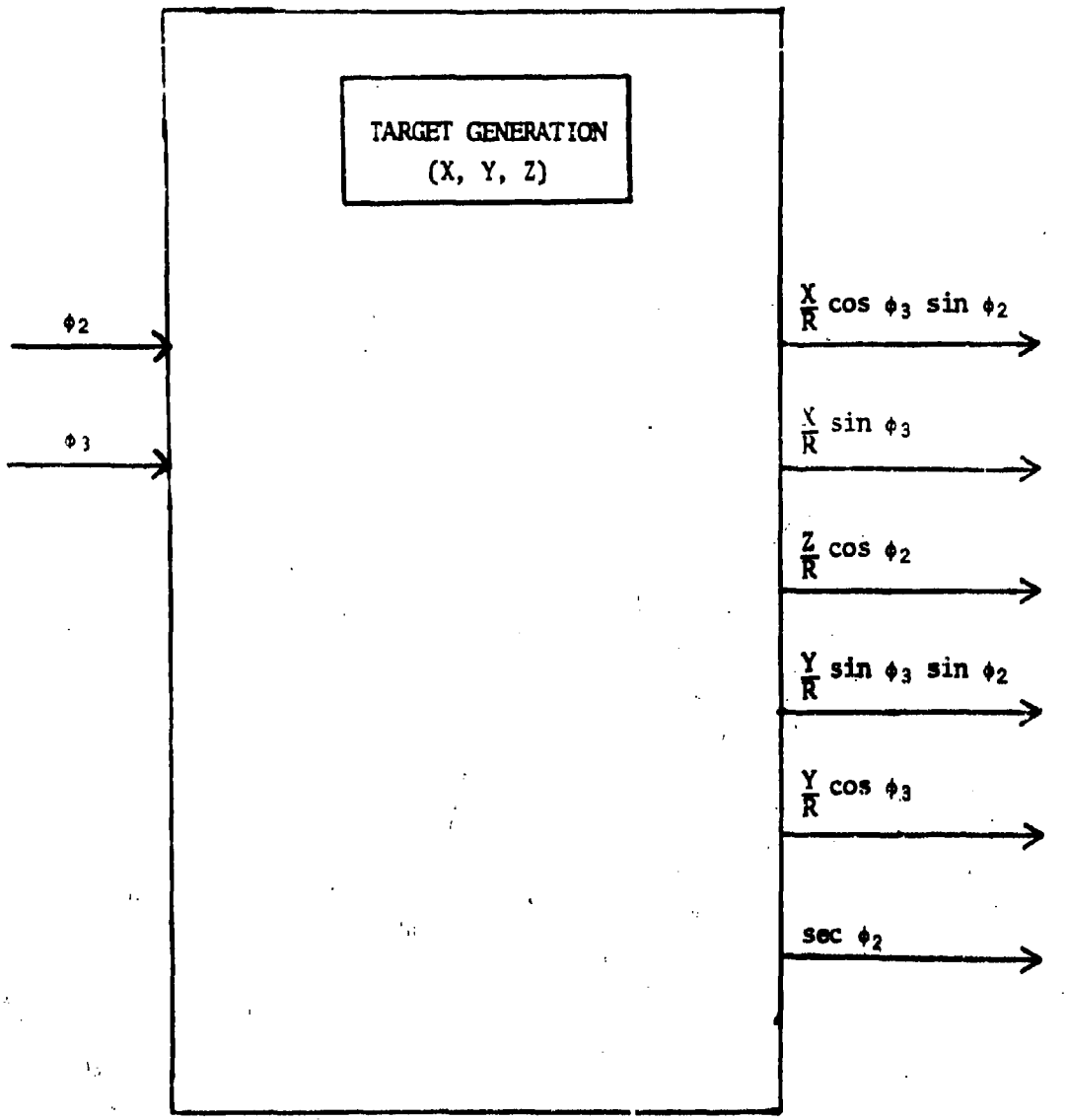
1. It was assumed that the gyro speed was constant. This is the case in a gyro when the motor-driving torque just balances the friction torques so the gyro spins at a constant rate. Thus sin ϕ_1 and cos ϕ_1 were obtained by running in oscillator at the required spin frequency.

2. The equation $\dot{\phi}_1 = \omega_1 + \sin \phi_2 \dot{\phi}_3$ was approximated by the equation, $\dot{\phi}_1 = \omega_1$, i.e., it was assumed that $\sin \phi_2 \dot{\phi}_3 \ll \omega_1$.

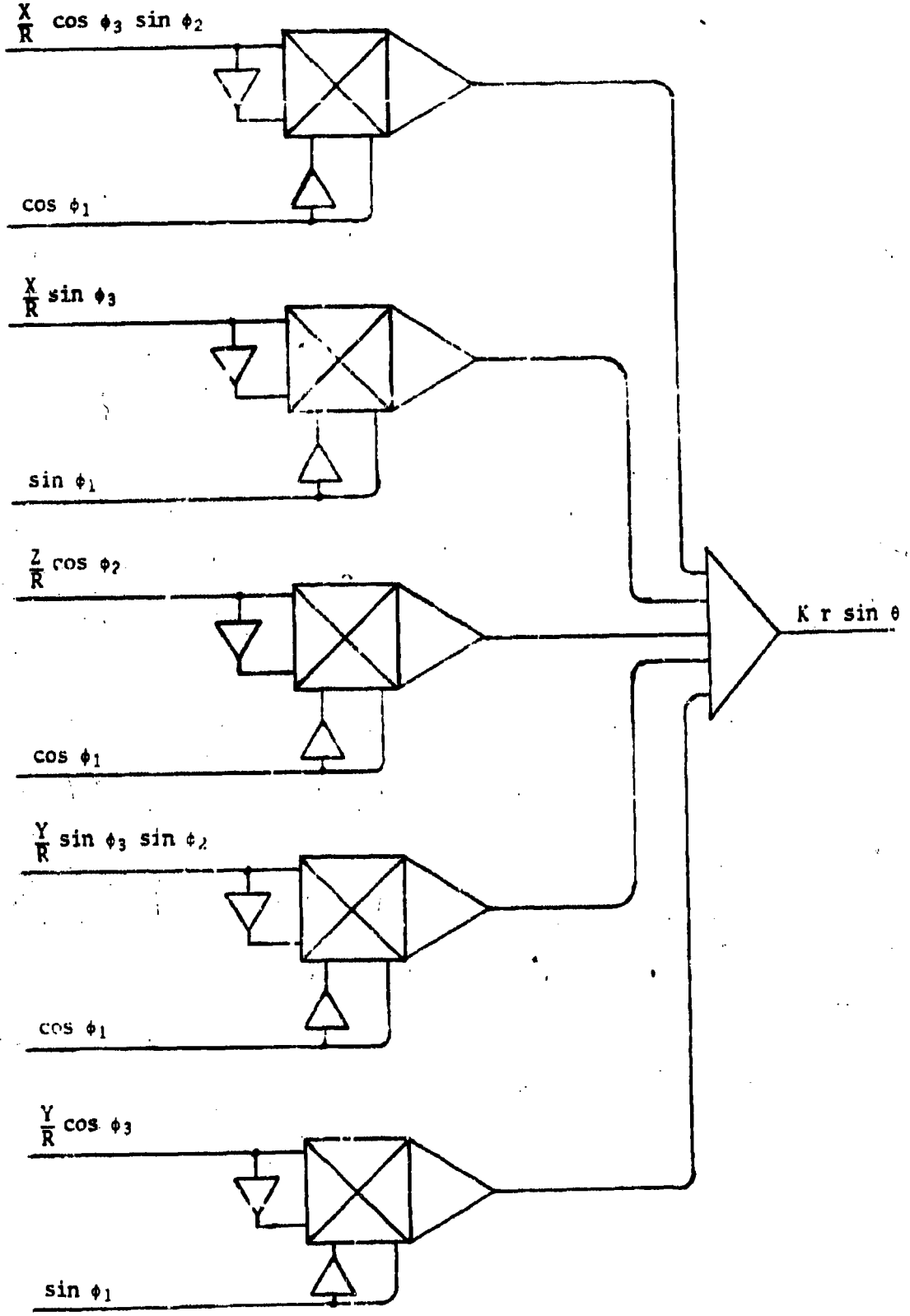
The computer mechanization diagrams are presented on pages 13 to 18. Since this particular mechanization is part of a hybrid missile simulation, some of the computations are shown as digital. It should be clear to the reader how on all-analog simulation could be obtained from the given mechanization. The parameters such as spin frequency, moments of inertia and loop gains were obtained from experimental data taken from the seeker hardware. After all parameters were obtained, the simulation was verified by comparing the response characteristics of the hardware with those of the simulation.

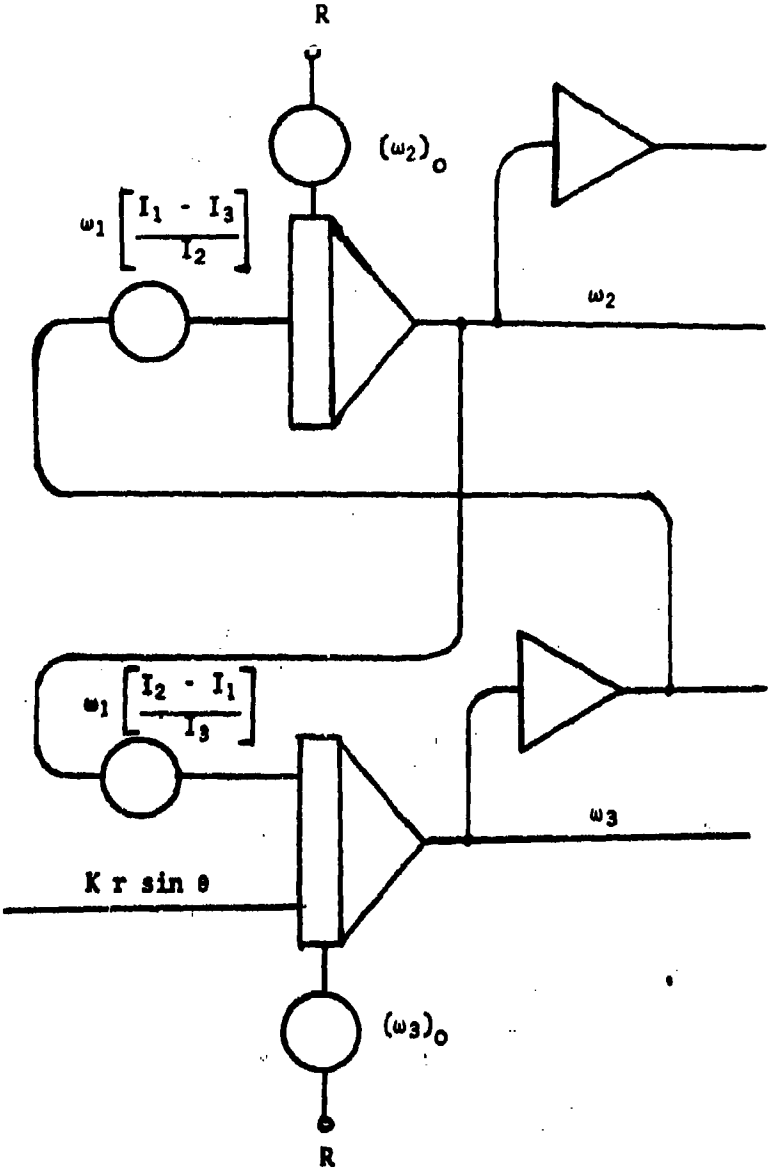


ω_1 REGULATED OSCILLATOR

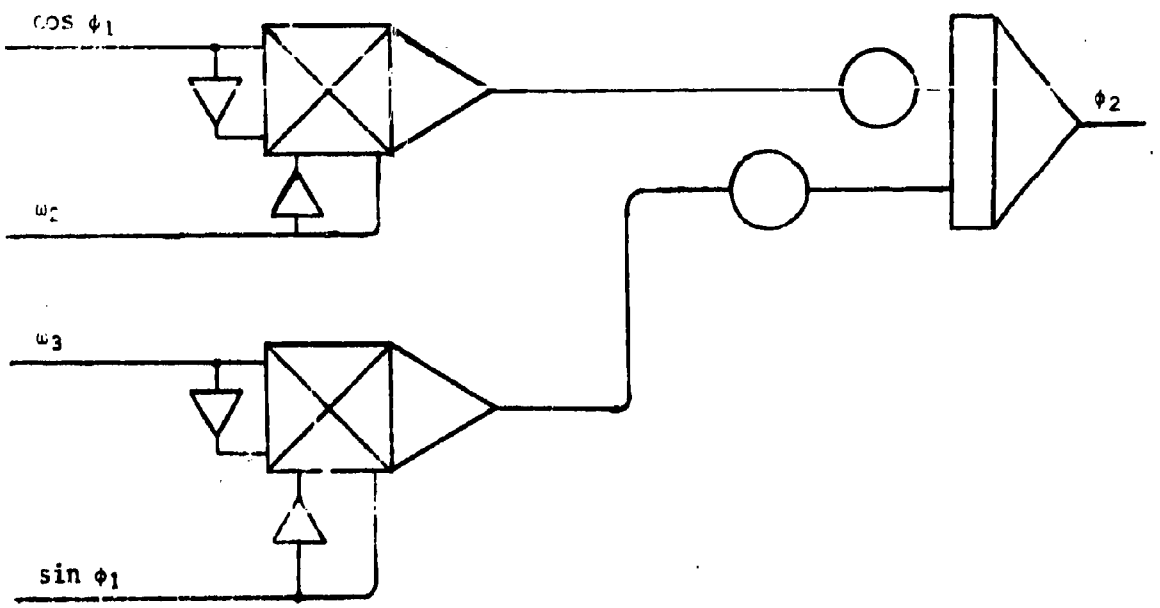


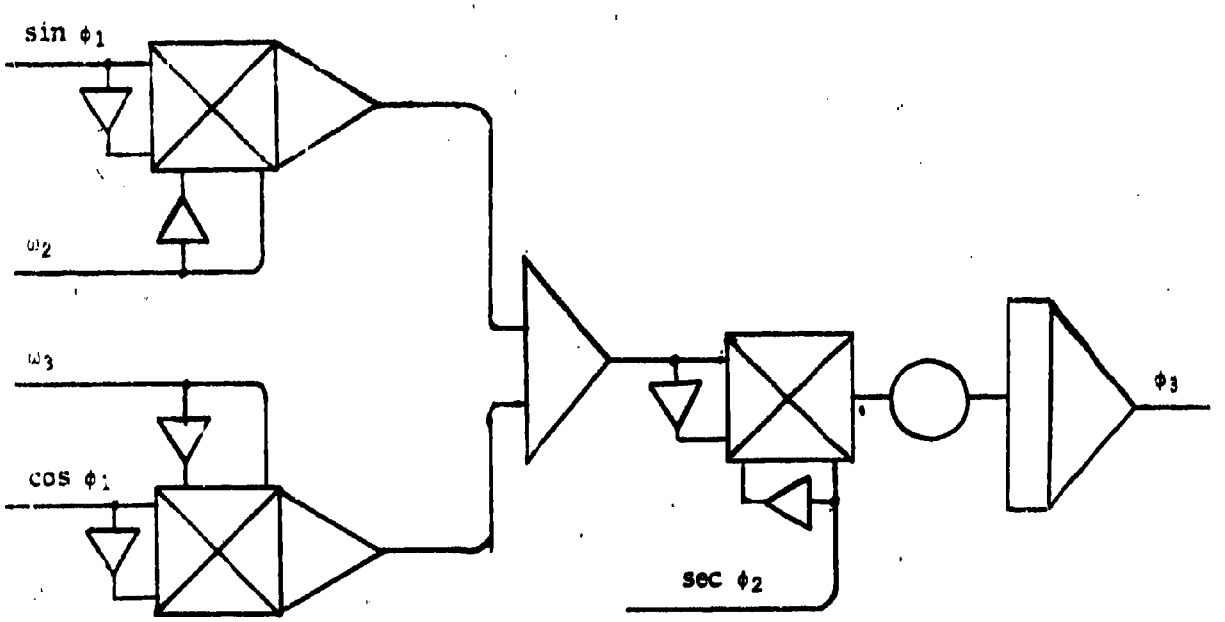
DIGITAL COMPUTATION

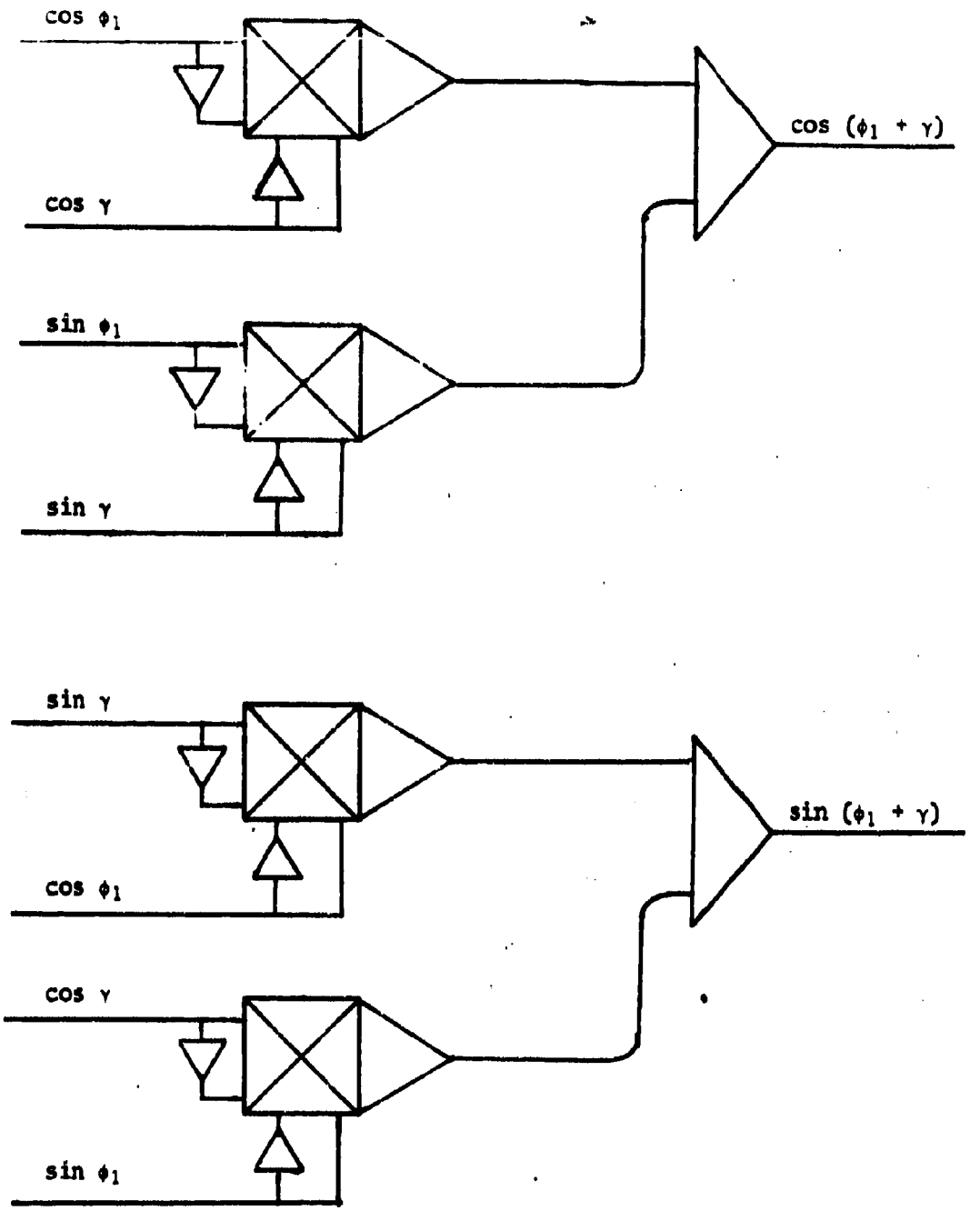

$\frac{X}{R}$ cos ϕ_3 sin ϕ_2
cos ϕ_1
$\frac{X}{R}$ sin ϕ_3
sin ϕ_1
$\frac{Z}{R}$ cos ϕ_2
cos ϕ_1
$\frac{Y}{R}$ sin ϕ_3 sin ϕ_2
cos ϕ_1
$\frac{Y}{R}$ cos ϕ_3
sin ϕ_1
K r sin θ



GYRO NUTATION OSCILLATOR


cos ϕ_1
ω_2
ω_3
sin ϕ_1
ϕ_2


sin ϕ_1
ω_2
ω_3
cos ϕ_1
sec ϕ_2
ϕ_3



REFERENCE COIL SIGNALS

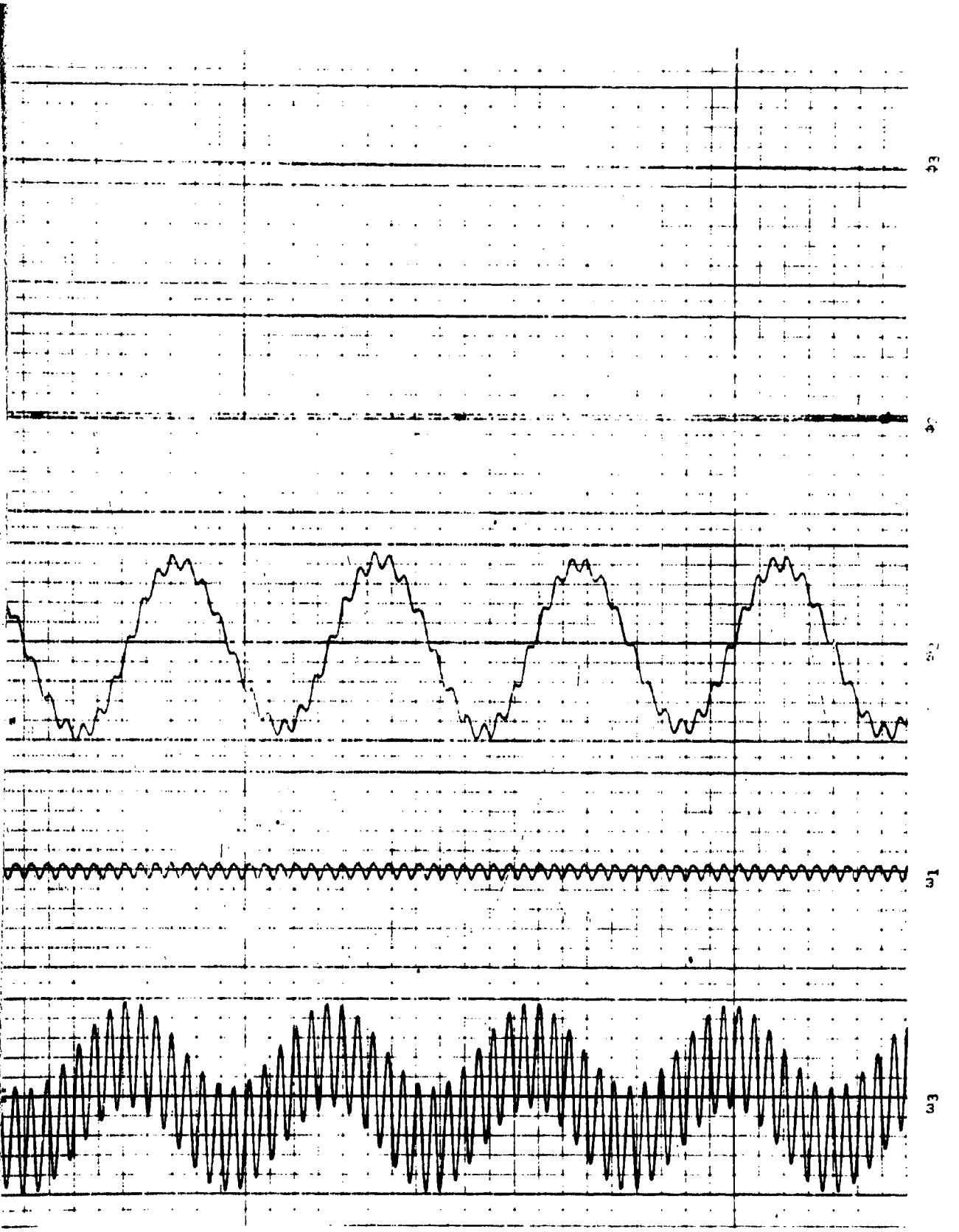
Some simulation recordings and description of results are presented on the following pages.

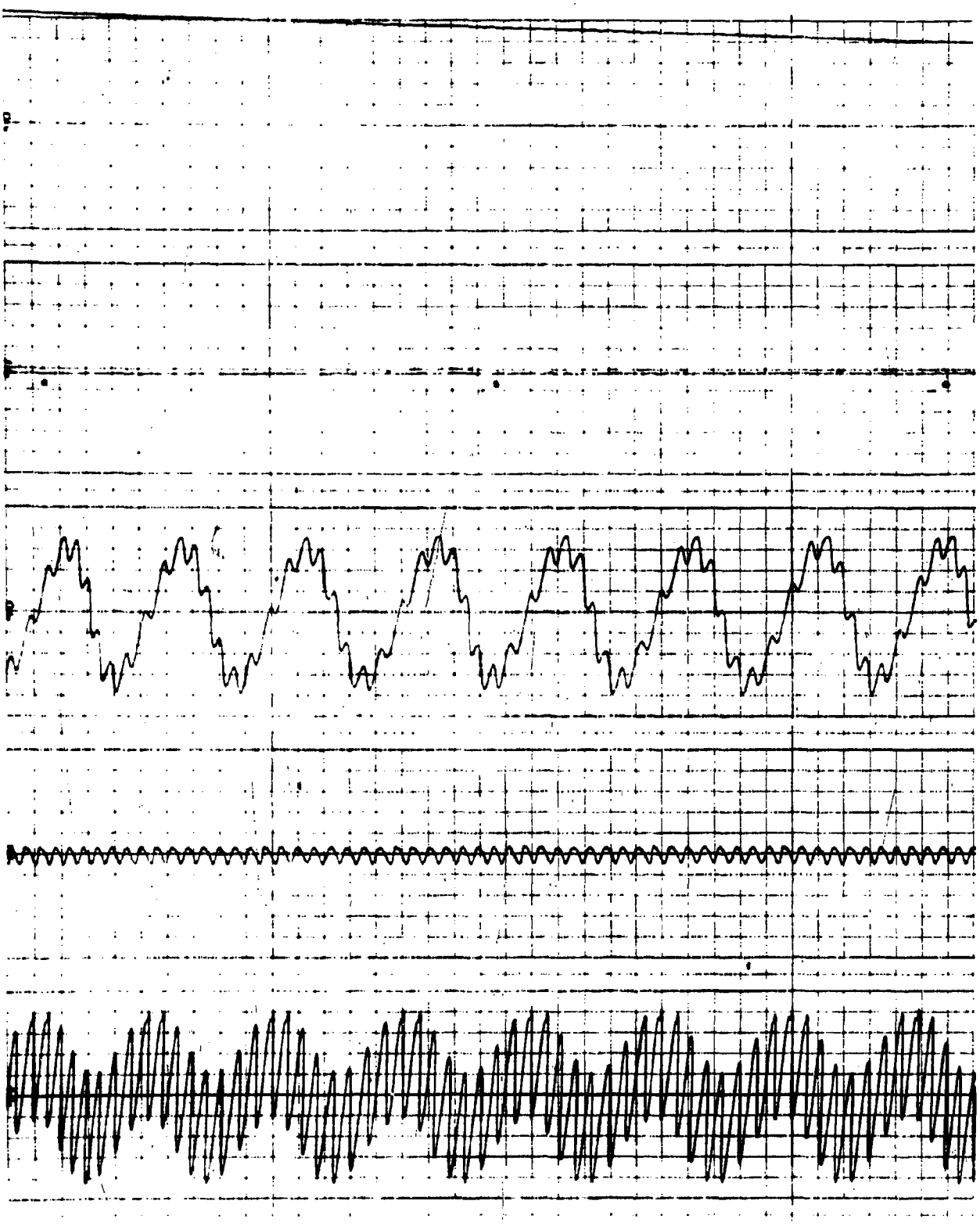
Pages 21 and 22 show the response of the gyro alone to a constant amplitude sinusoidal torque at spin frequency applied about the $\bar{g}_3$ axis and phased so as to precess the gyro in the ϕ_3 plane. The difference in the two recordings is the result of a change in the moment of inertia ratio $\rho = {}^{I_3}/_{I_1}$ which can be seen as a change in the nutation frequency on the ω_2 and ω_3 channels. The frequency of the nutation is determined by $\omega_1 (1 - \rho)$ while the amplitude depends on the initial conditions $(\omega_2)_0$ and $(\omega_3)_0$. (See, e.g., reference 3 for an analytical derivation). The gyro precession which is seen as a change in ϕ_3 is proportional to the amplitude of the applied torque.

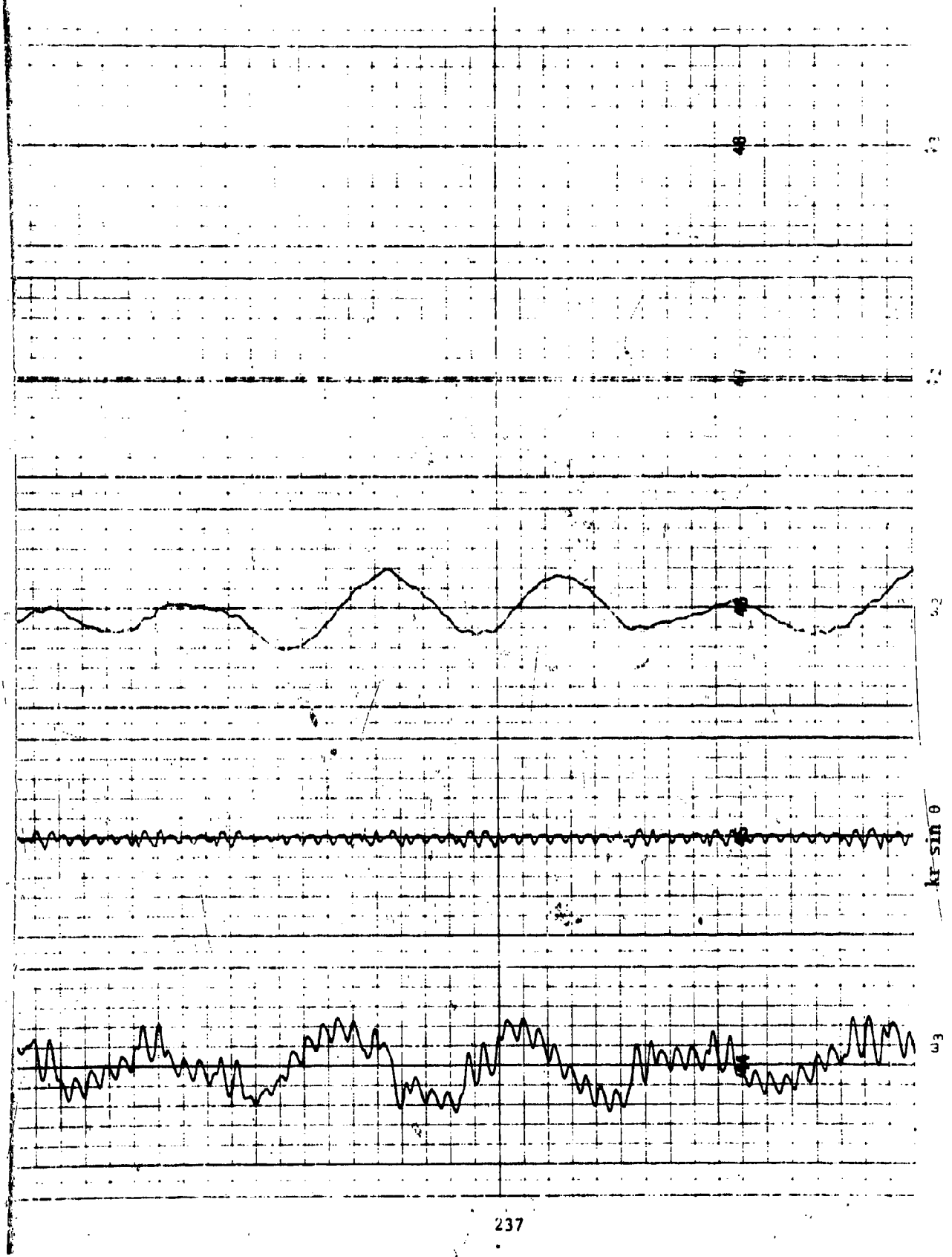
Page 23 shows the closed loop response for a given gain K_1 as a multiple of the feedback signal r sin θ. In this case the seeker was not tracking (as can be seen by ϕ_2 and ϕ_3), but was locked to a stationary target, X = C, Y = 0, Z = 0. Thus K_1 r sin θ has a small amplitude and a phase which is changing rapidly to compensate for directional changes of the spin axis from the line of sight. Page 24 shows the response to a target moving at a constant rate in the Y direction. This condition results in an error signal K_1 r sin θ which has a constant amplitude, reflecting the constant target rate, and a fixed phase dictated by the

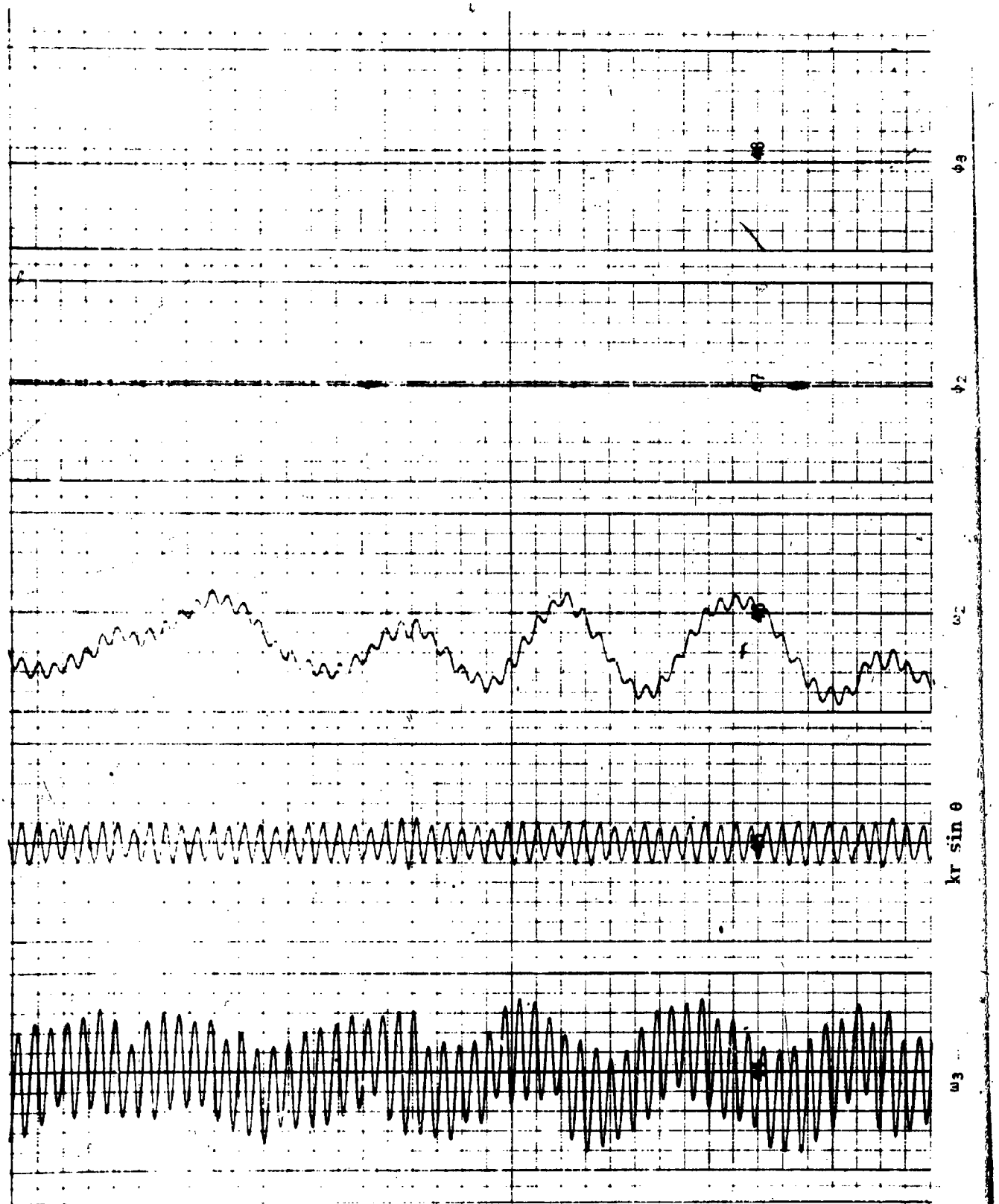
Y direction. (Note that since the torque applied to the gyro is about the $\bar{s}_3$ axis, it appears on the ω_3 channel.) The demodulated outputs of the error signal for target rates in the Y and Z direction are shown on pages 25 and 26 respectively. In this case γ was chosen as 0°, so that r sin θ was in phase with one reference signal and 90° out of phase with the other reference signal, for each tracking condition. Thus the perpendicular tracking directions result in alternate full value and zero value readings on the demodulated outputs R_1 and R_2 as shown on pages 25 and 26 . The slow rise of the R_1 and R_2 signals is due to filtering on the demodulated outputs and not to the demodulator circuits.

Error signals can also be produced by gyro motion such as gyro drift. If we assume a constant gyro drift rate and stationary target conditions then an error signal is produced to overcome the drift. Page 27 shows the simulation results for these conditions. In this case drift was produced by an appropriate torque in the ϕ_3 plane. The result of the error signal, Kr sin θ, generated by these conditions can best be observed on the demodulated outputs. One can think of these outputs as false tracking commands caused by gyro drift.

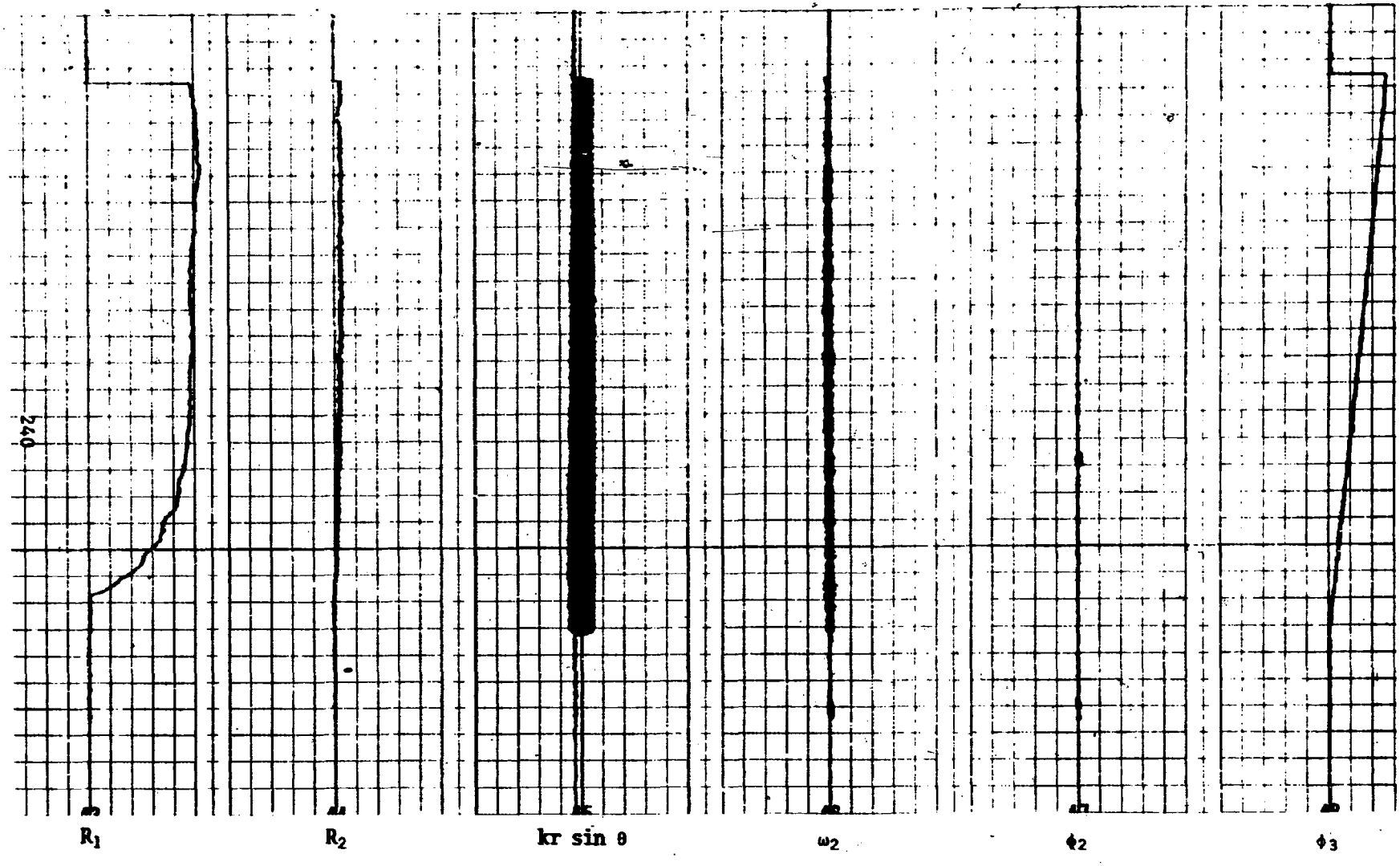

ω3
ω1
ω3

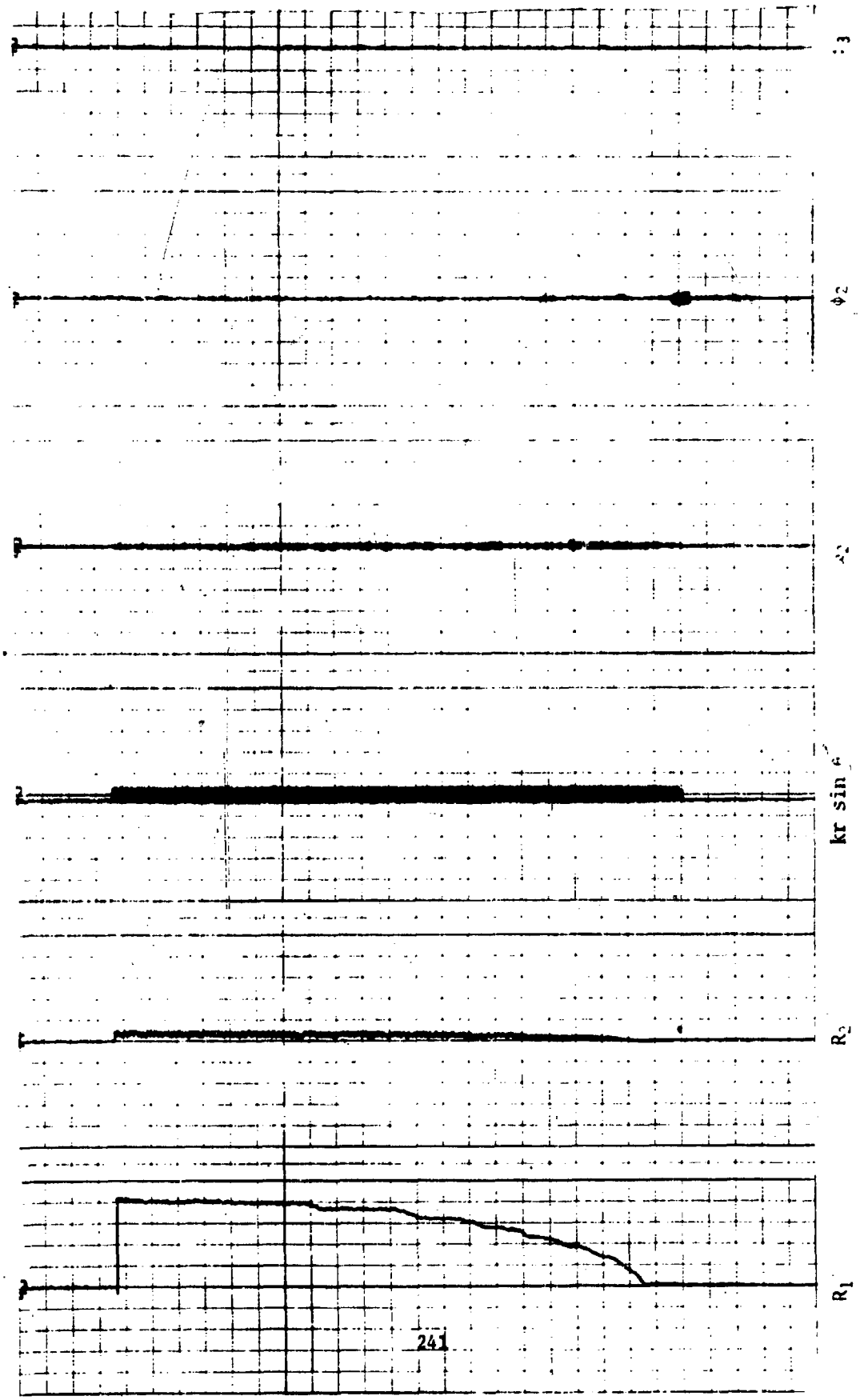



ω3
ω2
kr sin θ
ω3


ϕ_3
ϕ_2
ω_2
kr sin θ
ω_3


R1
R2
kr sin ϑ
ω2
ϕ2
ϕ3


240
R_1
R_2
kr sin θ
ω_2
ϕ_2
ϕ_3


ϕ_2
kr sin
R_2
R_1

REFERENCES

1. Mathematical Theory of the Sidewinder Gyro and Tracking Loop, Philco Ford, Oct 15, 1954

2. Tracking Accuracy of Infrared Trackers, Mengers, Paul E., General Electric Co., Ithaca, New York

3. The Gyroscope, Theory of Applications, Scarborough, J. B., Interscience Publishers, Inc. New York, 1958

4. Gyroscopes, Theory & Design, Savet, P. H., McGraw-Hill Book Co., New York, 1961

AN EMPIRICAL APPROACH TO ANALYSIS OF THE INTERACTION CHARACTERISTICS OF A SIX-COMPONENT ROCKET ENGINE TEST STAND

Aubrey W. Presson
U. S. Army Missile Command
Redstone Arsenal, Alabama

INTRODUCTION. LANCE Developmental testing within the T&RE Laboratory has been accomplished on test stands designed to measure the six components of thrust reaction. The basic problem inherent in such stands is that of restraining the engine with a measurement system that permits the engine six degrees of freedom, without the introduction of unknown effects upon the engine. This, of course, is further compounded by the requirement to supply propellants thru a high pressure plumbing system that shunts the measurement system. It is obvious that a thorough stand calibration program must be implemented to resolve this problem.

The facility requirements stated for the present phase of LANCE are presented in Table I. Paralleling these are our estimates of our then existing capability. These estimates were derived from extensive calibration tests performed on the original test stand. Only a casual observance is required to realize that this represents a significant step forward. A less casual but limited preliminary error analysis indicated more clearly the difficulties involved and concluded that a portion of the requirement was clearly beyond the state-of-the art. Briefly this is indicated, when load cell accuracy requirements are derived from consideration of the angularity and position requirements. Considering only the load cell capacities, dictated by the thrust magnitude and the practical geometry of the stand, the position requirement means a vertical load cell resolution of 0.05% F.S. and the angularity requirement a side load cell resolution of 0.3% F.S. The thrust and side force requirements are much less severe - approximately by a factor of 3 times.

This dilemma was resolved by a joint decision to proceed on a best efforts basis. It is this effort that I will summarize this afternoon.

II. PRINCIPAL DESIGN FEATURES. The present Thrust Measurement System, (Figure 1) like its predecessor system, contains seven load cells. This is the TP series stand with component convention illustrated. The four vertical cells are parallel and symmetrically placed about the vertical centerline or Z axis. These cells react thrust as well as the moments about the two horizontal axis designated MX and MY. Side forces designated X and Y are reacted by the horizontal cells placed on or parallel to these axes. The pair of cells paralleling the Y axis also reacts the roll moment (MZ). Each load cell is assembled with flexures to permit maximum compliance with all modes of loading except those acting directly along the cell axis. Thus, the magnitudes of force interaction inherent in the system are minimized.

Figure 2 illustrates additional features. These views are of the present system. Features to be noted are:

1. The unsupported long elbow sections of pipe in the propellant line approaches to the engine.

2. Arrangement of the three lines approaching the engine symmetrically about the vertical centerline in an attempt to balance the restraints.

3. Displacement of the plumbing from the centerline to permit application of single point calibration loads.

4. Inclusion of rupture disc housings (appearing as boxes between the lines and engine) that permit the pretest installation of these discs without disturbing the propellant line connections to the measurement system.

Load cell placement and alignment of the system and especially the alignment of the calibration input devices are of critical concern. Optical tooling was used to control these factors to precisions of better than .005 inches in position and .05 mil radians in orientation.

The basic calibration scheme developed for the original stand was first employed in the calibration of this stand. The premise of this calibration attempt was that there are many sources of interaction which combine to produce the net effect on the system. These include each of the components of force input, the static pressurization of the propellant lines, the dynamics of flow through these lines and the effects of temperature over the conditioned range of -40 to 160°F. Thus, each source was tested and its effects observed. It was further presumed that these independently derived effects could be summed to express the net effect. This approach is illustrated in the slide by the multiple calibration devices.

III. REVISED CALIBRATION METHOD. Time will not permit a detailing of the extensive calibration program by which it was determined that this latter premise was invalid. Perhaps it is sufficient to say that this led to a phase that is often referred to as an agonizing reappraisal. If the basic premise that there were no synergistic characteristics was invalid, how then could the stand be calibrated. Alternate schemes were considered but the one adopted involved the use of an existing program from the Computation Center files. This program - a so called Nonsimple Stepwise Multiple Linear Regression Analysis shortened herein to MLR Analysis was used to fit the data to empirical expressions of the input components in terms of the outputs. These expressions need be limited only by the capacity of the program which permits up to 59 independent terms. The scheme was attempted on data then available to us and a very close fit was obtained. This data did not cover the full ranges of

interest, however, as a recalibration plan was formulated and loading hardware was designed to permit the acquisition of data sufficient for this purpose. The plan was first implemented in January 1969 for ambient temperature condition and analysis of results were completed in March 1969. Due to a stand renovation, a repeat of ambient condition calibrations have been made as well as a calibration at each of the temperature extremes. Analysis is in progress on this data. The scheme developed for the January 1969 effort was used thru out and is the primary subject to be presented to this panel.

The loading scheme for this series of tests involved a series of input vectors whose locations, attitudes and magnitudes were closely controlled and/or measured. The input assembly, (Figure 3) starting at the "hard" point, involved a bi-directional translation device with planned displacements indexed by a series of dowel pin holes. This device was centered on the vertical centerline through the use of optical tooling. Upon this device was mounted a hydraulic jack which was linked to the input load cell through a universal flexure. A rod extended from the load cell to another flexure near the engine mounting fixture and another smaller bi-directional translation device connected this flexure to the mounting fixture. It, too, was indexed for the planned displacements. During loading operations, initial displacements were set with these translation devices to effect either vector displacement, vector angularity, or a combination of both. Then, to assure that any change in these initial conditions was known, displacement gages were used to monitor any lateral displacements above the upper flexure or below the lower flexure. The loads were then cycled under control of a servo loop to create the load sequence depicted in Figure 4. To combine the static pressure effects each of the four cycles were run at different line pressure conditions as noted. Four cycles of this type constituted a test run. The recalibration plan involved 30 runs.

Digital data acquisition and processing techniques were employed throughout the calibration process except for the displacement gages used to monitor the input rod attitude. The analog data obtained from these were manually reduced and entered into the digital analyses. Many of the features contained within the computer programs are illustrated in Figure 5. Automatic normalizing of all output data at the zero input point eliminates the negative thrust portions of our load cycles. Redundant data sections are also eliminated by an edit routine. The low range bridge of the input cell defines the sustain level sections indicated while the high range bridge defines the original boost level segments indicated by solid black lines. A range limiting feature that was subsequently added redefines the boost level to conform to the designated upper range segments.

For each of these segments, the 200 sample/sec acquisition rate was reduced by editing and averaging to approximately 1 sample/sec for boost data and 10 s/s for sustain data. Computation of input and output components were then made at these data rates and taped for input into

the MLR Analysis. Inputs were based on initial and monitored change of rod attitudes while outputs were functions of the stand geometry and the load cell results. Tabulations and CRT type plots were prepared to aid in analysis.

The MLR Analysis determined the empirical equation coefficients for each component in turn. Data from 25 of the 30 calibration runs was considered collectively. The determined coefficients were then programmed to provide predictions of the input components which were then compared to the known inputs and a residual error computed. When applied to runs 26 thru 30, a validity check was made of the total process. When applied to all 30 runs in turn a standard error is obtained for each run. It is the combination of these errors that determined our estimate of measurement system capability.

Several forms of the MLR Analysis were performed as indicated in Table II. To understand this table, you must realize that these are the stand output parameters that are considered as independent variables in each of six equations involving the known input values or dependent variables. The equation form, as noted, is Input Component = $\Sigma\ A_n T_n$ with A_n the unknown coefficients and T_n the terms selected from the chart. The analysis is performed on each of these equations in turn to determine a best fit set of unknown coefficients, one for each of the independent terms. Having determined these values by calibration, it is assumed that the equation holds for unknown input conditions and thus these conditions can be predicted from any set of output values contained within the calibration ranges. The original premise was that irrelevant terms would be effectively eliminated by the analysis and a 59 term form was chosen. The highest order terms used were second degree terms and their cross products. This provided a very good fit for the 25 runs used in the analysis but prediction for runs 26 thru 30 were very poor. It was reasoned that this was due to the inclusion of too many irrelevant terms and the equations were reduced to a 16 term form. (Diagonal shading.) All cross products are eliminated. This produced significantly improved prediction results even though the standard errors from the MLR Analysis were increased. This form was adopted for the sustain range of data. An addendum note should be made at this point. In setting the program controls, no test was made on the exclusion of terms. Present analysis includes this feature but its effectiveness is not yet known.

The next significant step was taken when it was observed that even though the standard error for an over-all fit was acceptable, the fit at the upper limit of the range was at times unacceptable. This led to a revised analysis based on data contained within the interval of 25 to 40K. The 16 term form was selected for this analysis. The residual error plots seemed to suggest a correlation with the ratios of Z input with each of the components. Thus, two forms containing 23 terms in the equations were attempted next. These forms include the 16 terms

and those shown in dots in Table II. Alternates shown involve the products or ratios with "Z". Results were only slightly improved in terms of the standard errors obtained, however the residual error plots pointed-up a preference for either of these over the 16 term form for most of the six components. This will be demonstrated in future figures. Attempting further improvement, two additional forms with 33 terms were used. The first attempt added the terms identified with horizontal shading. The second was broken into two sub-analyses - the 16 term form plus an analysis of the residual errors in terms of the balance of 17 terms. Significantly improved standard errors were obtained with the first of these but again the results of both produced poor predictions for the confirmation runs. A comparative charting of the five forms as they apply to runs 26 through 30 are shown in Table III. The best equation results are in shaded areas for each component. The 16 term form for Y and MX; the 33 term for MZ and the 23 term ratios for the balance. The graphs for Run 27 which are typical of the runs examined will illustrate further the reasoning for our preferences. These are residual error vs. input plots from the 16 term analysis with data points plotted. A computer determined best fit is line plotted thru these points and appears as a broken line. Similarly determined best fit plots for each of the 23 term forms have been manually transferred to these plots for comparison. Comparison should be made between the line plots. The ideal result would appear as a zero error through out the range. Figure 6 is the X component plot. The 23 term results are obviously best with maximum error for the mean data fit of approximately 3 pounds. Figure 7 is the Y plot with a maximum error of 8 pounds. The Z plots (Figure 8) show a slight preference for the 23 term ratio form. All data is under 50 pounds of error in the presence of inputs ranging to 40K of thrust, 700 lbs. of side force and 2000 ft-lbs of moments. The MX plots (Figure 9) show a mean error of less than 30 ft-lbs. The MY plots (Figure 10) show error ranging to 100 ft-lbs for this run but this is somewhat larger than the normal MY error. The MZ plots (Figure 11) show little preference between these three forms but the standard error is less than 4 ft-lbs for all forms.

The foregoing description suggests a much more direct path than that of our actual experience. From an experiment standpoint there were several areas that could have been explored in greater depth but were ignored since this was primarily an attempt to define the specific capabilities of the stand as they related to the LANCE facility requirements. A final summary of the attainment of this goal is presented in a comparison of the facility requirements with the demonstrated capability (Table IV). Based on these results and our continued surveillance of the stand through verification testing, we are confident that our data quality must be rated equal to or superior to that obtainable at any facility of this type.

A concise summary of the major points of the calibration problem, the implemented solution and the clinical questions to be asked of this panel is now in order.

THE PROBLEM. Since it is desired to resolve a force vector in six degrees of freedom it is necessary to evaluate the interaction effects of the electromechanical measurement system in six degrees of freedom. For this purpose an experiment is required that will adequately measure these effects throughout the range of interest and determine the precision characteristics of the total measurement process.

THE IMPLEMENTED SOLUTION.

A. An experiment of 30 input vectors involving differing combinations of the six components of force input and static pressurization levels with ranging of each parameter to near maximum expected value.

B. An evaluation of the six empirical "best fit" transfer equations relating the observed input and output data. The terms of the equations were arbitrarily selected and a coeffients derivation made by a multiple linear regression analysis.

C. A derivation of precision limits based on the combination of the overall experiment data fit precision with standard deviation values for the laboratory standard and for the transfer to the field standard. This combination is by the square root of the sum of variances method.

THE CLINICAL QUESTIONS.

1. Is method valid?

2. If so, are the thirty tests excessive or inadequate?

3. Are more practical methods known?

TABLE I. REQUIREMENT VERSUS CAPABILITY ESTIMATE AT START OF XRL PROGRAM

Parameter	Requirement (3 sigma	Capability Estimate (95% confidence)
Side forces (lb)	± 50	± 105
Sustain phase thrust (lb)	± 160	± 100
Boost phase thrust (lb)	± 200	± 250
Boost phase vector location (in.)	± 0.03	± 0.125
Boost phase vector angularity (mrad)	± 0.372	± 2.6

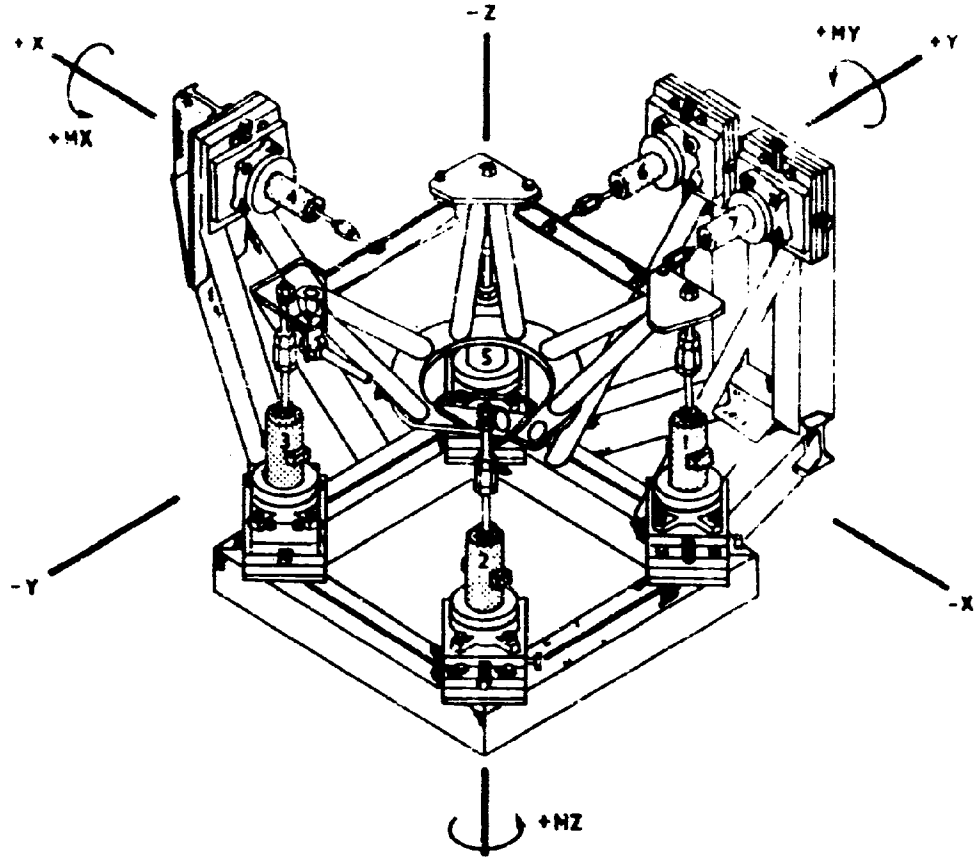


FIGURE 1. FORCE MEASUREMENT SYSTEM FOR TECHNICAL PROTOTYPE LANCE WITH DESIGNATION OF COMPONENTS CONVENTION

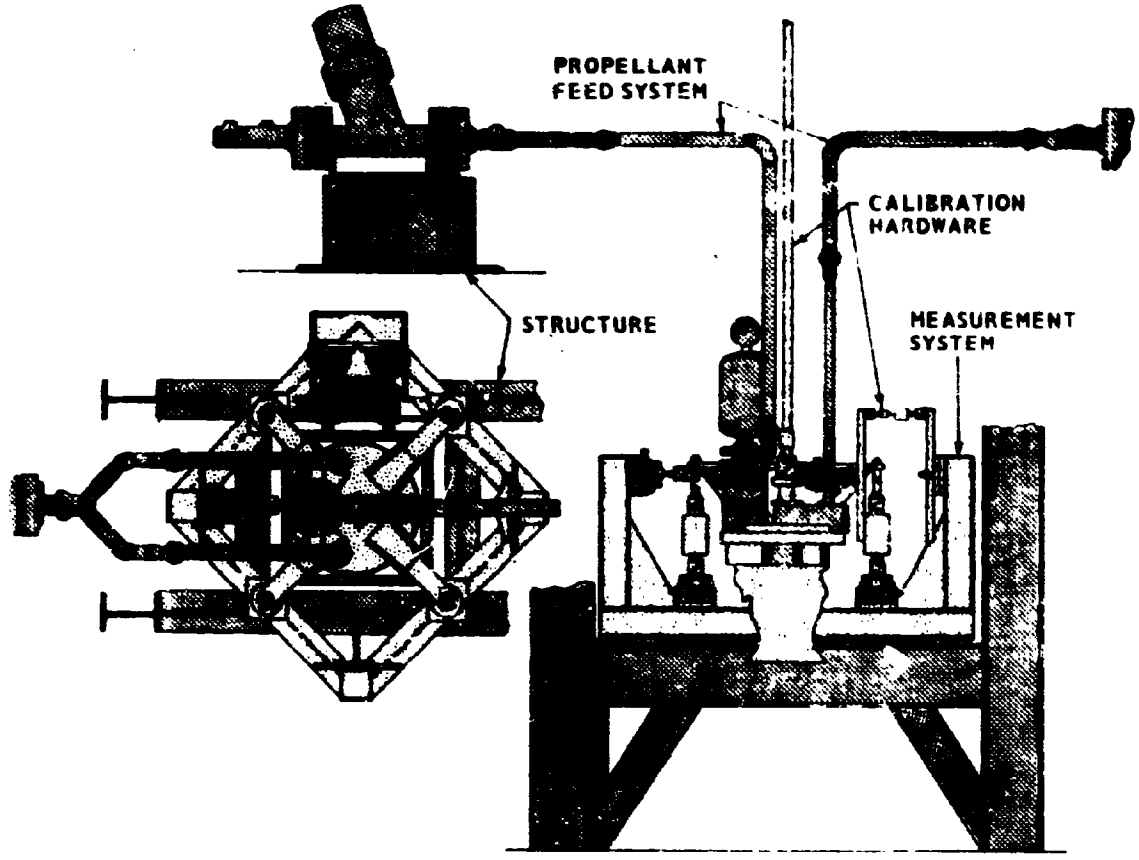


FIGURE 2. PRESENT FORCE MEASUREMENT SYSTEM

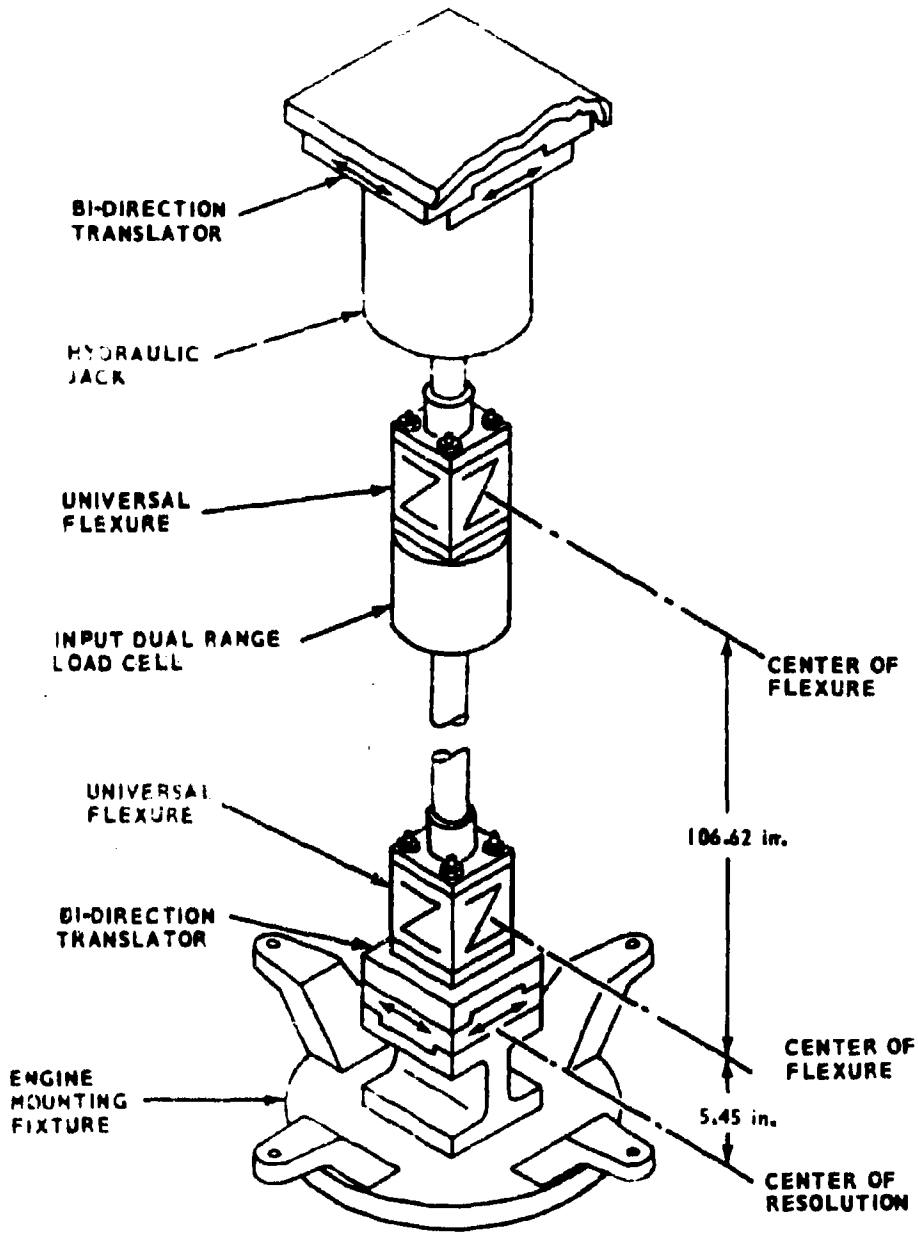


FIGURE 3. CALIBRATION FORCE INPUT ASSEMBLY

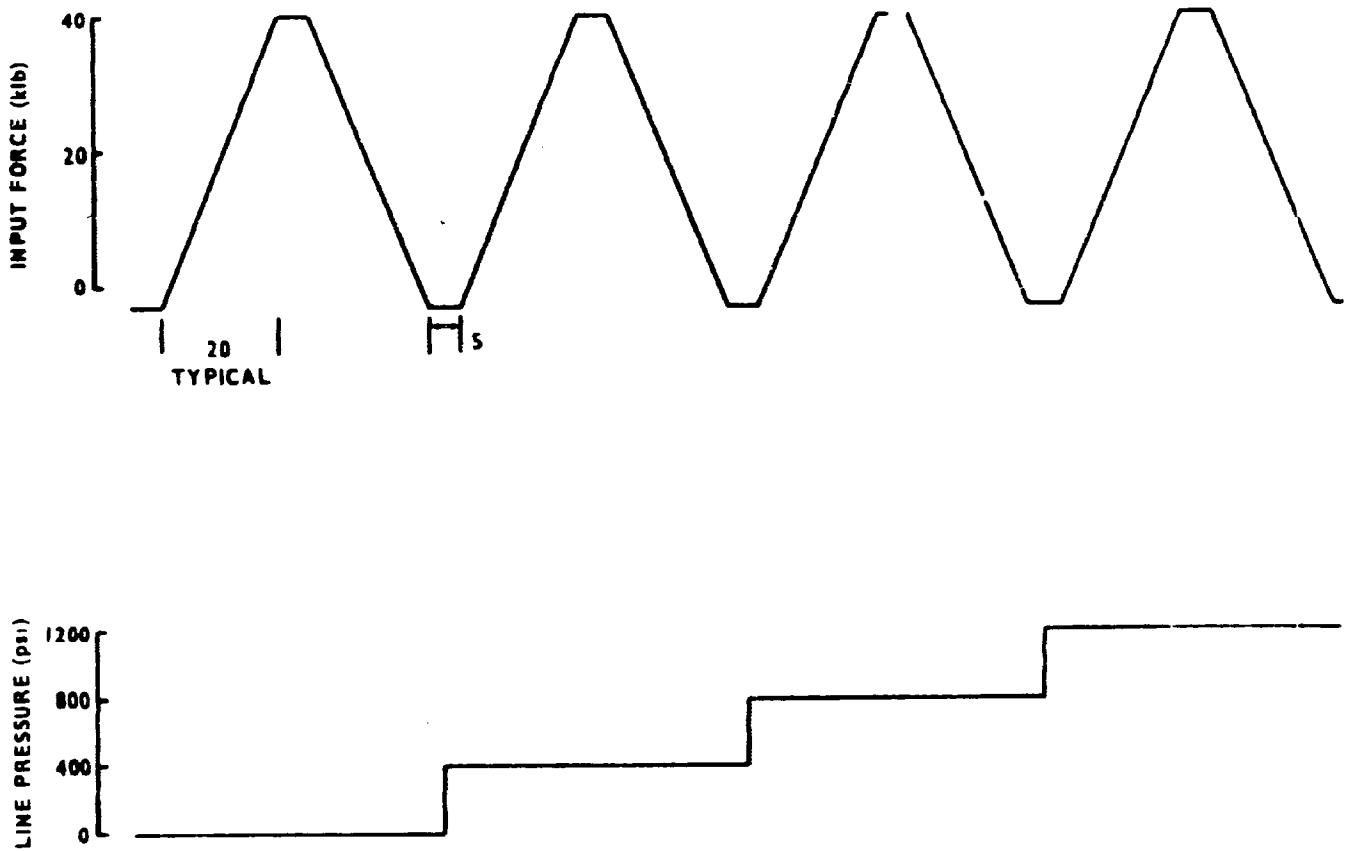


FIGURE 4. CALIBRATION TEST RUN LOADING SEQUENCE

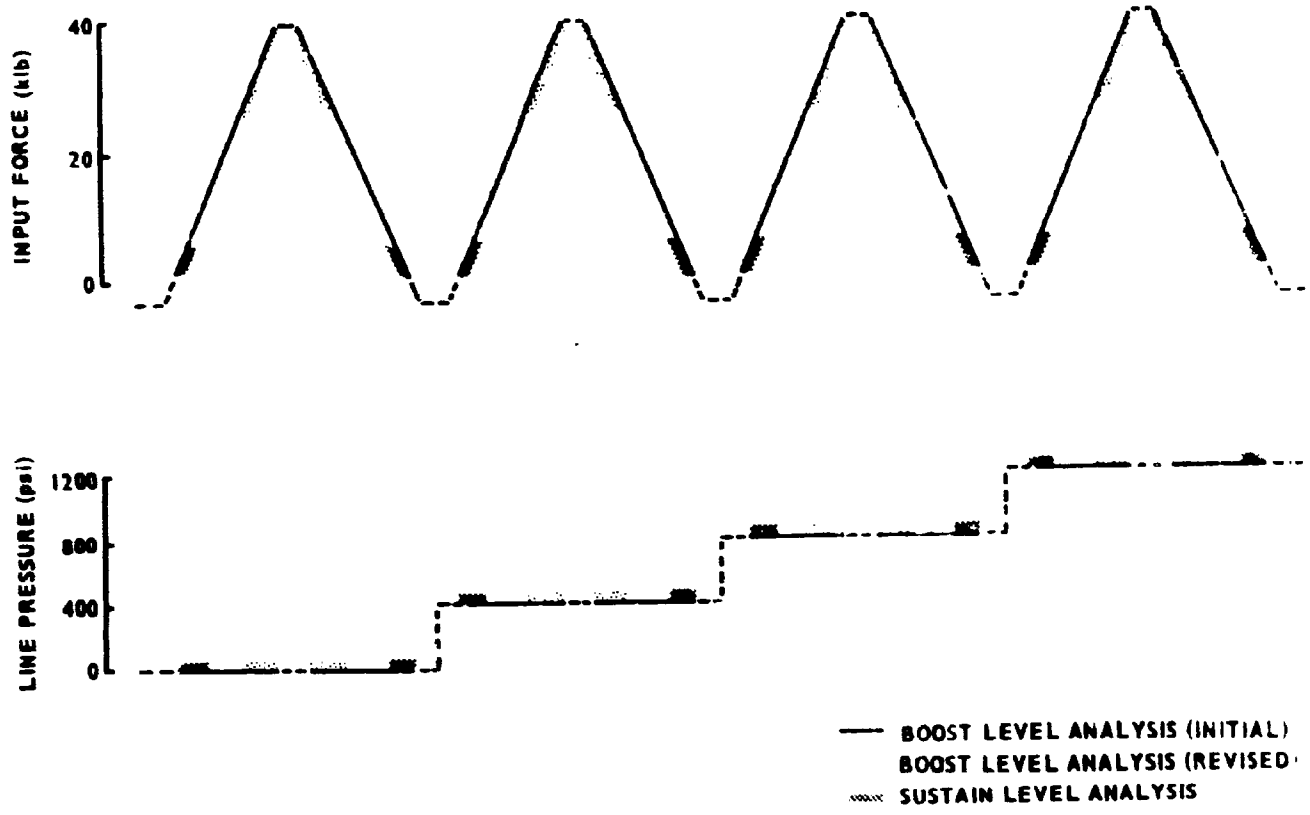
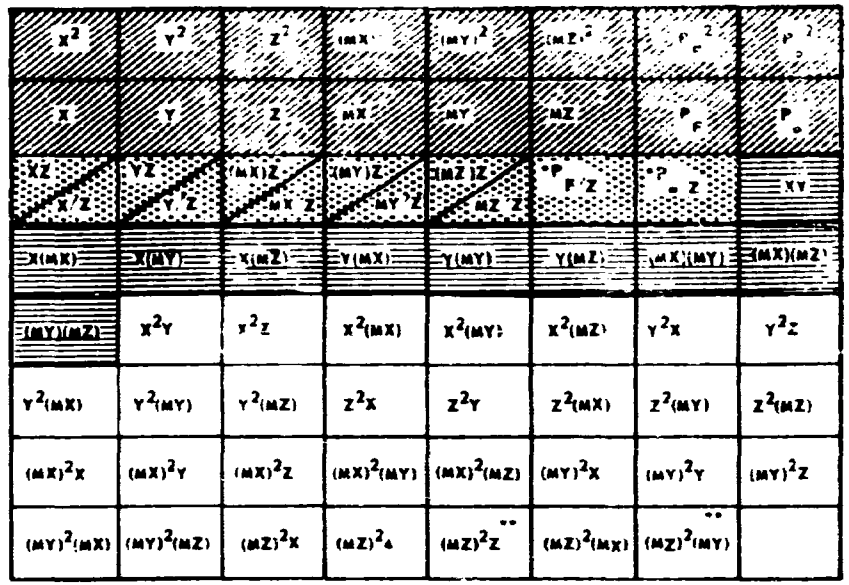


FIGURE 5. SEGMENTS OF LOADING SEQUENCE CONSIDERED BY ANALYSIS PROGRAM



* NOT INCLUDED IN 59 TERM FORM
**DELETED TO PERMIT INCLUSION OF P_F^2 AND P_o^2

a) OUTP[illegible]URCE COMPONENTS ARE X, [illegible] MX AND MZ
b) [illegible]AT PRESSURES ARE P_o AND [illegible] AND UNIFORM RESPECTIVELY
c) TERMS CONTAINED IN 16-TERM FORM ARE BLOCKED IN [illegible]
d) TERMS CONTAINED IN 23-TERM FORM ARE BLOCKED IN [illegible] WITH ALTERNATE TERMS SHOWN
e) TERMS CONTAINED IN 33-TERM FORMS ARE BLOCKED IN [illegible]
f) EMPIRICAL EQUATIONS ARE

INPUT COMPONENT = $\Sigma A_n T_n$

WITH A = UNKNOWN COEFFICIENTS
T = TERMS DEFINED IN CHART

g) THE MLR ANALYSIS DETERMINES VALUES OF A [illegible] THAT PROVIDE A BEST COMPROMISE FOR A DATA CONSIDERED IN ANALYSIS

TABLE II. VARIOUS FORMS OF MLR ANALYSIS CONSIDERED

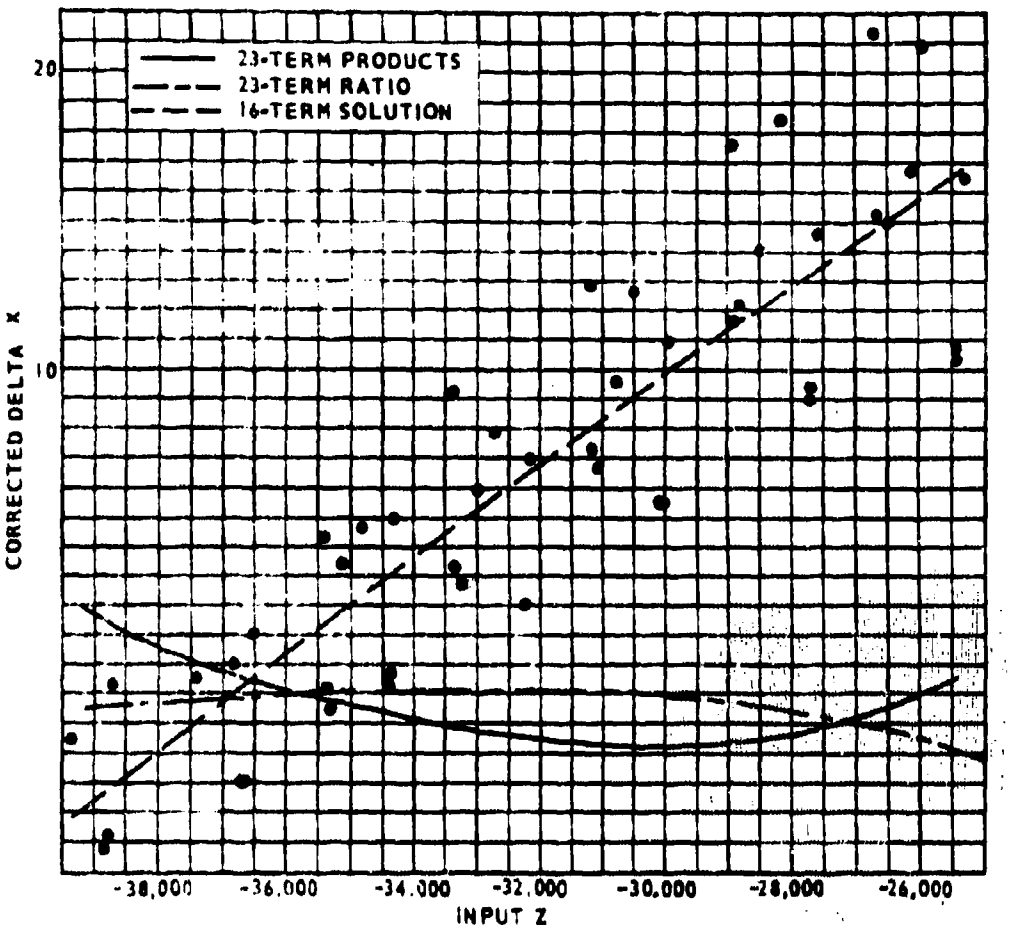


FIGURE 6. RESIDUAL ERROR PLOTS, X COMPONENT

EQUATION FORM	X COMPONENT (lb)					Y COMPONENT (lb)					Z COMPONENT (lb)				
	RUN 26	RUN 27	RUN 28	RUN 29	RUN 30	RUN 26	RUN 27	RUN 28	RUN 29	RUN 30	RUN 26	RUN 27	RUN 28	RUN 29	RUN 30
16-TERM	6	10	10	10	15*	12*	5*	9*	21	14*	18	26	14	26	19
23-TERM PRODUCTS	3*	7	9	7*	17	12*	8	9*	22	15	16*	28	11*	14	34
23-TERM RATIOS	3*	6*	9	7*	17	12*	8	9*	22	5	16*	25*	11*	14	35
33-TERM	21	16	5*	8	20	13	21	13	19*	16	24	79	20	15	44
16 AND 17	24	25	18	8	17	22	29	19	22	16	39	59	20	41*	35
	MX COMPONENT (ft-lb)					MY COMPONENT (ft-lb)					MZ COMPONENT (ft-lb)				
16-TERM	34*	27	70*	79*	100*	21	55	34	33	47	3.5	3.5	2.4	4.1	4.1
23-TERM PRODUCTS	35	25*	70*	79*	102	17*	82	31*	30	45	3.6	1.9	2.2	4.1	4.0
23-TERM RATIOS	35	25*	70*	79*	103	17*	83	31*	29*	45	3.5	2.1	2.2	4.1	
33-TERM	233	256	219	83	119	43	18*	59	31	37*	0.8	0.9*	0.4*	0.6*	0.6*
16 AND 17	216	237	139	84	119	20	74	43	37	40	0.6*	10.6	3.5	3.6	

*BEST RESULTS FOR TEST RUN INDICATED
"BEST EQUATION" RESULTS IN SHADED BLOCKS

TABLE III. STANDARD ERRORS RESULTING FROM M.I.R. PREDICTIONS, BOOST RANGE 25 TO 40 KILOPOUNDS

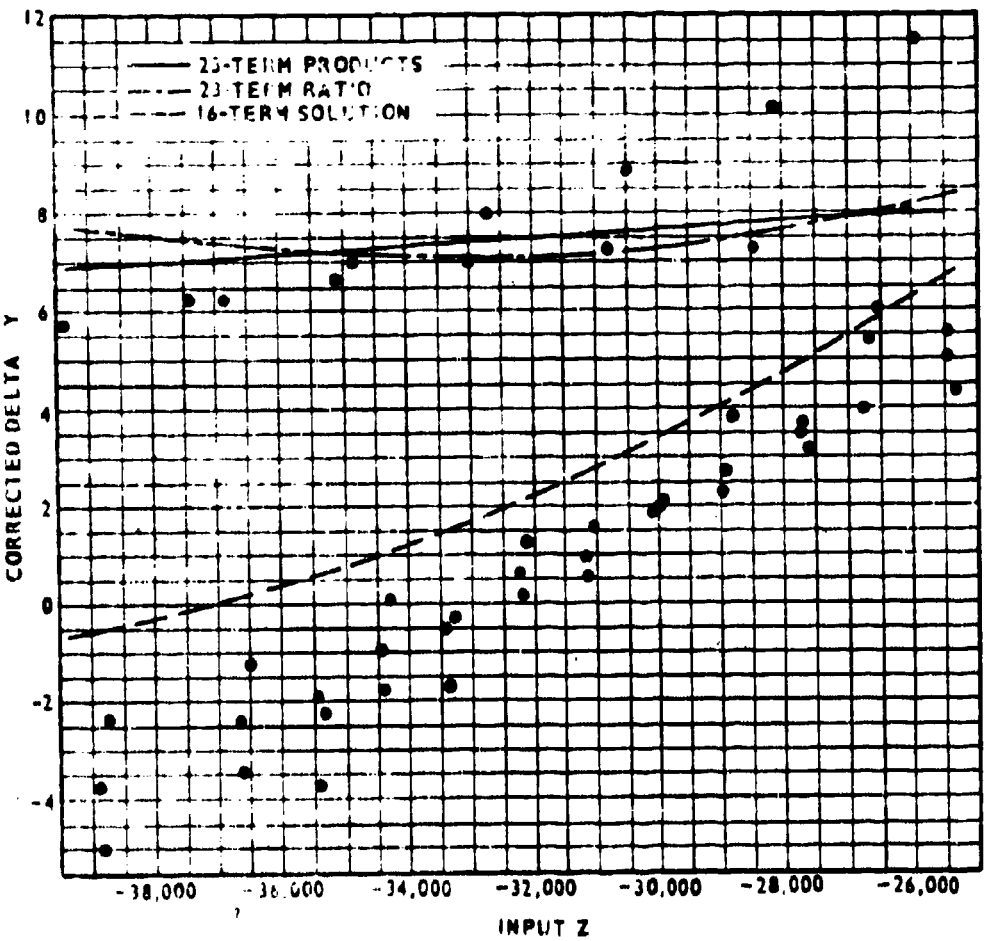


FIGURE 7. RESIDUAL ERROR PLOTS, Y COMPONENT

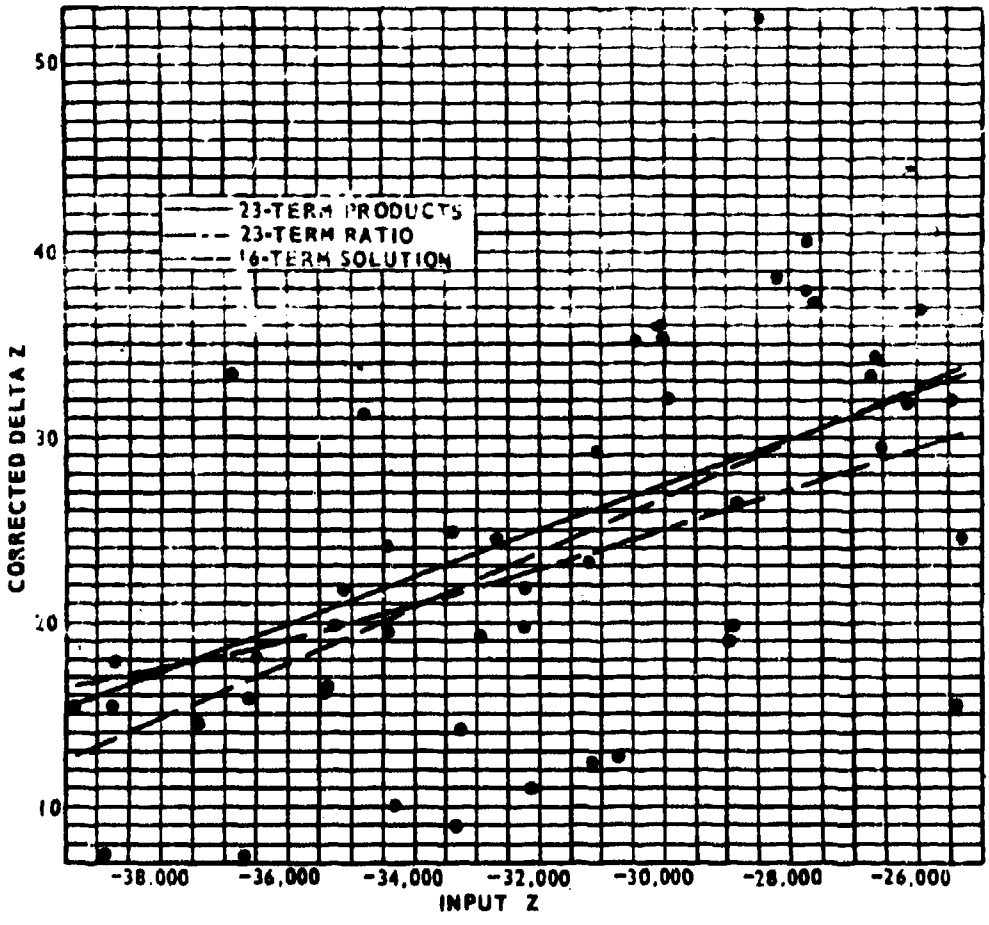


FIGURE 8. RESIDUAL ERROR PLOTS, Z COMPONENT

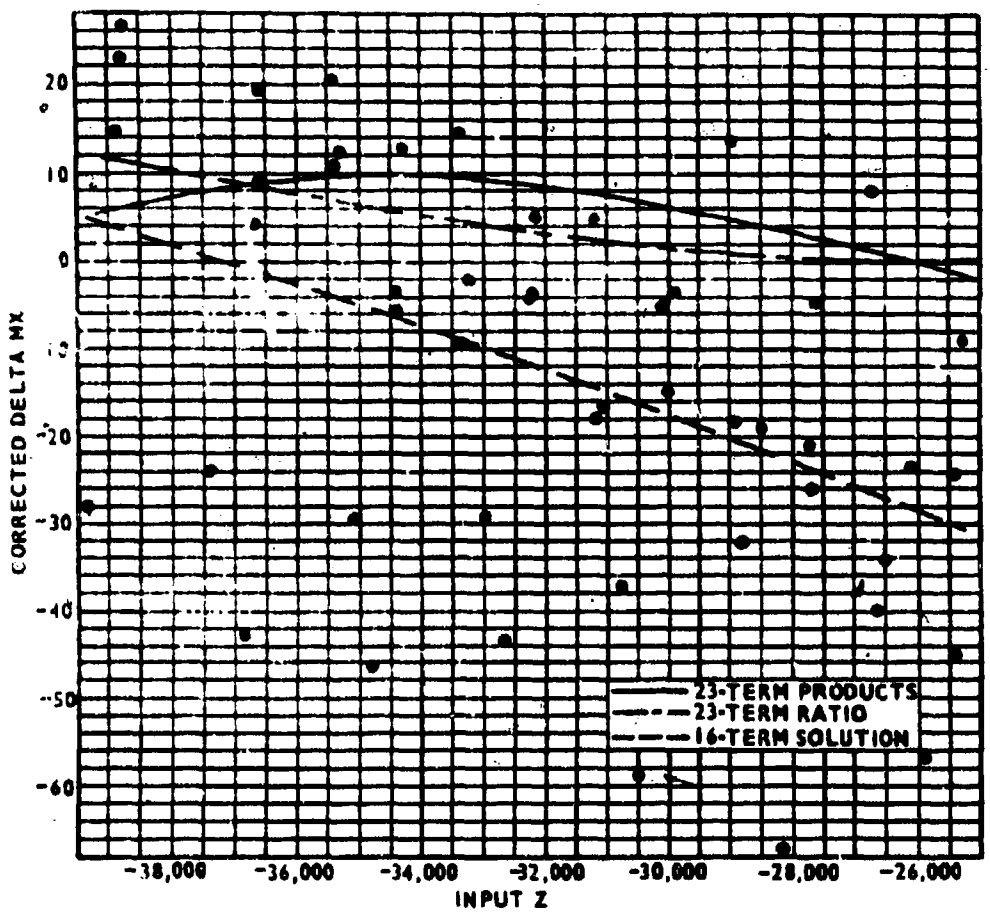


FIGURE 9. RESIDUAL ERROR PLOTS, MX COMPONENT

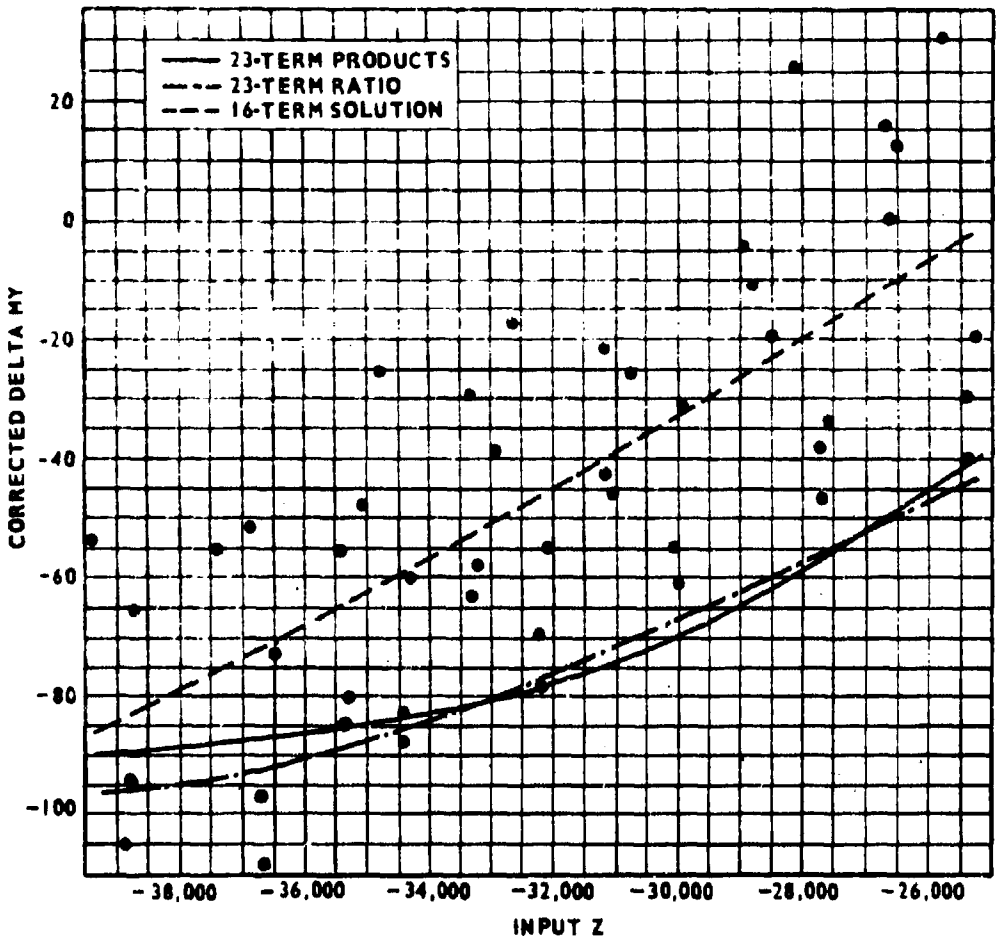


FIGURE 10. RESIDUAL ERROR PLOTS, MY COMPONENT

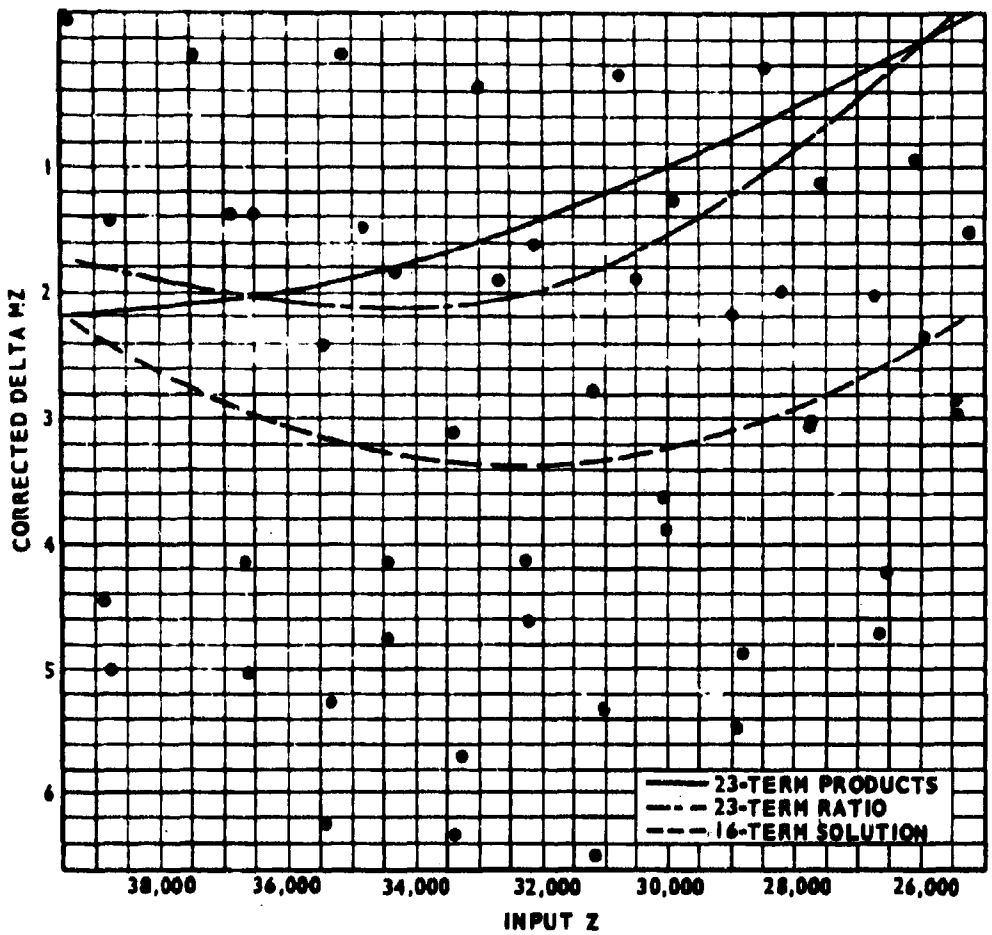


FIGURE 11. RESIDUAL ERROR PLOTS, MZ COMPONENT

TABLE IV. REQUIREMENT VERSUS DEMONSTRATED CAPABILITY AT END OF CALIBRATION ANALYSIS

Phase	Parameter	Requirement	Capability
Boost	Side force (lb)	±50	±15
	Vertical force (thrust) (lb)	±200	±115
	Vector location (in.)	±0.03	±0.06
	Vector angularity (mrad)	±0.372	±1.2
Sustain	Side force (lb)	±50	±20
	Vertical force (lb)	±160	±32
	Vector location (in.)	--	±0.18
	Vector angularity (mrad)	--	±4

INTERLABORATORY STUDY OF A METHOD FOR MEASURING AMMONIUM PERCHLORATE PARTICLE SIZE

Bernard J. Alley
Army Propulsion Laboratory and Center
Research and Engineering Directorate
U. S. Army Missile Command
Redstone Arsenal, Alabama 35809

ABSTRACT

An interlaboratory study of a method for measuring the particle size distribution of finely ground ammonium perchlorate was conducted by the Analytical Chemistry Working Group of the Interagency Chemical Rocket Propulsion Group (ICRPG). The primary objective of the study was to determine the suitability of the method for use as a standard specification procedure by evaluating its precision. Single analyses of two different ammonium perchlorate samples, having weight median diameters in the range of 20-30μ were made by each of nine laboratories, using the same liquid sedimentation technique and equipment. The random error within laboratories and the systematic error among laboratories were resolved, and confidence intervals were placed on the determination of specific surface areas and weight mean diameters. The random error estimate was acceptably small; however, the systematic error estimate was so large that the method is not recommended for use as a standard specification procedure.

INTRODUCTION

Ammonium perchlorate (AP) is widely used as an oxidizer in composite and composite-modified double-base propellants. The particle size distribution of the AP has a pronounced effect on the propellant processing characteristics and ballistic properties, and therefore must be precisely measured and controlled. The recent use of finely ground AP in high burning rate propellants places greater demands on the precision of particle size analysis.

This article has been reproduced photographically from the author's manuscript.

Preceding page blank

Several analysis methods were evaluated and compared during a program [illegible] at the Army Propulsion Laboratory and Center.[1] One of these methods, based on a liquid sedimentation analysis using an apparatus manufactured by the Mine Safety Appliances (M-S-A) Company,[2,3] gave precise results and was recommended for the general analysis of fine AP.[1] The method was used with success by a number of laboratories throughout the propulsion industry.

The interlaboratory study (Round Robin) described here was subsequently conducted by the Analytical Chemistry Working Group of the Interagency Chemical Rocket Propulsion Group (ICRPG), with nine laboratories participating. The objectives were: (1) to determine the suitability of the M-S-A method for use as a standard specification procedure based on an estimation of its precision; (2) to determine the effectiveness of a simple experimental design; and (3) to evaluate and compare the performances of the participating laboratories.

EXPERIMENTAL

Each of the nine laboratories was sent three samples of nominal 20-30μ AP mixed with an inert polymer, one selected at random from each of three different batches. One of the samples was provided simply for practice prior to initiation of the Round Robin. The other two, designated materials A and B, were to be analyzed in accordance with the detailed Round Robin procedure. The instructions specified that the analysis be conducted by a skilled operator, and that the A and B samples be analyzed on different days.

Briefly, the particle analysis procedure[1] was as follows. A 15-mg sample of the material was dispersed with a surfactant and suspended in a feeding liquid composed of 60% chlorobenzene and 40% benzene by volume. The particle suspension was placed on top of chlorobenzene in a special centrifuge tube. The larger particles were allowed to fall under the influence of gravity, and the smaller particles were centrifuged. All of the particles were collected in a uniform bore capillary at the bottom of the centrifuge tube. The diameter schedule used in the Round Robin and a typical analysis are shown in Table I.

The sediment height at each particle diameter of the schedule was measured as a function of settling times precalculated from Stokes law.[4] The percentage by weight (volume) of particles greater than each successive diameter was calculated by dividing the corresponding sediment height by the total height at the end of the analysis. It will be noted from the table that the percentages are not independent.

Table I

Typical Ammonium Perchlorate Particle Size Analysis with the Mine Safety Appliances (M-S-A) Analyzer

Diameter (μ)	Sedimentation Mode	Sedimentation Rate (rpm)	Sedimentation Time (min, sec)	Sediment Height (Relative)	Weight % > Diameter
200	Gravity		0, 4.1	0	0
149			0, 7.5	0.5	1.45
105			0, 15.0	1.0	2.90
74			0, 30.2	2.2	6.38
52			1, 1	5.5	15.94
37			2, 1	10.8	31.30
25			4, 25	16.5	47.83
18	Centrifuge	300	0, 27	21.0	60.87
9		600	0, 58	26.2	75.94
5		1200	0, 52.5	29.5	85.51
3		1800	1, 19	31.5	91.30
2			2, 2	32.5	94.20
1.2			5, 16	33.5	97.10
0.6		3600	6, 2	34.2	99.13
0.4			9, 55	34.5	100

RESULTS AND DISCUSSION

The particle size distribution data are given in Tables II and III. The data for laboratories 3 and 8 were omitted from the calculations of averages and the estimates of variances (S^2) and standard deviations (S) because of the outliers[4] in Table III and the abnormal shapes of their particle size distribution curves.

The average particle size distribution data are plotted on log-probability scale in Fig. 1. The shapes of the curves are typical of those obtained for finely ground unimodal ammonium perchlorate. The difference between the particle size distributions of the AP in the two materials was purposely made small so that the random errors for their respective analyses could be assumed to be equal.

Table II

Particle Size Analysis Data for Ammonium Perchlorate in Material A
(Values in wt. % > diameter)

Diameter (μ)	Laboratories									Average[a]	S^{2} [a]	S [a]
	1	2	3	4	5	6	7	8	9			
200												
149					0			0		0	-	-
105		0		0	0.59	0	0	3.16	0	0.08	0.0497	0.2230
74	0	0.86	0	2.12	5.14	0.31	2.19	7.91	0.88	1.64	3.071	1.752
52	5.17	6.90	3.77	9.39	15.41	3.38	10.38	19.30	8.21	8.41	15.33	3.915
37	17.24	20.26	10.46	24.85	27.47	15.38	23.77	35.44	23.75	21.82	18.92	4.350
25	34.48	35.34	18.83	42.42	46.05	32.00	38.52	51.90	41.06	38.55	24.55	4.955
18	50.29	46.55	37.24	58.79	58.89	45.85	56.01	59.81	53.37	52.82	29.54	5.435
9	68.97	68.53	68.62	83.94	82.21	72.92	78.42	77.53	75.95	75.85	37.00	6.083
5	81.03	82.76	83.68	90.30	93.48	86.15	89.89	91.77	87.39	87.29	19.19	4.381
3	89.66	91.38	90.79	93.33	98.62	92.00	94.54	96.52	92.96	93.21	8.088	2.844
2	93.39	93.53	94.14	96.06	100.00	95.69	96.72	98.10	95.89	95.89	4.912	2.216
1.2	95.69	95.69	96.23	97.88		98.46	98.36	99.37	98.24	97.76	2.447	1.564
0.6	98.56	99.57	100.00	100.00		99.38	100.00	100.00	99.71	99.60	0.2705	0.5201
0.4	100.00	100.00				100.00			100.00	100.00	-	-

[a] Laboratories 3 and 8 omitted.

Table III

Particle Size Analysis Data for Ammonium Perchlorate in Material B
(Values in wt. % > diameter)

Diameter (μ)	Laboratories									Average[b]	S^2 [b]	S [b]
	1	2	3	4	5	6	7	8	9			
200	0			0.35			0	0		0.05	0.0175	0.1323
149	1.45			2.48	0	0	1.41	4.69	0	0.76	1.028	1.014
105	2.90	0		4.26	2.88	0.49	4.24	17.19	2.81	2.51	2.802	1.674
74	6.38	4.65	0	9.93	7.05	4.41	10.73	33.75	6.25	7.06	5.942	2.438
52	15.94	13.95	7.07	22.34	18.27	10.29	20.62	49.37[a]	20.00	17.32	17.81	4.220
37	31.30	31.01	14.13	38.30	31.09	24.51	35.87	63.44	36.25	32.62	21.41	4.627
25	47.83	49.61	28.27[a]	53.90	52.24	43.14	53.95	76.56	53.13	50.54	15.91	3.989
18	60.87	58.14	53.00	66.67	65.38	57.84	69.49	80.94	65.00	63.34	19.87	4.458
9	75.94	76.74	81.27	86.17	83.65	78.43	86.72	86.56	83.13	81.54	19.89	4.459
5	85.51	87.98	92.93	94.33	93.31	89.22	94.35	96.87	91.25	90.85	11.66	3.415
3	91.30	94.57	96.82	97.52	98.40	94.12	97.74	99.37	95.94	95.66	6.324	2.515
2	94.20	96.12	99.29	99.29	99.04	97.55	99.47	100.00	98.13	97.69	3.735	1.933
1.2	97.10	97.67	100.00	100.00	99.68	99.02	100.00		99.38	98.98	1.330	1.153
0.6	99.13	99.22			100.00	100.00			100.00	99.76	0.1627	0.4034
0.4	100.00	100.00								100.00	-	-

[a] Outliers by Dixon's test.
[b] Laboratories 3 and 8 omitted.

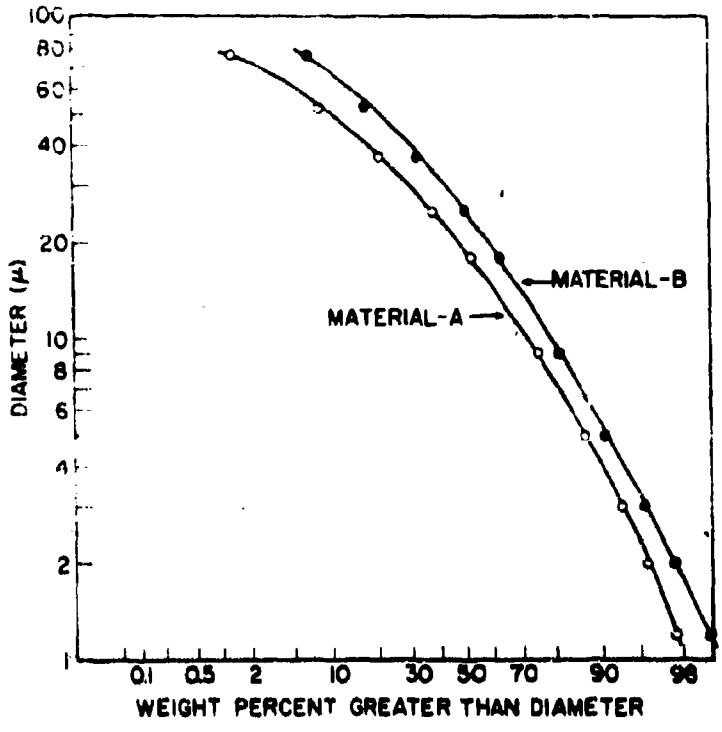


Fig. 1. Ammonium Perchlorate Particle Size Distribution Curves

Of the large number of single-valued variables that can be calculated from the particle size distribution data, the two chosen for this program were specific surface area (S_w) and weight mean diameter (d_w). The specific surface area correlates well with propellant burning rates[1] and is very sensitive to variations in the diameters of small particles; the weight mean diameter is very sensitive to variations in the diameters of large particles. The S_w and d_w values, assuming spherical particles, were calculated from the data in Tables II and III by the following formulas:

$$S_w(m^2/g) = 3.077 \sum_{i=1}^{n} \frac{W_i}{d_i} \qquad (1)$$

$$d_w(\mu) = \sum_{i=1}^{n} d_i W_i \, . \qquad (2)$$

where W_i is the weight fraction of particles in the i^{th} size interval, and d_i is an average diameter of the i^{th} interval.

The S_w and d_w values were initially examined using graphical procedures developed by Youden.[5,6] A two-material XY plot for S_w is shown in Fig. 2. Horizontal and vertical lines were drawn through the medians of the points, and a 45° line was drawn through their intersection. The perpendicular distances between the points and the 45° line are a measure of the random error within laboratories, and the spread of points along the 45° line is a measure of the systematic error among laboratories.

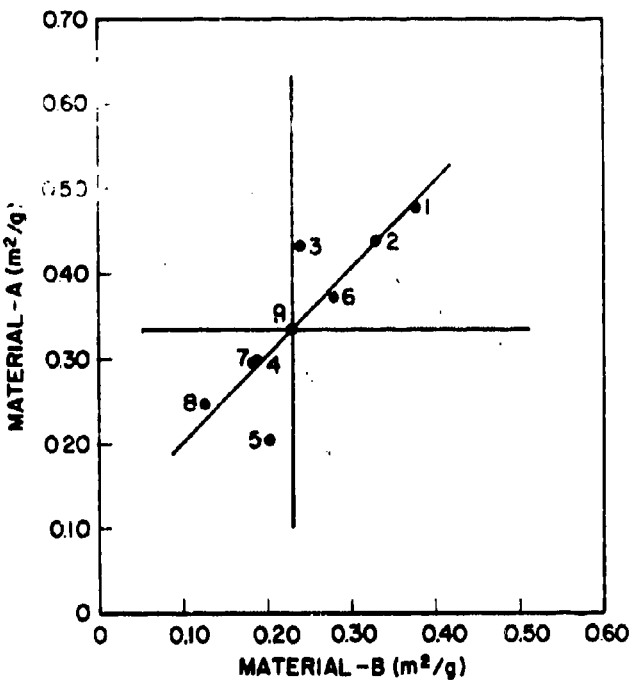


Fig. 2. Two-Material Plot of Ammonium Perchlorate Specific Surface Areas

The arrangement of points shows that the laboratories tended to get either high or low results on both materials. Moreover, the systematic error is appreciably larger than the random error. Most of the laboratories have a small random error, indicating that they did careful work. The differences between the two materials for laboratories 3 and 5 were found to be statistically significant at the 95% confidence level when compared with the average difference for all laboratories.

Figure 3 is the two-material XY plot for d_w. The laboratory 8 value differs markedly from the others, particularly for the analysis of material b, and was not considered when drawing the horizontal and vertical lines. The random error for d_w appears to be greater than that for S_w. It should be recognized, however, that the d_w values are larger and that d_w and S_w were calculated from the same particle size distribution data. The systematic error among laboratories is not noticeably larger than the random error within laboratories.

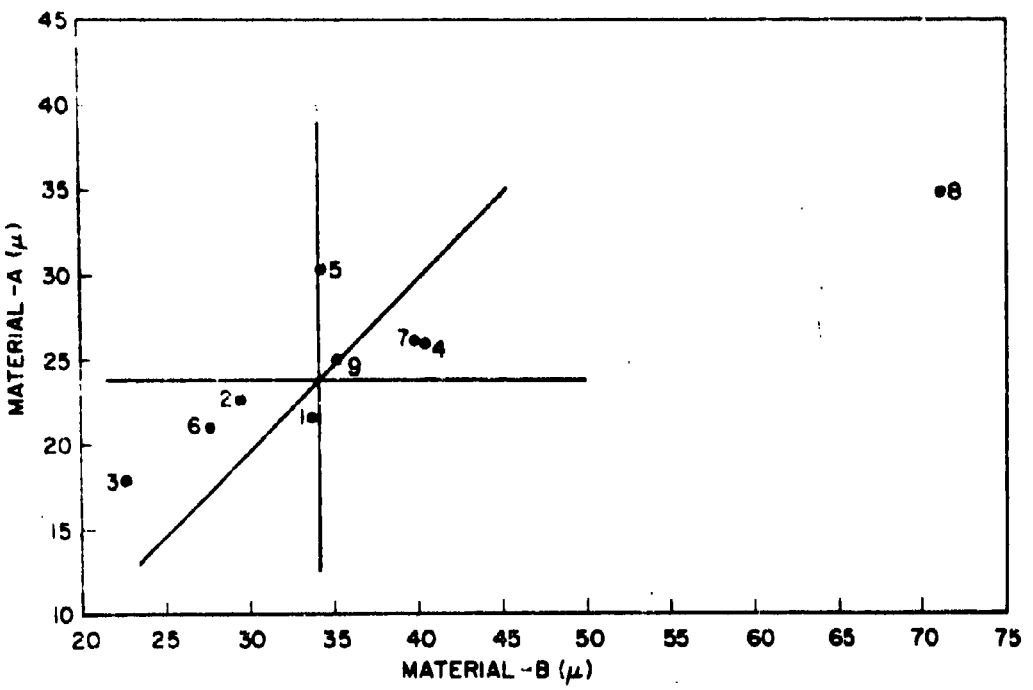


Fig. 3. Two-Material Plot of Ammonium Perchlorate Weight Mean Diameters

The extremely high d_w value for laboratory 8 clearly indicates some deviation from the recommended procedure, and this laboratory was omitted from some of the statistical analyses described later. Laboratory 3 was also omitted from some of these analyses because of its low d_w value and the outliers and abnormally shaped curves mentioned earlier.

An analysis of variance[6,7] is given in Table IV. The results, as expected, agree with the qualitative interpretations of the data in Figs. 2 and 3. The variance of S_w averages among laboratories is statistically significant when compared with the estimated random error variance (S_R^2); whereas that for d_w is not statistically significant. The laboratories x materials mean square $(MS)_{LM}$ is considered to be the best estimate of the true random error variance.

Table IV

Analysis of Variance (Laboratories 3 and 8 Omitted)

Source of Variation	DF	Specific Surface Area (m^2/g) SS x 10^4	Specific Surface Area (m^2/g) MS x 10^4	Weight Mean Diameter (μ) SS	Weight Mean Diameter (μ) MS	EMS
Among Laboratories (L)	6	806.13	134.36[a]	168.60	28.10	$\sigma_e^2 + \sigma_{LM}^2 + 2\sigma_L^2$
Between Materials (M)	1	272.27	272.27[a]	338.20	338.20[a]	$\sigma_e^2 + 7\sigma_M^2 + \sigma_{LM}^2$
Laboratories x Materials (LM)	6	46.74	7.79	45.59	7.60	$\sigma_e^2 + \sigma_{LM}^2$
Total	13	1125.14		552.39		

DF - degrees of freedom
SS - sum of squares
MS - mean square
EMS - expected mean square

$$S_R^2 = S_e^2 + S_{LM}^2 = (MS)_{LM}$$

$$S_L^2 = \frac{(MS)_L - S_R^2}{2}$$

$$S_d^2 = S_L^2 + S_R^2/n$$

[a] Statistically significant at the 95% or higher confidence level.

The estimate of the systematic error variance among laboratories (S_L^2) is also an important component of the estimate of the overall variance (S_d^2) of the method. The formulas for calculating S_L^2 and S_d^2 are given in the table. The expected mean squares (EMS) are for a random model, which was assumed in this case.

Estimated random and systematic error variances are given in Table V for various combinations of the S_w and d_w data. The S_R^2 and S_L^2 values for specific surface area determinations are not significantly affected by omitting laboratories 3 and 8, but they are significantly reduced in the case of weight mean diameter determinations. The value of S_R^2 for the determination of specific surface area is significantly reduced when laboratories 3 and 5 are omitted, as could have been expected from Fig. 2. However, a comparison with the 0.00137 value independently obtained by replicate determinations within laboratory 1 (the originating laboratory) indicates that the value of 0.000048 is not a good estimate of the true random error variance.

Table V

Estimated Random and Systematic Error Variances

	S_R^2	S_L^2
Specific Surface Area (m^2/g)		
All Laboratories	0.001171	0.006192
Omit Laboratories 3 and 8	0.000779[b]	0.006328[b]
Omit Laboratories 3 and 5	0.000048	--
Within Laboratory 1[a]	0.001374	--
Weight Mean Diameter (μ)		
All Laboratories	48.78	62.94
Omit Laboratories 3 and 8	7.60[b]	10.25[b]
Within Laboratory 1[a]	1.73	--

[a] DF = 5.
[b] Values used for calculating confidence intervals.

The S_R^2 for the replicate determination of weight mean diameter within laboratory 1 is significantly smaller at the 95% confidence level than S_R^2 for the Round Robin data with laboratories 3 and 8 omitted.

The precision of the method is defined by the confidence limits ($X \pm L_{\frac{1}{2}}$), where X is an analysis result and $L_{\frac{1}{2}}$ is one-half the length of the confidence interval, in this case at the 95% level. These one-half values are given in Table VI. They were calculated from the formulas in the table and the estimated S_R^2 and S_L^2 variance components noted in Table V. The degrees of freedom associated with S_d were estimated by Satterthwaites approximation.[7]

Table VI

Precision of M-S-A Analysis Method
($L_{\frac{1}{2}}$ at the 95% confidence level)

Number of Analyses (n)	Specific Surface Area (m^2/g)		Weight Mean Diameter (μ)	
	Within a Laboratory[a]	Laboratory at Random[b]	Within a Laboratory[a]	Laboratory at Random[b]
1	0.0683	0.201	6.75	9.55
2	0.0483	0.196	4.77	8.47
3	0.0394	0.194	3.89	8.08
4	0.0341	0.193	3.37	7.88

[a] $L_{\frac{1}{2}} = t_f S_R / \sqrt{n}$, where t_f is Student's t with f degrees of freedom. (f = 6)

[b] $L_{\frac{1}{2}} = t_f S_d$ (f = 6.69 for specific surface area, f = 9.02 for weight mean diameter).

The estimated precision of analyzing AP samples within a single laboratory (random error) is given by the confidence intervals in the second and fourth columns of Table VI. These intervals apply for the analysis of AP samples having particle size distributions within the range shown in Fig. 1. Note that the precision improves with an increase in the number of replicate analyses.

The estimated precision of analyses, considering the random plus the systematic error, by any laboratory selected at random is given by the confidence intervals in the third and fifth columns. Assuming the participating laboratories are representative of the entire population of laboratories, these precision estimates determine the suitability of the method for use as a standard specification procedure. The error is larger than for analyses within a single laboratory because of the contribution of the systematic error variance (S_L^2). Nor is the precision improved much in this case by replicate analyses, because the replicates (n) reduce only the smaller S_R^2 component of S_d^2.

Perhaps in actual practice a higher degree of confidence than 95% would be desired. For higher degrees of confidence the value of t_f would be larger, and the confidence intervals would increase accordingly. The accuracy (bias) of the method could not be estimated in this Round Robin, because a standard AP sample of accurately known particle size is not available.

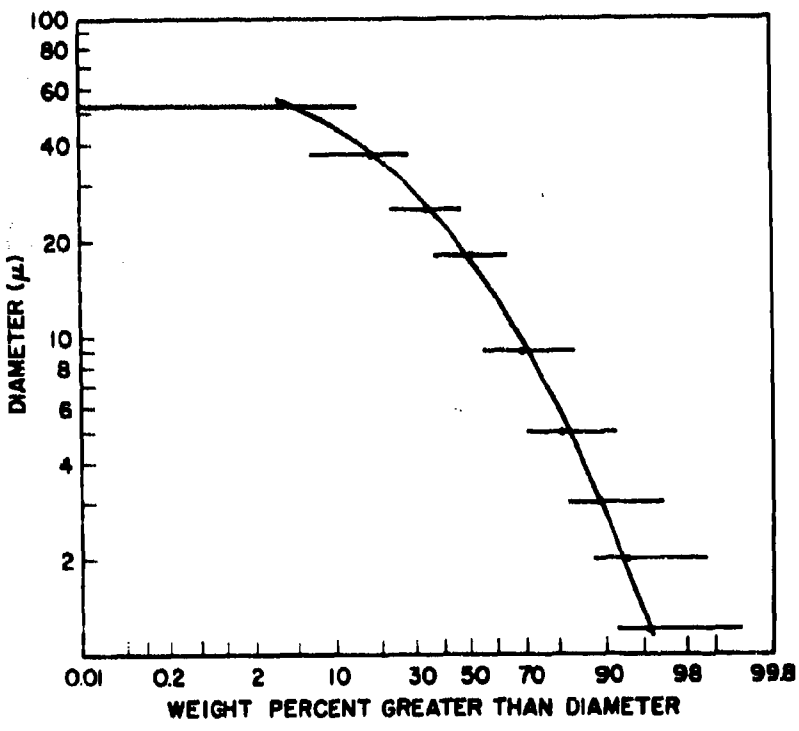


Fig. 4. 95% Confidence Intervals on the Percentage Points of the Particle Size Analysis of Ammonium Perchlorate in Material A by Laboratory 1

Confidence intervals can also be placed on the percentage points of a particle size distribution curve. This is illustrated in Fig. 4 for the analysis of a single sample of material A by laboratory 1. The confidence intervals were calculated using the variance estimates in Table II. Such a precision estimate is of value for determining whether the variations in particle size distribution are due to the analysis procedure or the grinding process.

CONCLUSIONS AND RECOMMENDATIONS

The precision of analyses within a single laboratory is considered adequate, and the fact that a number of laboratories are successfully using the procedure supports this conclusion. However, the method is not recommended as a standard specification procedure for the particle size analysis of nominal 20-30μ ammonium perchlorate, because of the large systematic error among laboratories in the determination of specific surface area. The great difference between these two errors could be due to some deficiency in the analytical procedure that permits laboratories to introduce their own variations. One likely area of inconsistency is in the dispersion of the AP particles, but there are no known alternative techniques that would not also affect the accuracy of analyses.

The simple experimental design, without replication, encouraged laboratories to participate, thus enabling a more reliable estimate of systematic error among laboratories. Past experience has shown that this systematic error is almost always significantly greater than the random error within laboratories.

The comments and suggestions of the clinical session panelists are particularly solicited with respect to the following elements of the Round Robin statistical analysis:

(1) Estimation of the degrees of freedom associated with S_d. and S_L.

(2) Determination of confidence intervals or regions for particle size distribution curves.

(3) Criteria for the rejection of extreme laboratories and data.

(4) Experimental design for Round Robins and possible alternatives.

REFERENCES

1. Alley, B. J., Dykes, H. W. H., and Howard, W. W., "Particle Size Analysis of Fine Ammonium Perchlorate and Correlation with Propellant Burning Rates (U)," Report No. RK-TR-68-4, U. S. Army Missile Command, Redstone Arsenal, Ala., April 1968 (Confidential report).

2. M-S-A Particle Size Analyzer, Operating Procedures and Applications, Mine Safety Appliances Company, Pittsburgh, Pennsylvania.

3. Whitby, K. T., "A Rapid General Purpose Centrifuge Sedimentation Method for Measurement of Size Distribution of Small Particles," Heating, Piping, and Air Conditioning, 61, June 1955.

4. Dixon, W. J., and Massey, F. J., Jr., Introduction to Statistical Analysis, McGraw-Hill Book Co., New York (1957).

5. Youden, W. J., "Graphical Diagnosis of Interlaboratory Test Results," Industrial Quality Control, Vol. 15, No. 11 (1958).

6. Youden, W. J., "Statistical Techniques for Collaborative Tests," The Association of Official Analytical Chemists, Washington, D. C., 1967.

7. Hicks, C. R., Fundamental Concepts in the Design of Experiments, Holt, Rinehart, and Winston, New York (1964).

NEW ANALYSES AND METHODS LEADING TO IMPROVED TARGET ACQUISITION REQUIREMENTS INVOLVING SYSTEMS, GEODETIC AND RE-ENTRY ERRORS, AND INCREASED WEAPONS EFFECTIVENESS FOR CONVENTIONAL WEAPONS*

Hans Baussus-von Luetzow
U. S. Army Engineer Topographic Laboratories
Fort Belvoir, Virginia

SUMMARY. After a cursory critique of currently used methodology for the study of target-accuracy requirements for artillery weapons, this research note is concerned with the development of analytical methods and two different though interrelatable and essentially additive optimization concepts. If implemented within the context of TACFIRE, these methods are conservatively estimated to provide on the average a 30% greater weapons effectiveness. Although the intra and extra weapons systems employment parameters are interdependent, variable, and changing, an integrated operational optimization is achieved. The methods outlined are also useful in weapons R&D and related systems analyses. Furthermore, the rather cogent requirement and related recommendations or conclusions arrived at may be of considerable significance for certain R&D and combat development activities.

FOREWORD. It was originally contemplated to finalize this study in 1967. The author who was also investigating more powerful methods in connection with burst and target height variabilities and the use of conventional cratering and nuclear weapons became, however, increasingly convinced that rudimentary or short-cut methods had to be considered unsatisfactory. A more rigorous and mature approach required time and concentration in view of the slow progress made in the past and also because of a satellite systems study performed during 1967. As to the word "improved" in the title, this should rather be interpreted as "less restrictive," Implementation of the methods and concepts developed would undoubtedly lead to a significant increase of Army weapons effectiveness. In addition, the new methods are expected to have some ramifications pertaining to a variety of R&D and combat development activities. The technical responsibility for this study is exclusively the author's who appreciates USAETL's continued interest in this kind of effort.

*This article appeared as Research Note No. 35, U. S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia.
The remainder of this article has been reproduced photographically from the author's manuscript.

NEW ANALYSES AND METHODS LEADING TO IMPROVED TARGET ACQUISITION REQUIREMENTS INVOLVING SYSTEMS, GEODETIC AND RE-ENTRY ERRORS, AND INCREASED WEAPONS EFFECTIVENESS FOR CONVENTIONAL WEAPONS

1. Introduction.

1.1 The essential ideas underlying this report were developed in August 1966 after an evaluation of the following material: "Target Acquisition Accuracy Requirements, 1965-1975 (U)" (1)[1]; "A Model for Determining Target Location Accuracy Requirements" (2); "Trip Report to CDC Artillery Agency" (3); and "A Technical Analysis to Support Map Accuracy Requirements" (4).

1.2 According to Ref. (3), additional contractual work, to start in July 1967 and expected to last one year, was considered necessary by USACDC in order to improve the methodology report (1). It led to the report "A Study of Target Location Accuracy Requirements for Artillery Weapons – Army 1975 (U)" (5). This study, conducted at the Combined Arms Research Office, Fort Leavenworth, Kansas, and coordinated with the USACDC Artillery Agency, Fort Sill, Oklahoma, applied the methodology of Ref. (1) to all artillery weapons of the 1966-1975 time frame.

1.3 The methodology in both Ref. (1) and Ref. (2) is essentially restricted to the 2-dimensional problem of fragmentation projectiles with impact fuzes and thus less applicable with respect to height bursts. It consists of computing a measure of effectiveness f (see Ref. (1), B-1, eq. (2)) and a fractional coverage C (C^1 in Ref. (2)) so that the fraction of casualties $F = f \cdot C$. Although it has not been spelled out explicitly, f is the probability of hitting the target which is computed by dividing the common area between target and effects pattern, a, by the target area A_T. The determination of f involves the use of a quantity A_L called the lethal area. A_L and C are calculated under the assumption of a uniform target distribution. As to multiple volleys, the assumption is made that the percentage reduction g in F will be directly proportional to the respective g_n in F_n (n volleys). Through the use of this methodology, Spears strives to arrive at the conclusion that "Changes in single-volley coverage of a target by a weapons effects pattern (a quantity relatively easy to determine) can be used as a basis for determining critical reductions in *effectiveness* of multi-volley fire (a quantity difficult to determine accurately)." Through the introduction of $\frac{a}{A_p} \sum_i A_{Li}$ as a measure of the

1. Numbers in parentheses appearing in the text refer to "LITERATURE CITED," p. 13, while numbers in parentheses on the right margin refer to equations.

average fraction of the firepower which hits the target, i.e., the total casualty potential reduced by the factor "common area between target and effects pattern divided by the effects pattern area" and use of the Poisson distribution with $= \frac{a}{A_p} \sum_i A_{Li}/A_p$ in the exponential. Spears arrives at a new non-factorized, or mixed, F. This derivation is, however, not permissible and conflicts accordingly with the former result $F = f \cdot C$. At best, it is an approximation within an approximate framework. Reference (4) states that the assumption of proportionality between casualties and fractional coverage is the weakest link in the methodology employed in Ref. (1) and (2) and criticizes various other assumptions made. Under the criterion that the target acquisition does not degrade the accuracy of the weapons systems itself by more than 10% and under the assumption that the map accuracy or error is the principal contributor to the weapon site survey error and the target location error, it is concluded in Ref. (4) that present map accuracies can be relaxed or that the Class A National Map Standards have about twice the required precision. This result has been obtained by simple calculations based on the assumption that the total variance is the sum of the individual variances including the geodetic one (target location error). This assumption is wrong and is one of the basic weaknesses of all analyses so far, apart from a rather primitive methodology.

1.4 In view of the shortcomings enumerated in paragraph 1.2 and in order to provide a sound basis for decision making, this report was undertaken. Objectives of the report are as follows:

1.4.1 A rigorous mathematical-statistical analysis involving a direct, physical approach to the problem. In particular, this analysis shall be independent of assumptions regarding target distributions and simplifications involving, e.g., proportionality between casualties and fractional coverage, and casualty potential.

1.4.2 Inclusion of multiple shots and multiple volleys without loss of rigor or generality.

1.4.3 Optimization for multiple volleys as a new and most significant discovery.

1.4.4 Consideration of inhomogeneous target distributions and its change after the first volley.

1.4.5 Incorporation of meteorological-error variances.

1.4.6 Utilization of non-circular distribution parameters.

1.4.7 Application of non-isotropic fragmentation patterns.

1.4.8 A total systems optimization or marginal utility analysis involving the whole range and employment spectrum of a weapons systems, i.e., a grand optimum.

1.5 A contemplated Part II of this study will include a supplemental analysis for height bursts (time and ambient fuzes) including vertical target location errors.[2] Parts III and IV will deal with cratering conventional weapons and nuclear weapons respectively, and a partially different methodology will be required in these areas.

1.6 The optimal aiming pattern analysis together with the optimal overall weapons-systems employment concept developed in this report allow—on the average—a considerable relaxation pertaining to stringent target-acquisition requirements in general and map-accuracy requirements in particular. They tend to shorten the firing engagement time and are also advantageous in case of ammunition shortage. An exception would be hardened-point targets. According to the experience gained (as mentioned in footnote 4), target-location errors can be very large, and identification and location problems will probably exist for longer distances if direct distance and azimuth measurements are performed, though to a lesser extent. Meteorological errors are also not supposed to become negligible under many combat conditions. In view of the above, numerically fixed and extreme accuracy requirements synonymous with sophisticated and very expensive equipment which very often does not live up to expectations under realistic conditions are unnecessary. The R&D process in the areas of more accurate mapping and target location being essentially independent of that pertaining to new weapons systems should be pursued at a normal technological pace and should not overemphasize accuracy but rather concentrate on versatility, reliability, and survivability. This is also consistent with a recent directive of the Army Chief of Staff.

As exhibited by this study, the intra and extra weapons systems employment parameters are interdependent, variable, and changing but nevertheless allow a continuous integrated operational optimization. In so far, the study is also of significance for the Geographic Intelligence and Topographic Support Systems Study (GIANT), the development of the Position and Azimuth Determining System (PADS), and the development of the Long Range Position Determining System (LRPDS). Finally, the methods outlined can serve as a research and weapons-systems analysis tool for both the Combat Developments Command and the Materiel Command.

---

2. In this regard, it may be worthwhile to mention that, according to Ref. (3), O. S. Spears, Scientific Advisor to CDC, Artillery Agency, has stated the following: "It is not that we don't consider the vertical component important, we simply realize that it is a difficult problem to solve. Once we get a complete handle on the horizontal accuracies, we will be able to start tackling that problem more intelligently."

2. **Individual Hit Probabilities.** The fragment-damage pattern of a particular artillery shell is not isotropic as can be inferred from Fig. 1. It depends also on range, height of burst, and impact angle. Tabulations contain, in general, isotropic data including distance from burst, total number of effective fragments, and average number of effective fragments per area unit.

An example from Ref. (6) is given below:[3]

Fragment Damage of Shell, HE, 155-mm, M107;
Initial Fragment Velocity 3,500 f/s
Source: Army TM 9-1907, Table XXXV

Dist from Burst in Feet	Total Number of Effective Fragments	Average Number of Effective Fragments Per Sq Ft	For the Lightest Effective Fragment	
			Weight ozs	Vel f/s
r	N	B	m	v
20	1880	.374	.0108	2340
30	1740	.154	.0148	2000
40	1640	.0816	.0195	1740
60	1450	.0321	.0310	1380
80	1300	.0162	.0440	1160
100	1220	.00971	.0562	1030
150	1040	.00368	.0832	845
200	940	.00187	.109	738
300	770	.00068	.166	598
400	640	.00032	.235	503
700	420	.00007	.515	340

From individual, i.e., unaveraged, fragment patterns, it is possible to determine through the use of sampling techniques individual hit probabilities. Thus, $p_1 (F, \rho, \phi)$ would be the average probability that a person or item with cross section F which is located at a distance ρ and azimuth ϕ from the burst suffers exactly one hit. In this respect, the azimuth is to be counted counterclockwise from the line of fire. By $p_1 (F, \rho, \phi) = p_1 + p_2 + \ldots$. we designate the probability of at least one hit. With reference to human beings, it would be possible to drop the letter F. For identification purposes, we denote a semi-fixed pattern of human beings by superscripts and have

3. An excellent introduction into kinds and characteristics of explosives (munitions) is given in Ref. (7).

Source: Army TM 9-1907

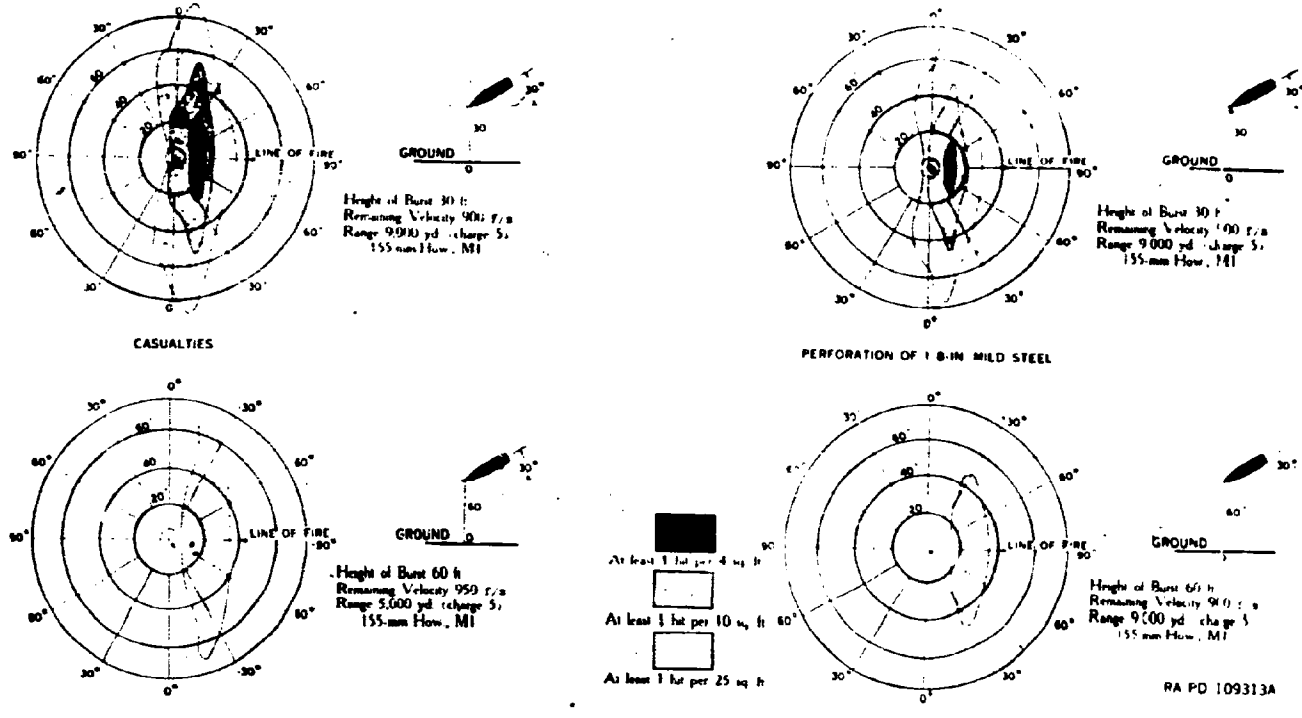


Fig. 1. Damage pattern – 155-mm HE shell, M107.

thus for example, $p_1^{14}(\rho, \phi) = p_1(\rho_{14}, \phi_{14})$. It is to be kept in mind that these probabilities refer to constant range, height of burst, and impact angle.

3. **Distributions and Distribution Parameters.** Weapon distribution parameters for a specific range are the line of fire and lateral standard deviations S_r and S_ϱ. As already mentioned, height uncertainties are considered negligible in this investigation (Part I). In addition, we have target-location errors depending, e.g., on map accuracy, target identification, and location,[4] and meteorological errors. The corresponding distributions may for simplicity be described by the two parameters σ_T and σ_M. Since we restrict ourselves to normal distributions, we may establish the relation

$$\text{var } r_T + \text{var } r_M = \text{var } \varrho_T + \text{var } \varrho_M = \sigma_T^2 + \sigma_M^2 = \sigma^2 \qquad (1)$$

It is important to remember that σ_T^2 and σ_M^2 need not be considered constant for a certain range. Hence, σ^2 may allow a few classes of variability depending on circumstances.

4. **Formulation of Multiple Volley Optimization Problem for Stationary Personnel Distribution.** In Fig. 2, the general target coordinate system for aiming purposes is denoted by x, y. At the origin, the combined target location distribution

$$f(\bar{\xi}, \bar{\eta})\, d\bar{\xi}\, d\bar{\eta} = \frac{1}{2\pi\sigma^2}\, e^{-\frac{1}{2}\frac{\bar{\xi}^2 + \bar{\eta}^2}{\sigma^2}}\, d\bar{\xi}\, d\bar{\eta} \qquad (2)$$

has a maximum. The aiming point for the first volley is represented by 0_2, with coordinates a_1, b_1; and the respective gun-aiming points separated by the distance c are

4. The accuracy of a class A map of 1/50.000 scale within a single sheet can be expressed approximately by a standard deviation of 25m. Though this is not a negligible parameter and accuracies decrease with reference to lower quality maps, additional errors enter in case of target identification (which includes determination of a reference point for the whole target configuration) and target location on the map. The latter type of error can be very sizable, and standard deviations of the order 500m have been found according to Ref. (8), (9), and (10). For simplicity, we lump map, target identification and target location variances together into var r_T. Smaller σ_T's are, of course, to be expected in case of a direct link including distance and azimuth measurements between observer and a suitable target reference point. Directional σ_T's might also be generated by moving targets and target configurations.

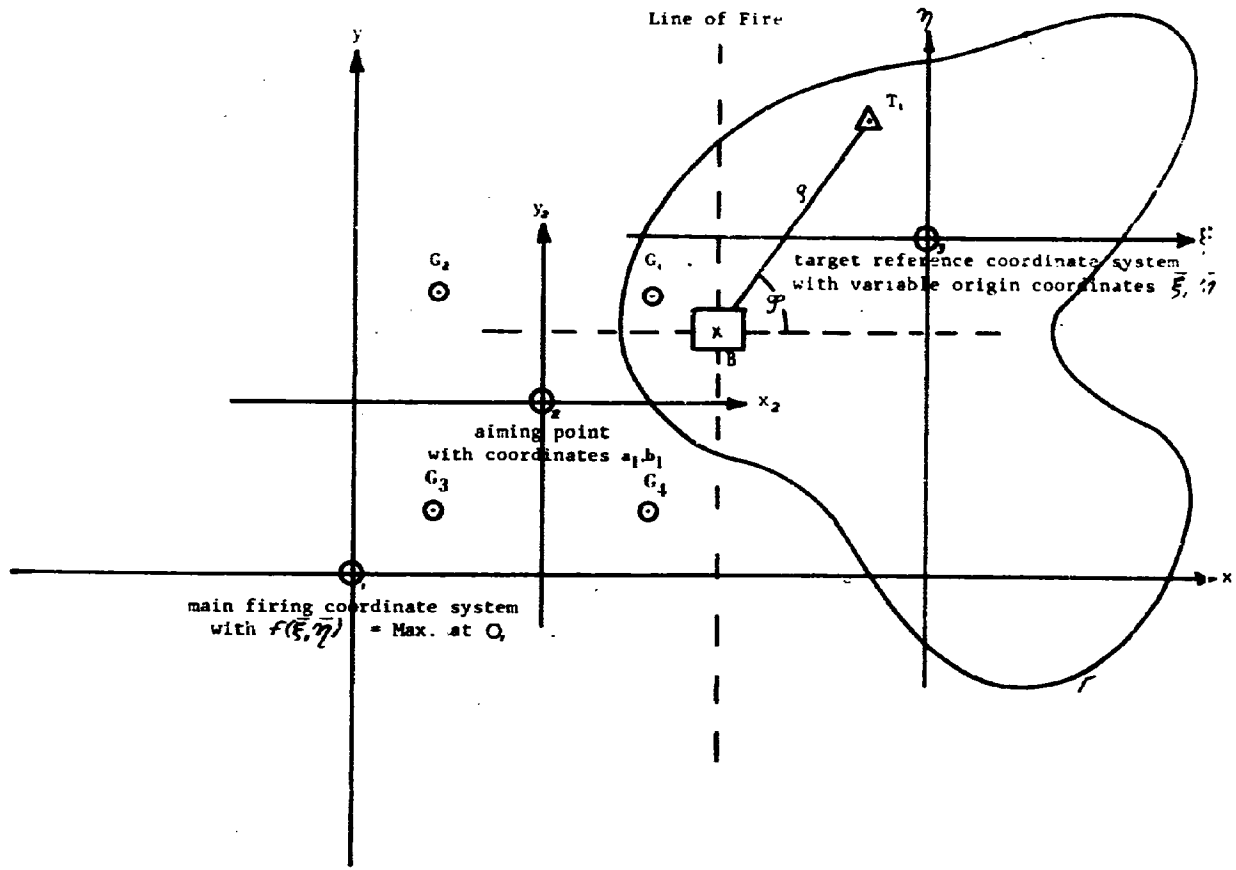


Fig. 2. The general target coordinate system.

G_1, G_2, G_3, G_4. The burst point for which a total impact probability is to be computed and which lies in a finite area element (for numerical purposes) is B. Only one individual target, T_1, is indicated within the target area with boundary Γ. The distance from B to T_1 is ρ and the azimuth is ϕ commensurable with the denotations of para. 2.

For the first volley, we have four burst distributions designated by

$$\lambda_1(x,y;a_1+\frac{c}{2}, b_1+\frac{c}{2}; S_r, S_\varrho)dxdy, \lambda_2(x,y;a_1-\frac{c}{2}, b_1+\frac{c}{2}; S_r, S_\varrho), \text{ etc.} \tag{3}$$

For sufficiently small area elements $\Delta x \Delta y$, we arrive then in integral form at an intermediate average probability of hitting T_1 at least once,

$$P_{i1}^1 = \int_{-\infty}^{+\infty}\int \sum_1^4 \lambda_\nu \cdot p_1^1(\rho,\phi)\,dxdy = \int_{-\infty}^{+\infty}\int \sum_1^4 \lambda_\nu\, p_1(x, y, \bar{\xi}, \bar{\eta}, \xi_1, \eta_1)\,dxdy \tag{4}$$

and, since 0_3 obeys a distribution law, at

$$P_1^1 = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int \sum_1^4 \lambda_\nu\, p_1(x, y, \bar{\xi}, \bar{\eta}, \xi_1, \eta_1)\, f(\bar{\xi}, \bar{\eta})\,dxdyd\bar{\xi}d\bar{\eta} \tag{5}$$

For k volleys and μ individual targets, we obtain the total expected casualty result

$$n_1 \sum_1^k = \sum_{\tau=1}^{\tau=k} \sum_{\mu=1}^{\mu=N} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int \sum_{\nu=1}^{\nu=4} \lambda_\nu(x,y;a_\tau,b_\tau)\, p_1(x,y;\bar{\xi},\bar{\eta};\xi_\mu,\eta_\mu)\, f(\bar{\xi},\bar{\eta})\,dxdyd\bar{\xi}d\bar{\eta} \tag{6}$$

The optimization conditions can be formulated as

$$\frac{\partial}{\partial a_\tau} n_1 \sum_1^k = 0 \; ; \qquad \frac{\partial}{\partial b_\tau} n_1 \sum_1^k = 0 \tag{7}$$

For a circular and homogeneous (uniform) target distribution conditions (7) reduce to fewer equations, i.e., the respective aiming pattern, consisting of a set of k origins 0_2 would be invariant under a rotation about 0_1.

After the first volley which has in many cases a surprise effect, a degradation with respect to P_1 can be expected which can be expressed as a transition from ground to air

burst[5] and by an empirical reduction factor $R < 1.00$. Taking this into consideration, a more general result corresponding to eq. (6) reads as

$$n_{1\sum_1^k} = \sum_{\mu^{(1)}=1}^{\mu^{(1)}=N} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sum_{\nu=1}^{\nu=4} \lambda_\nu \cdot p_1(x,y;\bar{\xi},\bar{\eta};\xi_{\mu(1)},\eta_{\mu(1)}) f(\bar{\xi},\bar{\eta}) dx dy d\bar{\xi} d\bar{\eta}$$

$$+ \sum_{\tau=2}^{\tau=k} \sum_{\mu^{(2)}=1}^{\mu^{(2)}=N} R \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \sum_{\nu=1}^{\nu=4} \lambda_\nu \cdot \pi_{\mu(2)} p_1^c(x,y;\bar{\xi},\bar{\eta};\xi_{\mu(2)},\eta_{\mu(2)}) f(\bar{\xi},\bar{\eta}) dx dy d\bar{\xi} d\bar{\eta} \quad (8)$$

In eq. (8), the first term refers to the first volley. The second term containing the reduction factor R reflects a changed hit probability function and includes the factor $\pi_{\mu(2)}$.

A particular $\pi_{\mu(2)}$, say $\pi_{3(2)}$, requires the computation of the individual P_1^3 from eq. (5). It is then

$$\pi_{3(2)} = 1 - \kappa P_1^3 \quad (9)$$

where κ denotes the (average) probability that an individual, hit at least once, remains at the initial position. The index (2) in $\pi_{3(2)}$ indicates the transition from the initial target configuration to a second, more protective one.

It should be mentioned that, in connection with an evaluation of eq. (8), an average $n_{1\sum_1^k}$ for typical target distributions under consideration of protective obstacles can be determined. It is also possible to classify targets by size and concentration indices (cf. para. 7). Furthermore, it is possible to split P_1 up into probabilities for exactly 1 hit, 2 hits, etc.

5. **Probability Distributions.** The probability distribution associated with $n_{1\sum_1^k}$ of eq. (8) can easily (though approximately) be found by setting $\bar{p} = \frac{n_{1\sum_1^k}}{N}$ where N

5. According to Ref. (6), p. 181, ground bursts generally are more effective against material and personnel in case of no shielding by revetments, but personnel in foxholes or trenches should be attacked by air-burst fire.

is the total number of individual targets. According to Kendall (11), we have then the moments about the mean $N\bar{p}$

$$\mu_2 = N\bar{p}\bar{q} \qquad \bar{q} = 1 - \bar{p}$$

$$\mu_3 = N\bar{p}q(\bar{q} - \bar{p})$$

$$\mu_4 = 3N^2\bar{p}^2\bar{q}^2 + \bar{p}\,\bar{q}\,N(1 - 6\,\bar{p}\,\bar{q})$$

from which a Pearson Type I curve can be calculated which is evaluable in terms of the incomplete B-function. Alternatively, the binomial distribution and associated tables might be used.

6. **Symmetric Aiming Patterns.** As an illustration, some symmetric aiming patterns applicable for circular and completely homogeneous conditions are shown in Fig. 3. Equivalent solutions would result through an arbitrary rotation.

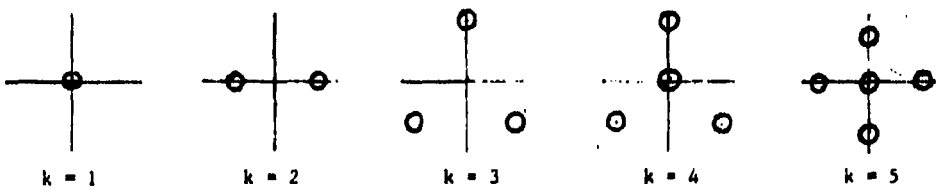


Fig. 3. Symmetric aiming patterns.

7. **Computation and Utilization Considerations.** A particular kind of quasi-circular target could profitably though not exclusively be characterized by r (range to origin 0_1) S (target size, 3 indices), D (quadrant density, 3 indices), σ (combined target location error, 3 or 4 indices), k (number of volleys). From these data, k azimuth and angular height corrections for the optimal aiming points would be immediately available. Of course, the optimization computations should be conducted by a large-scale digital computer, i.e., not in the field. The corrections = functions of variable input data would be available as stored digital information and incorporated in TACFIRE procedures. Probability statements depending on k could be added. Too many indices are to be avoided. As to D, there are 3^4 variations (with three elements of the fourth class and repetitions). Some of these variations can be omitted because of practical reasons. Nine quadrants leading to 3^9 complexions would be prohibitive. From a

scientific standpoint, it so appears that an effective employment of a single artillery weapons system (battery) represents a rather formidable problem.

8. Views on Optimal Weapons Systems Employment.

8.1 With reference to optimization considerations, we assume the existence of the following scheme:

$$\left.\begin{array}{lll} r_1 & h_{11}\, U_{11}(k_{11}) & \ldots\ldots\ldots\ldots\ h_{1n}\, U_{1n}(k_{1n}) \\ r_2 & h_{21}\, U_{21}(k_{21}) & \ldots\ldots\ldots\ldots\ h_{2n}\, U_{2n}(k_{2n}) \\ \vdots & & \\ & - & \\ r_n & h_{m1}\, U_{m1}(k_{m1}) & \ldots\ldots\ldots\ldots\ h_{mn}\, U_{mn}(k_{mn}) \end{array}\right\} \qquad (10)$$

In this discrete scheme, valid for a particular weapons system (e.g., artillery battery), the symbols r, h, U, and k denote range, relative frequency of employment, mean weapons effect for a particular type of target, and number of volleys respectively. Strictly speaking, there has to be a greater number of discrete schemes with associated scheme frequencies in order to account for variations in target size and target location. This involves an additional frequency matrix with elements $j_{a,b}$. The total mean weapons systems effect can then be formulated as

$$\bar{U} = \sum_{\alpha,\beta} h_{\alpha,\beta} \sum_{a,b} j_{a,b}\, U_{\alpha,\beta;a,b}(k_{\alpha,\beta;a,b}) \qquad (11)$$

We shall distinguish between $\bar{U}$ and $\hat{U}$ with $\hat{U}$ considered optimized by the analysis outlined in para. 4. In other words, $\bar{U}$ does not imply the utilization of optimal aiming patterns.

If we apply the rather usual criterion of 30% damage or casualties with a 90% assurance, we arrive at

$$\bar{U}_{30\%/90\%} \cdot \quad \bar{K} = \sum_{\alpha,\beta} \sum_{a,b} k_{\alpha,\beta;a,b} \qquad (12)$$

and

$$\bar{U}_{30\%/90\%} \cdot \quad \hat{K} = \sum_{\alpha,\beta} \sum_{a,b} \hat{k}_{\alpha,\beta;a,b} \qquad (13)$$

with $\bar{U} \approx \hat{U}$ and $\hat{K} < K$.

A reasonable measure for the effectiveness increase expressed in percent is evidently

$$\eta = 100 \frac{K - \hat{K}}{\hat{K}} \tag{14}$$

8.2 A different approach would consist of stipulating a constraint, say

$$\hat{K}^{(2)} = \sum_{\alpha,\beta} \sum_{a,b} k^{(2)}{}_{\alpha,\beta:a,b} \leq A \tag{15}$$

and to compute the $k^{(2)}$'s in such a way that

$$\hat{U}^{(2)} = \text{Max.} \tag{16}$$

For the purpose of comparison, we may assume

$$\hat{K}^{(2)} = \hat{K} \tag{17}$$

The optimization expressed by eq. (15) and (16) implies *a forteriori*

$$\hat{U}^{(2)} > \hat{U} \tag{18}$$

and a relatively greater expansion of volleys with respect to closer range targets and those involving smaller σ's. On the other hand, for some targets with less favorable characteristics, the 30%/90% criterion might not be fulfilled.[6] What can be said with certainty is that the utilization of optimal aiming patterns makes the ground optimization described in para. 8.2 quite attractive. It is conservatively estimated that optimal aiming-pattern utilization incorporated in TACFIRE would result in a 15% increase in systems' effectiveness. The systems' overall optimization would yield an additional 15% increase and thus lead to a combined improvement of 30%.

6. This is, however, not a serious limitation since it can be partially or completely overcome by a greater A in eq. (15). This would particularly apply to defensive positions with a large ammunition supply.

LITERATURE CITED

1. U. S. Army Combat Developments Command, Artillery Agency, "Target Acquisition Accuracy Requirements, 1965-1975 (U)," Fort Sill, Oklahoma, June 1965.

2. Spears, Otis S., "A Model for Determining Target Location Accuracy Requirements," USACDC, Artillery Agency, Fort Sill, Oklahoma (paper presented at the Operations Research Symposium, Fort Monmouth, N. J., 29-31 March 1966).

3. USACDC, Artillery Agency, "Trip Report to CDC Artillery Agency," Fort Sill, Oklahoma, 12-19 April 1966.

4. Autometric Operations, Space and Information Systems Division, Raytheon Co., "A Technical Analysis to Support Map Accuracy Requirements," Alexandria, Virginia, 23 July 1966.

5. Booz-Allen Applied Research, Inc., "A Study of Target Location Accuracy Requirements for Artillery Weapons—Army 1975 (U)," Fort Leavenworth, Kansas, June 1967.

6. Department of the Army Technical Manual TM 9-1907, "Ballistic Data Performance of Ammunition," U. S. Government Printing Office, Washington, D. C., July 1948.

7. "Weapons Systems Fundamentals NAVWEPS OP 3000 (Volume 2)," U. S. Government Printing Office, Washington, D. C.

8. "Evaluation of Photo Based Map Substitute Products of Ft. Sill, Field Test Evaluation," Army Map Test, August 1966.

9. "Ground Observer Probabilities of Acquisition Adjustment," Volume 1 of 2 (ACN 7395), USACDCEC, Fort Ord, California, September 1968.

10. Shaffer, F. B., "Excerpts from the Accuracy of Maps in Tactical Operations," Stanford Research Institute, Draft 1967.

11. Kendall, M.G., "The Advanced Theory of Statistics," Volume I, Charles Griffin and Co., London, 1948.

An Air Defense Comparative Model

Robert E. Shannon
James P. Ignizio
James L. Stimach

University of Alabama Research Institute
Huntsville, Alabama

ABSTRACT

This paper proposes a new and unique approach for conducting comparative experiments or evaluations between existing or proposed air defense weapon systems. It is based upon the game theory "minimax" philosophy and provides several distinct advantages over the use of computer simulation methods. Submodels for objectively determining the optimal deployment of the defense and the optimal attack routes to be used by the attacking aircraft are discussed.

INTRODUCTION

Development and deployment of air defense systems having a large degree of effectiveness against high altitude aircraft, has resulted in increased interest in the operation of tactical aircraft at low altitudes [1]. As a result many weapons systems analysts have become deeply involved in analytical and experimental studies evaluating the effectiveness of existing or proposed defense systems for defeating the low altitude threat. Historically, war games have been extensively used to "model" military situations for

This article has been reproduced photographically from the authors' manuscript.

such experimental and evaluative purposes. The different types or classes of war games that have been used are: (1) field exercises, (2) board games, and (3) computer simulations.

When considering air to ground conflict situations, experimentation using the computer simulation technique has proven to be the most feasible and efficient. Consequently, computer simulation models have evolved from very simple and basic models into those which are now very large, complex and time consuming. This increase in size and complexity has arisen due to the desire to approach, as near as practically possible, an exact model of the real life situation. Unfortunately, as realism has increased, so too has the computer time required to run the experiments.

This paper proposes a new approach for conducting comparative experiments or evaluations between existing or proposed air defense weapon systems. It is based upon the game theory "minimax" philosophy and provides several distinct advantages over existing computer simulation models. These include:

1. Less computation time required.
2. Fewer necessary assumptions and simplifications, hence greater realism.
3. Additional useful information is generated such as optimal defense system deployment, optimal attack routes, etc.
4. Real world scenarios, (actual situations) may be used.
5. Only one computer run per defense system is required.

MODEL PHILOSOPHY

The scenario used for the Air Defense Comparative Model should be a real world situation i.e. a specific piece of terrain which is to be defended. The concept of the low altitude attack is to utilize the masking effects of the terrain (hills, valleys, etc.) and of the earth's curvature, to prevent the defense from being able to detect and engage the attackers until the targets are reached [1]. Hence, any experimental evaluation of the defense system must take this into account.

The defensive problem in our scenario can be stated as follows:

1. Given a specific sector of terrain (with hills, valleys, etc.) which is to be defended by n or less defensive units.
2. Given the characteristics of the defense system (i.e. range, maximum and minimum elevation angles, azimuth scan angle, etc.).
3. Given the feasible locations and pointing angles for the placement of defensive units. (i.e. cannot be located in middle of lakes, bottom of ravines, etc.).
4. Find those n locations and pointing angles which (a) minimize the range from any attacker to a systems radar and (b) maximizes the visibility of the combined radar systems. This must take into account the masking effects of terrain features and earth's curvature.

Likewise, we can state the problem faced by the offense or attackers in our scenario. This can be done as follows:

1. Given a set of targets to be destroyed which are contained within a specific piece of terrain which is defended by a set of ground based air defense missile systems which are optimally deployed.
2. Given that the attacker or penetrator has complete knowledge from intelligence operations of these allocations and of the defensive capabilities.
3. Determine the best location to enter the defended sector, and then the least risk route to follow in order to reach a designated target. The least risk route is that which minimizes the visibility time and maximizes the survival probability.

In the proposed Air Defense Comparative Model, these problems are solved objectively and optimally by sub-models. The objective, optimal solution to both problems, is a unique feature of the proposed model. The tactics are not determined by educated guess as in other war game models. It should be pointed out however, that the two sub-models (optimal allocation of defense units and optimal attack route analyzer) can be used to set the scenario and tactics for other computer simulation models. It is a firm conviction of the authors, that where tactics are determined by educated guess, the experimenter may inadvertently penalize a system by his choices. Allocating or placing the defensive units by the use of the optimal allocation model on the other hand, allows each different system to capitalize on its strengths and minimize its weaknesses.

The concept of the proposed comparative model can now be stated. The philosophy followed is:

1. Determine the optimal defense system deployment for each system to be considered, based upon its own characteristics and the terrain features of the sector to be defended.
2. Determine the optimal attack routes against each defense system which minimizes the risk to the attacking aircraft.
3. Determine the risk incurred by the attacker for each defense system to be compared.
4. The defense system that maximizes the enemy's risk is the preferred system.

Maximization of risk to the enemy has been chosen as the measure of effectiveness for a very straightforward reason. The purpose of the air defense system is to protect field army value units such as supply depots, vehicle concentrations, artillery positions, troop concentrations, etc. The purpose of any offensive weapons system is to destroy a given set of targets (value units) with the least possible cost. It is a generally accepted fact that a defensive system cannot prevent a determined and powerful offense from destroying a given number of these targets if the offense is willing to pay the price. The defense objective then is to try to extract a high cost from the offense. In gaming theory terms then, the goal of the defense is to maximize the offensive cost while minimizing the defensive cost. Both offensive and defensive costs are direct functions of the risk incurred by the offense in carrying out its attack.

The proposed Air Defense Comparative Model is a Game Theoretic Model utilizing the maxi-min principle of optimality. [2] Stated simply the defense chooses that strategy

which maximizes minimum risk while the offense chooses that strategy which minimizes maximum risk. The value of the game is then calculated for each defensive system to be compared and the one which extracts the highest risk to the attackers, is the preferrable system.

An overall schematic of the model is shown in Figure 1. Due to space limitations, it will not be possible to give detailed descriptions of the sub-models in this paper. However detailed descriptions of the component sub-models, including computer programs, may be found in references 3, 9, 10, and 11. Short descriptions of the sub-models are given in the following sections.

MAVD MODEL

Basic to the proposed Air Defense Comparative Model is the visibility subroutine called MAVD (Minimum Altitude Visibility Diagram). MAVD is a new concept and subroutine for calculating the visibility of targets to the defensive system sensor units [3].

The input to the MAVD Model is an array of digitized topographic data which is stored on magnetic tape. The Army Map Service has expended a considerable amount of time and effort in the digitization, and storing on magnetic tape, of topographic data. An m x n grid of horizontal (m = 1,2,...,i) and vertical (n = 1,2,...,j) lines is overlaid over the topographical map of the piece of terrain of interest. The spacing or grid interval between the lines is equal. The standard army battle map is a transverse Mercator projection of the Gauss-Kruger type [4]. The primary coordinate system for the map is a square grid system called the Universal Transverse Mercator grid [5]. Points of interest can be located on the map by their UTM grid coordinates. The UTM grid will appear on any map as a

square grid system where the numerical values of the coordinates of a point are positive, and increase as one moves the point east and north. For good terrain definition, it has been found that the grid spacing should not exceed about 1,500 feet or 300 meters. The local altitudes above sea level for the grid points thus defined are read off the topographical map and entered along with their grid point designation (i,j) as inputs.

MAVD (Minimum Altitude Visibility Diagram) is a geographic representation of the minimum local altitude at which a target may fly above the local terrain and still be visible to the given air defense sensor. Thus a MAVD value of 150 feet at point 5, 45 (the i,j grid representation of a specific point on the terrain) means that any aircraft at 150 feet altitude or above is visible to the sensor, or conversely any aircraft below 150 feet altitude is not visible (either masked by terrain irregularities or the curvature of the earth) to the sensor.

The MAVD routine is used to compute all the MAVD values for every designated point (i.e., a point defined by the intersection of two grid lines) on a grid for all given sensor locations. Figure 2 represents an example of a MAVD display. The top figure (2a) is the original terrain map and the bottom figure (2b) is the MAVD display where each MAVD value is given for the corresponding point on the original terrain map. In three dimensions a surface through all the MAVD values could be represented and any aircraft on or above this surface is visible to the given sensor(s).

The values on the MAVD represent, as mentioned, the minimum altitude values at which a penetrator is visible to a sensor at any given grid point (intersection point represented by the intersection of an "i" and "j" line). The effect of the curvature of the earth's surface and all terrain irregularities are considered in the computation of these

values. The calculation procedure is straightforward and uses basic plane and analytical geometry techniques. The detailed computer program for this subroutine (written in Fortran IV) may be found in reference 3. The present program is capable of handling a 211x211 grid size.

DEFENSE ALLOCATION MODEL

The second sub-routine utilized is the Allocation program which provides a systematic, objective method for computing the optimal deployment of any air defense system. Allocation is defined, with respect to this paper, as the assignment (or placement) of air defense system sensors at specific points on the given piece of terrain. The optimal allocation is that deployment which maximizes the attacker's risk. It may be also thought of as that deployment which minimizes the probability that an attacking aircraft or missile penetrates the defensive system undetected.

A survey of the literature uncovered an almost negligible amount of effort towards devising any systematic, objective, assignment of sensor units to terrain. The majority of models surveyed assigned sensor locations at random or at best, use an educated guess based on an "analysis" of the terrain involved. This analysis consists of little more than looking at the terrain map and attempting to visualize the effect of placing a sensor at a certain point. Such methods of choosing sensor locations are far from optimum. It is highly subjective and consequently it is doubtful that any two people would choose the same locations.

The mathematical formulation of optimal deployment problems falls into a sub-class of non-linear, zero-one programming problems. Although this has been previously

recognized, [6] it has not been possible to apply the existing methods of zero-one programming to any practical size problem due to the severe limitations of these mathematical methods. A new procedure called Complementary Programming, was therefore developed as a part of this research [7] and is applicable to very large problem types. For example, it was used to compute the optimal deployment of radars within the continental United States. The results were then compared with those previously proposed by Smallwood [8] utilizing a much more involved and time consuming procedure. The Complementary Programming method achieved an improved deployment over Smallwood's "optimal method."

Tests conducted during the evaluation of an early version of the Allocation Program showed that optimal deployment was sensitive to both range and visibility. The tests showed that, for a large terrain area, the primary factor in deployment was range, and visibility was only secondary. This observation led us to divide the original allocation program into two separate programs. We have designated the first program as the Coarse Allocation Program and the second as the Fine Allocation Program. The main concern of the Coarse Allocation Program is the minimization of range distance while the purpose of the Fine Allocation Program is the maximization of radar system visibility. The two programs are then used sequentially (see Figure 3). A good analogy to this method is the process used in turning to a station on the radio. One first turns the selector to the vicinity of the station in one rapid motion. When the station vicinity on the dial is reached, you then fine tune the selector until the station is optimally received. The Coarse Allocation Program achieves an initial, coarse deployment based primarily on range considerations. This coarse deployment is then used in conjunction with the Fine Allocation Program to achieve a new final deployment based primarily on visibility considerations. The sequential operation of the two

programs thus provides a final deployment that has both minimized the range from any attacker to a systems radar and maximized the combined radar system visibility.

It now becomes necessary to define a measure of the "goodness" of the coverage or visibility of the sensor for the points within its defined sector. Visibility was previously defined as a measure of the ability of a sensor or sensors to detect a target (or targets) within the air space over a given terrain area. This measure can be represented by a range of numerical values from zero to one and will be called the Visibility Value. A value of zero will be defined as there being no visibility over a given grid point for a specified altitude range. For example, if a grid point is not within the sector or range of a certain sensor, a zero is given to the Visibility Value for that sensor for the grid point. Another example of zero Visibility Value would be if the MAVD value for a grid point (for a given sensor) was 10,000 feet and the probability of an attack at that altitude or above was zero. We then would assign a zero to the Visibility Value. A value of one would require that, for the grid point, there exists visibility for all possible altitudes of attack.

The method used to convert MAVD values to visibility values is simple. First a limit is set on the altitude values of interest. Since the emphasis for this paper is on low altitude attacks and since an attack at high altitudes is visible to almost any sensor allocation, it is unnecessary to consider any altitudes above a specified ALT MAX. ALT MAX will be assigned a value for which there is (a) essentially zero probability of attack at altitudes $\geq$ ALT MAX or (b) considering terrain altitudes and irregularities, there is an almost certain probability of detection of any targets above ALT MAX.

The Visibility Value would then be calculated as:

$$\text{Visibility Value} = \frac{\text{ALT MAX} - \text{MAVD}}{\text{ALT MAX}}$$

where if the MAVD exceeds ALT MAX we assign a Visibility Value of zero i.e., we do not allow negative Visibility Values. Thus the Visibility Value is proportional to the per cent of air space that the sensor can see between the point on the surface of the local terrain and ALT MAX.

The Visibility Values as computed are then written on the computer drum in the order shown below:

Visibility Values on Drum

B Grid Point \ Sensor	1	2	3.	4	5	.....	240
1,1	.000	.900	.905	1.000	.500	.....	.000
1,2	.100	.000	.810	.150	.400	.....	.200
1,3	.000	.800	.000	.300	.200	.....	.800
.	.	.	.	.	.	.....	.
.	.	.	.	.	.	.....	.
.	.	.	.	.	.	.....	.
50,50	.900	.900	.000	.600	.750	.....	1.000

Thus each column represents the Visibility Values for a possible sensor location for all points on the grid. Our objective is then to combine a specified number of the above possible sensors so that the resulting "sum" of their coverage is maximum.

As perviously mentioned a new heuristic programming method called Complementary

Programming was derived for accomplishing this. The method is based upon the basic principle of the union of sets from set theory, where each of the columns in the above table of Visibility Values is considered an ordered set. The development, justification and complete computer programs for accomplishment can be found in references 9 and 10. The present computer program will handle a situation with 199x199 grid size over the terrain, 38,601 possible candidate radar locations and/or 463,212 possible candidate location - pointing angle combinations (if radar has less than 360° azmith capability).

ATTACK ROUTE MODEL

Having determined the optimal deployment of the defense, the next step is to turn our attention to the offense. As stated earlier, the problem of the offense may be stated as, "given an airspace over a specific piece of terrain that is defended by a ground based air defense system, find the least risk route that may be taken over this terrain to reach an assigned target. Based on the mini-max principle, it is assumed that the air defense system is optimally deployed over the terrain and that the offense has complete knowledge of both the defense system deployment and capabilities. The least risk route solution would then specify at which point(s) to enter the defended air space, the path to follow through the air space to the target, and the probability of survival. Such a computer model has been developed [11] and will now be briefly described.

A survey of the literature showed that very little had been accomplished in the area of the systematic, objective determination of optimal attack routes. Furthermore, of the few methods proposed, none was capable of handling anything except very small problems. It was therefore, decided to provide a relatively new approach rather than

try to build onto or refine an older approach. The optimal attack route problem was formulated in terms of a classical network problem. This was a natural approach in view of the grid overlay used for the terrain description in the MAVD and Defense Allocation models. With regards to the network description, we can state our problem as: "Determine the least risk route through the network, where we may enter the network at any outer node (intersection of grid lines) and travel on any branch (grid line between nodes) in either a forward or lateral direction.

The risk in traveling from one node to another in the network is then expressed as follows:

$$R = f(V, R_a, t)$$

where:

R = Risk

V = Visibility factor (i.e. is the target visible or not)

R_a = Range from target to defense system

t = The time in which the target is visible to the defense system.

Consequently, a least risk path would in general minimize the time the target is visible, minimize the number of times the target is visible, and maximize the range to the defense system (for the times in which the target is visible). Under this description, each node of the network may be assigned a value of risk. The "cost" of going from one node to the other is then the difference in risk from one node to the next, or the probability of survival from one node to the next.

Under the network formulation, several methods for solving the classical "shortest route through a network" are available. The most efficient methods are linear programming

and dynamic programming. While either of these methods can solve a small problem, it was found necessary to utilize dynamic programming for the larger real world problem because of computer storage requirements. For example, the solution of a problem by linear programming (Hungarian algorithm) would require storage of at least N^2 points (where N is the number of grid points). The storage requirements of dynamic programming are more on the order of 4N. Since determination of a flight path is a multi-stage decision process, dynamic programming was particularly well suited to the problem.

The method of dynamic programming is discussed in detail in the literature [12] and thus will be touched on only briefly. Generally speaking, dynamic programming is a method of solving multi-stage decision decision problems. Unlike linear programming, there is no standard mathematical formulation of the problem. It is a general approach and the particular equations used must be developed to fit each separate case at hand.

As stated, we use the same grid overlay as used in the MAVD and Allocation Models. But under our dynamic programming formulation we let each row (i.e. nodes i,j with i constant) represent a decision stage (see Figure 4). Each stage in turn, has a number of states associated with it. In our case, the states of each stage are simply the nodes of each row. In general, the states are simply the various possible positions in which the aircraft might be at any stage of the problem. In a multi-stage problem with discrete stages, (as in this problem) decisions are to be made at the beginning of the stages. The policy decision to be made at each stage is the destination for the next stage i.e. which state in the next stage. It is dependent upon the situation at the time of decision, upon the decision itself and upon the stage of the system. Each decision

affects not only the next stage but all subsequent stages. The solution of the problem is a sequence of decisions that yields the least risk route. This is essentially Bellman's principle of optimality, "An optimal policy has the property that whatever the initial state and initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."

The particular version of dynamic programming used in this problem computes the flight path in a "backward" manner. That is, one first starts at the target and then determines the optimal paths from each state in the previous stage to the target. Once this is done, the optimal paths from each state in the M-2 stage to the M-1 stage is computed. At each stage only the values of optimal paths need to be stored. This procedure is repeated until we are at the initial stage (i.e. row one). At this point all of the optimal paths from any of the entry points to the target are available.

As with the allocation model discussed in the previous section a two phase sequence is used. The first phase or calculation of the course attack route is primarily predicated on minimizing visibility (or risk) and the second phase or calculation of the fine course route is primarily to minimize exposure time. The solution procedure requires data in the form of two matrices. These matrices are (a) visibility matrix and (b) missile flight time matrix.

The visibility matrix provides the probability of detection for each node of the terrain sector. The aircraft altitude, h, and the MAVD values for each node are compared. If the MAVD value is greater than h, the aircraft is invisible and the risk is zero. It should be noted that the MAVD value to be compared with the aircraft altitude is always the minimum value of all the radar sites within range of the node. If the MAVD value

is less than h, then the visibility value for that node has some probability value associated with it. True aircraft visibility is not just a line of sight yes or no variable. Range from the target, reflective surface of the target, transmitter power, size of the antenna, etc. are all variables which may affect the visibility of a target to a radar. A review of the variables affecting visibility indicates that the effect of each is dependent to a large degree, upon the range of the aircraft from the radar. For this reason, the range of the target from the radar having the best MAVD value was selected as the best single variable to measure visibility. The relationship of range to probability of detection can be expressed graphically and is obtained from an analysis of the performance specifications of the missile system under consideration. The range value is thus converted to a visibility probability value based on each missile systems specifications. The risk is set equal to the probability value for the specified range. The detailed development of the model with the procedure coded in FORTRAN V language is given in reference 11.

SUMMARY

This paper has proposed a new and unique approach for conducting comparative experiments or evaluations between air defense weapon systems. The sub-models which were briefly discussed were developed as means of improving existing digital, computer stimulation experiments. It is believed, however, that these submodels and developed methodologies can be utilized as the basis for a completely independent, "unified air defense system comparison model." Such a model could be used for realistically analyzing and evaluating air defense systems in what we believe would be a far more economical and accurate manner than is presently available from simulation models.

ACKNOWLEDGEMENT

The sub-models discussed in this paper were developed as a part of the research under contracts DA-AH01-67-C1630 and DA-AH01-69-C1015, "Air Defense Modeling Studies," supported by the Systems Analysis Office, Future Missile Systems Division, Research and Engineering Directorate of the U. S. Army Missile Command, Redstone Arsenal, Alabama.

REFERENCES

1. B. Kovit, "Low-Altitude Penetration," Space/Aeronautics, 76-83, (May 1965).

2. J. Von Neuman and O. Morgenstern, Theory of Games and Economic Behavior, Princeton, University Press, Princeton, N. J., 1947.

3. R. E. Shannon and J. P. Ignizio, "Minimum Altitude Visibility Diagram--MAVD," UARI Report No. 50, University of Alabama Research Institute, April 1968.

4. M. R. Weldon, "Map Oriented Terrain Simulation," U. S. Army Missile Command Report No. RF-TM-64-34, November 1964.

5. "Universal Transverse Mercator," Technical Manual No. T. M. 5-241-8, Headquarters Department of the Army, July 1958.

6. B. J. Dunn et. al., "A Model for Allocating Interceptors with Overlapping Batteries: A Method of Nonlinear Programming," Braddock, Dunn, and McDonald, Inc., Report BDM-241-68-T, April 1968.

7. J. P. Ignizio and R. E. Shannon, "The Complementary Programming Method for Solving Certain Non-Linear, Zero-One Programming Problems," UARI Report No. 52, University of Alabama Research Institute, August 1968.

8. R. D. Smallwood, "Minimax Detection Station Placement," Operations Research, Vol. 13, 632-646, July-August 1965.

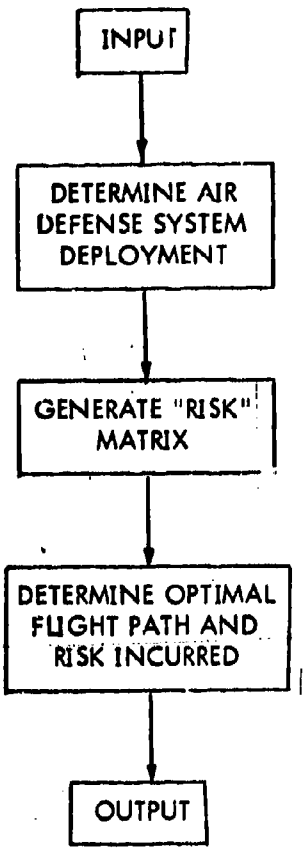
9. R. E. Shannon and J. P. Ignizio, "A Method to Achieve Optimum Air Defense Sensor Allocation," UARI Report No. 51, University of Alabama Research, April 1968.

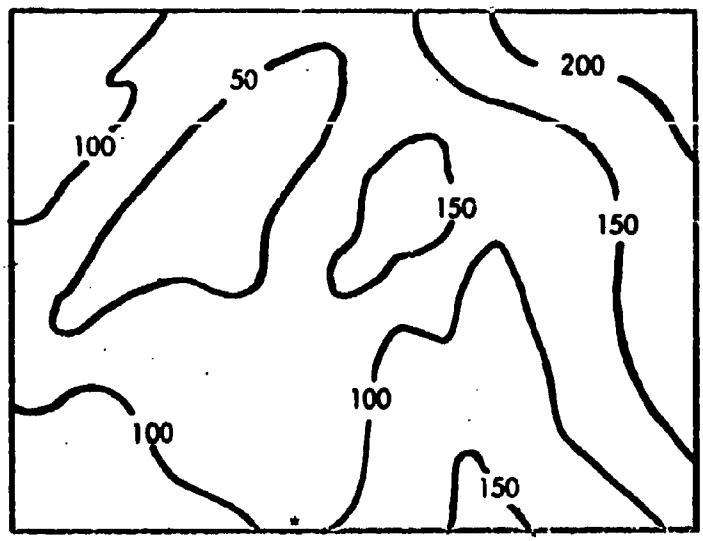
10. J. P. Ignizio and R. E. Shannon, "Improved Visibility and Deployment Models for Air Defense Studies," UARI Report No. 74 , University of Alabama Research Institute, February 1970.

11. J. P. Ignizio, R. E. Shannon and J. L. Stimach, "Determination of Optimal Aircraft Attack Routes," UARI Report No. 75 , University of Alabama Research Institute, February 1970.

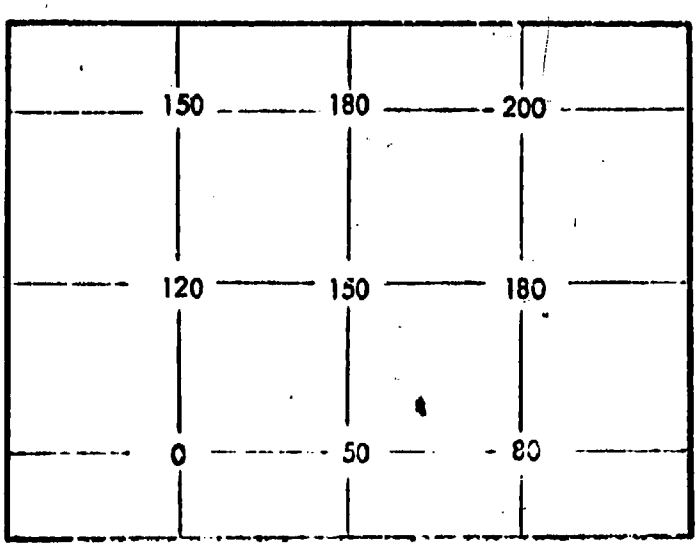
12. R. Bellman, Dynamic Programming, Princeton University Press, Princeton, N. J. 1957.

FIGURE 1


INPUT
DETERMINE AIR DEFENSE SYSTEM DEPLOYMENT
GENERATE "RISK" MATRIX
DETERMINE OPTIMAL FLIGHT PATH AND RISK INCURRED
OUTPUT

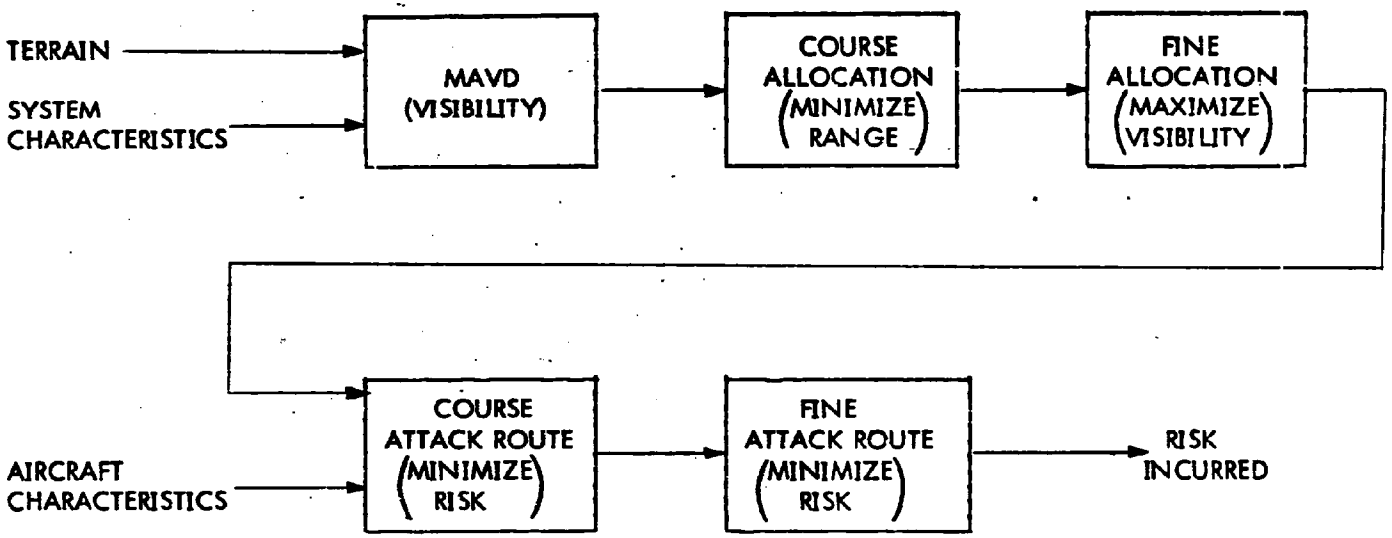


2a: TERRAIN MAP



2b: MAVD

FIGURE 2: MAVD



SCHEMATIC - GAME THEORY

COMPARATIVE MODEL

FIGURE 3

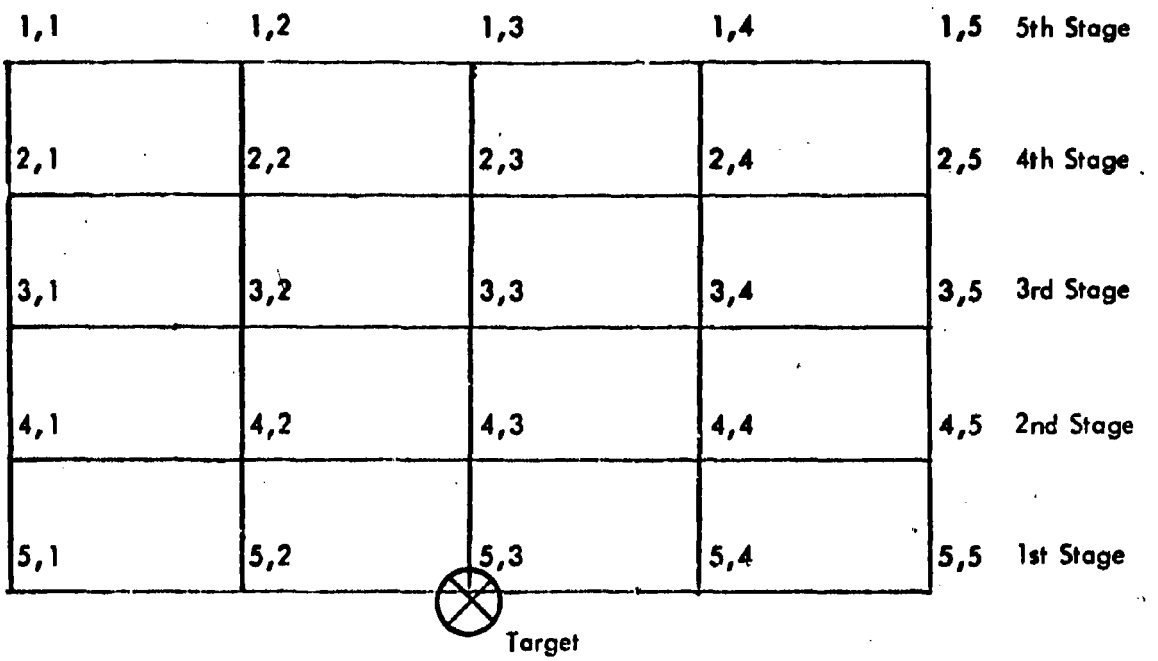

1,1
1,2
1,3
1,4
1,5
5th Stage
2,1
2,2
2,3
2,4
2,5
4th Stage
3,1
3,2
3,3
3,4
3,5
3rd Stage
4,1
4,2
4,3
4,4
4,5
2nd Stage
5,1
5,2
5,3
5,4
5,5
1st Stage
Target

FIGURE 4

PROBABILISTIC MANPOWER PLANNING FOR THE RESEARCH AND DEVELOPMENT ORGANIZATION

Larry H. Johnson
Redstone Arsenal, Alabama

INTRODUCTION

This paper proposes a statistical approach to one of the problems confronting all Army research, development and testing organizations. That is, manpower planning.

Though the problem addressed involves manpower, it should be noted that the mathematical techniques are applicable to all types of inventory by redefining the paramaters involved.

The question for long-range planning is not what should be done tomorrow, but rather what can be done today to cope best with the uncertain tomorrow. Management must understand the alternatives available to them, the risk associated with each, and choose rationally among the alternatives rather than plunge into uncertainty only on the basis of intuition or previous experience.

DEFINITION OF PROBLEM

If a given organization has a large number of programs planned for the future, it is usually reasonable to assume that the manpower requirement is somewhat normally distributed. However, most R & D organizations do not have a large number of outstanding programs and, therefore, the gain or loss of a single program can have gross affects on the required manpower. This problem requires that the "exact" probability distribution be known and solutions for this problem are not available in the literature.

The current need for management planning techniques with relatively few outstanding programs motivated the study described herein.

ASSUMPTIONS

This study makes four basic assumptions:

1. First it is assumed that the organization will not be required to perform every program for which current planning exists and that a subjective probability can be associated with the gain or loss of each program.

This article has been reproduced photographically from the author's manuscript.

3. Here it is assumed that each program is statistically independent. That is, the initiation or cancellation of one program does not affect the probability of any other program.

3. Third, it is assumed that personnel of common disciplines are reasonably interchangeable. For example, any mechanical engineer or civil engineer (within reason) could be utilized on any program where such skill is required.

4. Fourth, each program being considered has a proposed initiation date, and if this date is uncertain, it is assumed that a conditional probability for starting on various dates can be established.

DATA NEEDS

First, for each outstanding program, there must be a program plan which delineates each task to be performed, the time phasing, and the sequence in which tasks must be conducted. From this plan, the organization generates a time adjusted manpower array for performing all tasks under the development plan.

PROBLEM SOLUTION

In effort to promote appreciation of the problem at hand, an example problem will be utilized.

From the development plan, a time adjusted manpower array for performing all tasks has been generated and is presented in Figure 1. It is assumed certain that this project will be conducted, but the start date is uncertain. Utilizing the theory of Expected Values:

$$E(X_t) \quad (P_i)(P_{t/i})(X_{t/i}) + (P_j)(P_{t/j})(X_{t/j}) = \sum_{j=1}^{k} (P_i)(P_j)(X_j)$$

where $j = t/i$

and when "n" projects are considered:

$$E(X_t) = \sum_{i=1}^{n} \sum_{j=1}^{k} (P_i)(P_j)(X_j)$$

where: X_t = the units of manpower required during period t
P_i = capture probability

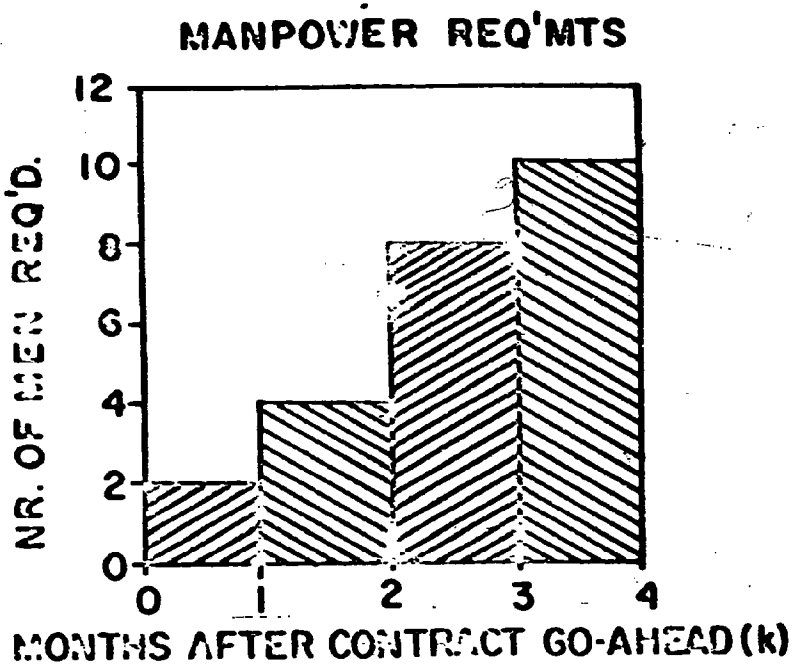
P_j = probability that project "i" starts in period "t"

X_j = number of manpower units required for project "i" during the period "t"

This process is illustrated graphically in Figures 2, 3 and 4.

Figure 1 - Time Adjusted Manpower Array and Probable Start Date

PROBABILITY OF CAPTURE (P_i) = 1



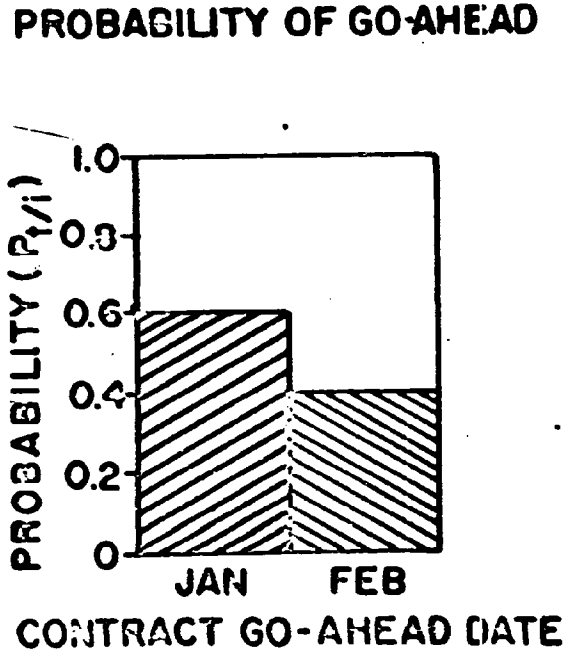
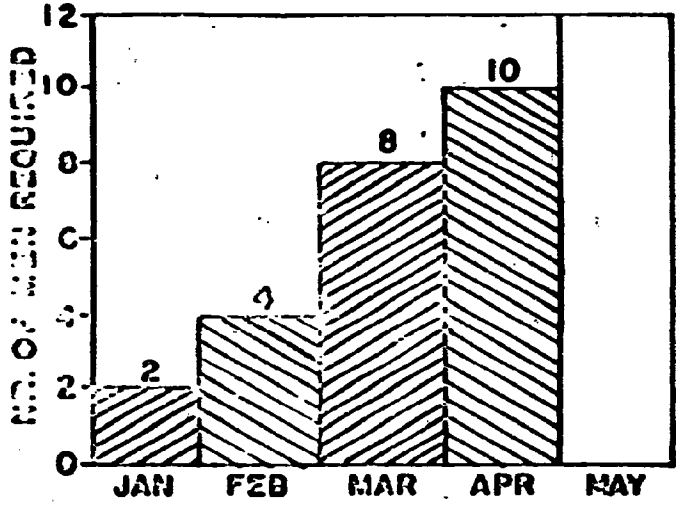
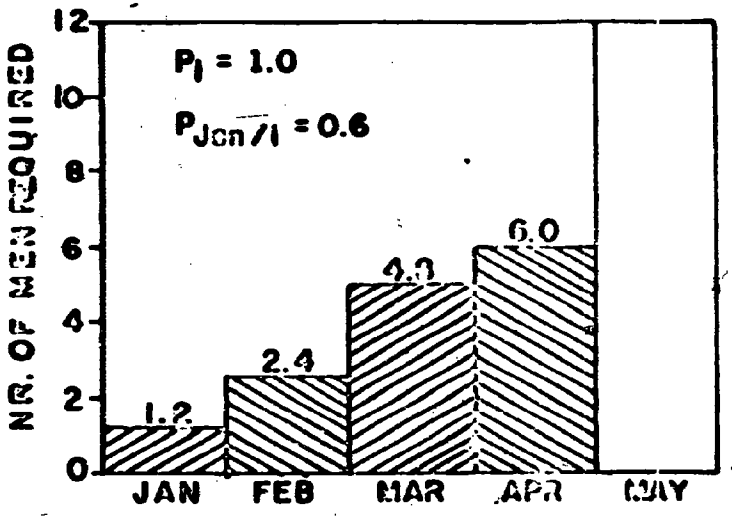


Figure 2 - Actual and Expected Manpower Requirements for January Start Date

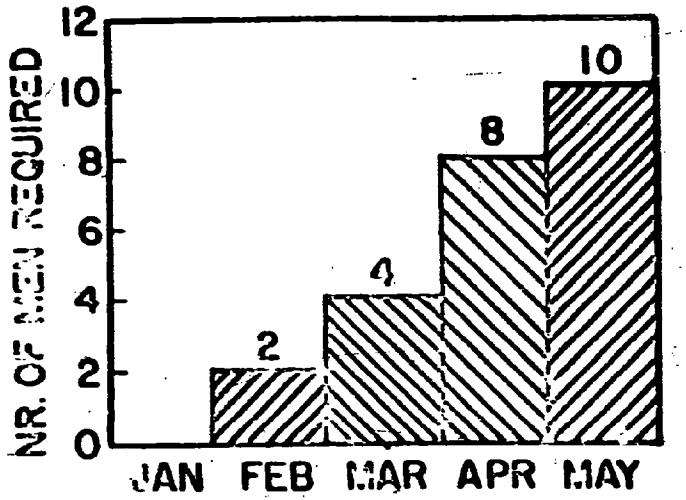


CONTRACT I MANPOWER REQ'MTS: JANUARY GO-AHEAD

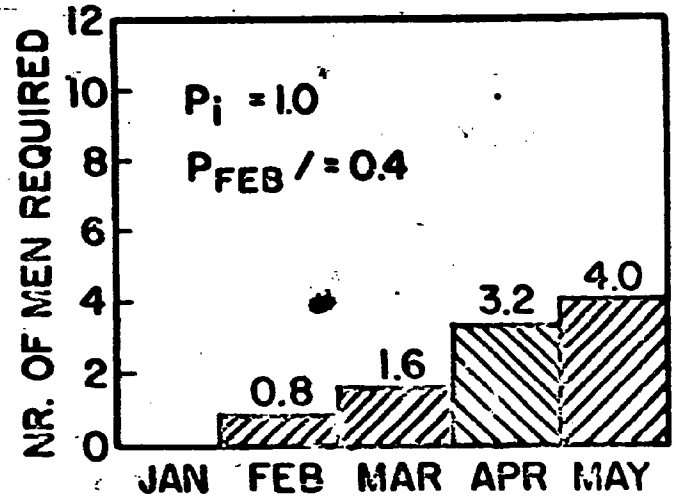


CONTRACT I EXPECTED MANPOWER REQ'MTS: JANUARY GO-AHEAD

Figure 3 - Actual and Expected Manpower Requirements for February Start Date

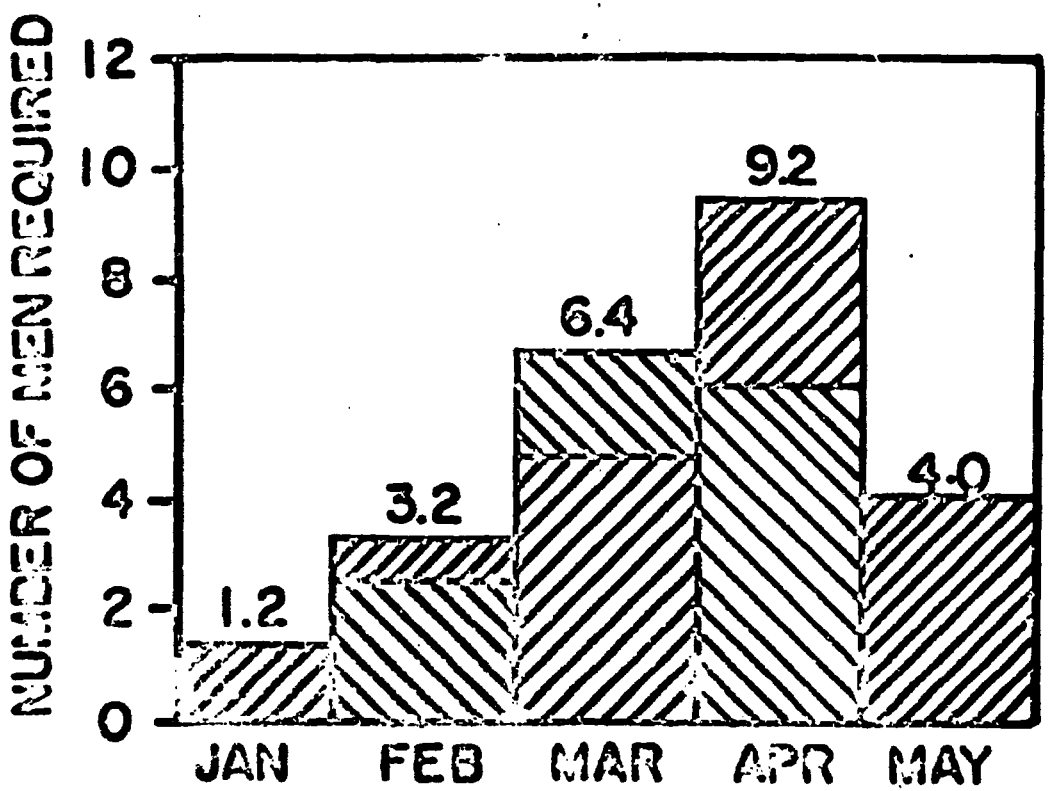


CONTRACT I MANPOWER REQ'MTS: FEBRUARY GO-AHEAD



CONTRACT I EXPECTED MANPOWER REQ'MTS: FEBRUARY GO-AHEAD

Figure 4 - Total Expected Manpower Requirement for Contract I



At this point the problem solution seems quite simple; however, because we used expected values, the "risk" associated with the solution cannot be evaluated unless there is a large number of outstanding projects (which usually is not the case for an R&D Organization). We must therefore develop a model which utilizes the entire statistical distribution. For this situation an enumeration process has been developed and is illustrated in Figure 5. Every possible workload is identified and the probability of occurrence for each is evaluated. Considering that the go-ahead date for the programs may not be "fixed" but rather can be expressed as a probabilistic function, we expand the enumeration procedure as illustrated in Figure 6 where the additional uncertainty is accounted for by t_1, t_2, etc.

A technique has thus been developed for enumerating the total array of possible workloads for an organization and the probability associated with each. This concept can readily be adapted to fit any particular problem that one may have.

MINIMUM COST METHOD

Given the total array of possible manpower requirements developed above, the corporation is now faced with the problem of determining the most economical method of performing any given workload. That is, if management were to assume that they knew specifically which one of the workloads will occur, how can they most economically perform the task realizing that if the workload exceeds capacity, they may choose to:

1. hire additional employees
2. work on an overtime basis or
3. subcontract a portion of the work.

If capacity exceeds the workload, management may choose to:

1. continue on an over-staffed basis or
2. lay-off some employees

Management must recognize that with each alternative, the total cost for conducting the workload will vary. For example, new employees must be trained, overtime costs premium wage, employee efficiency decreases with overtime and subcontract personnel may not be as effective as regular employees.

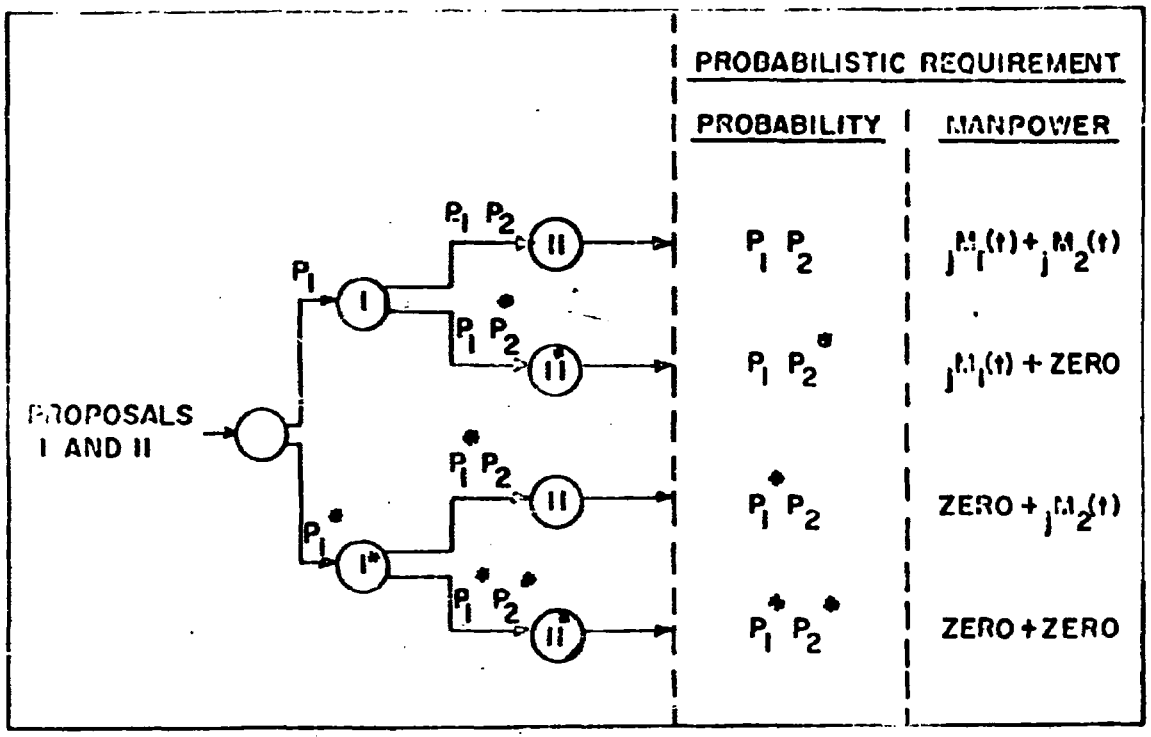
Total capacity for the organization during any period is therefore given by:

$$Z_t = \lambda_1 X_{1,t} + \lambda_2 X_{2,t} + \lambda_4 X_{4,t} + \lambda_5 X_{5,t} + \lambda_6 X_{6,t}$$

where:

λ_m is the efficiency factor for each type of manpower (overtime, subcontract, etc.)

Figure 5 - Enumeration Process When Initiation Date is Fixed
(Two Proposals)



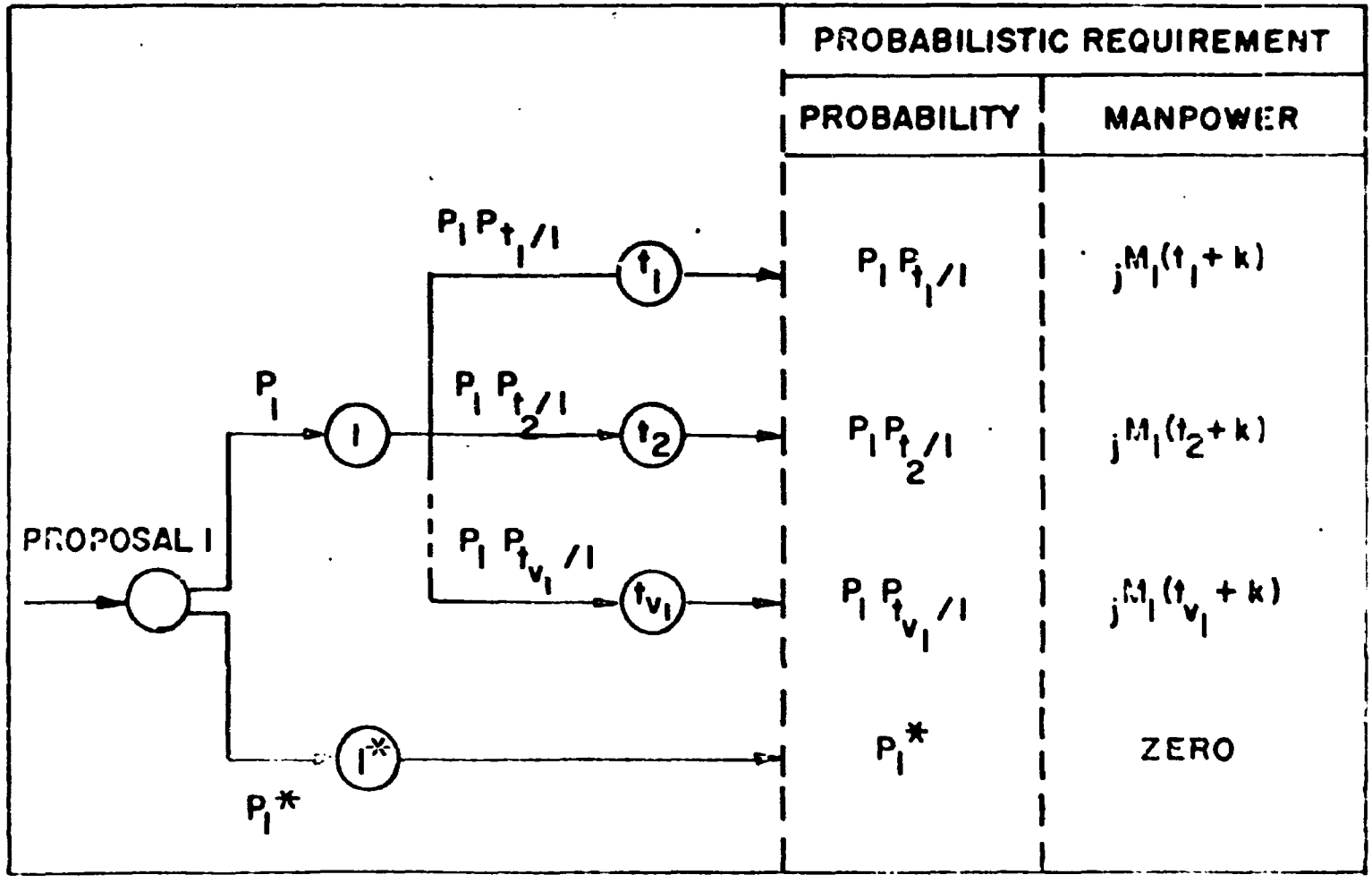
NOTE: P_1 = PROBABILITY OF GETTING PROGRAM

P_1^* = PROBABILITY OF NOT GETTING PROGRAM

$P_1 + P_1^* = 1$

${}_jM_1(t)$ = NR. OF TYPE j PERSONNEL FOR JOB 1 DURING PERIOD t

Figure 6 - Enumeration Process for Variable Initiation Date
(One Proposal)



$X_{1,t}$ is the effective capacity of manpower available during period t.

$X_{1,t}$ is the number of experienced employees available during period t.

$X_{2,t}$ is the number of new employees available during period t.

$X_{4,t}$ is the number of overtime units which can be worked by experienced employees during period t.

$X_{5,t}$ is the number of overtime units which new employees can work during period t.

$X_{6,t}$ is the number of subcontract personnel available during period t.

Note that $X_{3,t}$, which is the number of full time employees to be terminated during period t, is not included in the equation. Unit costs for each type of manpower must also be available.

The problem is one of determining the optimum manpower schedule for a given workload which permits the organization to operate for the duration of the planning period with minimum labor costs.

Requirements and data inputs for the minimization problem are ideal for solution by the Simplex Linear Programming Technique where the constraints, due to management policies, labor agreements, etc., limit the range of values for X_i. The Simplex not only provides a manpower plan for each workload but also the total cost for each plan.

Note that in some cases the Simplex may indicate to hire employees one month and terminate them the next. This is not fault with the mathematics but rather fault with management policies. The Minimum Cost Technique is therefore a good indicator for restraining management labor policies.

MINIMUM RISK METHOD

In the previous discussion it may be noted that the probability associated with each workload was ignored. How then can the element of risk be considered in the management plan?

Given the probabilistic manpower requirements, management needs a decision-making policy which allows them to plan for a theoretical workload and adjust with minimum consequence to the actual workload when it occurs. The Minimum Risk Method will provide such a plan.

Considering all of the possible workloads, the notion of risk is introduced where risk for a given workload is defined as the cost of making a transition, in later intervals, from some planned manpower level to the appropriate Minimum Cost Plan. In short, risk is the cost of adjusting to the workload that actually occurs. Once the workload becomes known, an adjustment or transition is made and the appropriate Minimum Cost schedule is followed.

The problem now is to identify a manpower planning level which minimizes total risk for the enumerated range of workloads.

Letting C_k^* represent the minimum cost for the k^{th} workload and C_k' represent the cost of adjusting from the planned level to the appropriate Minimum Cost Plan and completing the job, then the risk R_k is given by

$$R_k = C_k' - C_k^*$$

The expected risk (R_k') for each of the k workloads is given by

$$R_k' = P_k R_k = (P_k)(C_k' - C_k^*)$$

where P_k is the probability of occurrence determined by enumeration. The total risk R(•) for any manpower plan is therefore given by

$$R(\cdot) = \sum_k R_k' = \sum_k P_k(C_k' - C_k^*)$$

Our problem is now to identify a manpower plan which minimizes R(•). A dynamic programming technique for minimizing R(•) has <u>not</u> been developed; however, we can iterate a solution as illustrated in the following problem.

Example Problem

Suppose the scheduling period is 4 months and the initial number of experienced employees is 60. The organization has two outstanding project with estimated capture probabilities of 0.6 and 0.3 respectively and it will not be known until 1 January if the projects will be funded. The type A manpower requirements for each project are presented in Table 1. The projected manpower requirement without consideration of the two new contracts is also shown.

Table 1. [illegible] Manpower Requirements; Minimum Risk Problem

Identification of Manpower Requirements	Probability	January	February	March	April
Workload without new contracts	1.0	50	70	90	60
Contract I	0.6	10	10	20	30
Contract II	0.3	30	20	20	10

[illegible] company policies and labor agreements, constraints exist such [illegible]

$$0 \leq X_{1,t} \leq 100$$

$$0 \leq X_{2,t} \leq 20$$

$$0 \leq X_{3,t} \leq 20$$

$$0 \leq X_{4,t} \leq 30$$

$$0 \leq X_{5,t} \leq 30$$

$$0 \leq X_{6,t} \leq 20$$

[illegible] costs and efficiency factors for the various types of manpower are presented in Table 2. It is also assumed that new employees can be trained within one time period.

Table 2. Unit Costs and Efficiency Data for Minimum Risk Problem

Unit Costs ($)	Identification	Efficiency Factors
$C_1 = 100$	Experienced employees	$\lambda_1 = 1.00$
$C_2 = 130$	New employees	$\lambda_2 = 0.50$
$C_3 = 50$	Mandatory terminations	
$C_4 = 150$	Experienced employee overtime	$\lambda_4 = 0.70$
$C_5 = 150$	New employee overtime	$\lambda_5 = 0.35$
$C_6 = 170$	Subcontracts	$\lambda_6 = 0.80$

The problem is to determine theoretical values for $x_{1,1}, x_{2,1}$, etc., which minimize the total risk for this planning situation.

Referring to the complete enumeration technique, there are four possible workloads with probabilities of occurrence as calculated and shown in Table 3. The Minimum Cost plan for each of the four possible workloads, determined by the Simplex Technique, is presented in Table 4.

Since the primary interest for planning is full time employees, and the Minimum Cost plans of Table 4 indicate that portions of the workload should be subcontracted, the workload and Minimum Cost arrays are modified to reflect only the in-house efforts. The in-house efforts are determined by subtracting the subcontracts from Tables 3 and 4, and results of this operation are presented in Tables 5 and 6.

Table 3. Enumerated Workload Requirements for Minimum Risk Problem

Possible Workloads	Probability of Occurrence	Manpower Requirements			
		January	February	March	April
$W_1 = M_1^* M_2^*$	$p_1^* p_2^* = 0.28$	50	70	90	60
$W_2 = M_1 M_2^*$	$p_1 p_2^* = 0.42$	60	80	110	90
$W_3 = M_1^* M_2$	$p_1^* p_2 = 0.12$	80	90	110	70
$W_4 = M_1 M_2$	$p_1 p_2 = 0.18$	90	100	130	100

Table 4. Minimum Cost Manpower Plans for Minimum Risk Problem

Minimum Cost Plan for	Probability	January	February	March	April	Total Cost ($)
W_1	0.28	$X_{1,1} = 60$	$X_{1,2} = 60$ $X_{2,2} = 20$	$X_{1,3} = 80$ $X_{6,3} = 12.5$	$X_{1,4} = 60$ $X_{3,4} = 20$	31,725
W_2	0.42	$X_{1,1} = 60$ $X_{2,1} = 10$	$X_{1,2} = 70$ $X_{2,2} = 20$	$X_{1,3} = 90$ $X_{4,3} = 5.71$ $X_{6,3} = 20$	$X_{1,4} = 90$	39,157
W_3	0.12	$X_{1,1} = 60$ $X_{2,1} = 20$ $X_{6,1} = 12.5$	$X_{1,2} = 80$ $X_{2,2} = 20$ $X_{6,2} = 6.25$	$X_{1,3} = 90$ $X_{4,3} = 5.71$ $X_{6,3} = 20$	$X_{1,4} = 70$ $X_{3,4} = 20$	42,344
W_4	0.18	$X_{1,1} = 60$ $X_{2,1} = 20$ $X_{4,1} = 5.71$ $X_{7,1} = 20$	$X_{1,2} = 80$ $X_{2,2} = 20$ $X_{6,2} = 12.5$	$X_{1,3} = 100$ $X_{4,3} = 20$ $X_{6,3} = 20$	$X_{1,4} = 100$	51,982

Table 5. In-House Workload Requirements for Minimum Risk Problem

Possible In-House Workloads	Probability of Occurrence	January	February	March	April
W_1'	0.28	50	70	80	60
W_2'	0.42	60	80	94	90
W_3'	0.12	70	85	94	70
W_4'	0.18	74	90	114	100

Table 6. Minimum Cost Plans for In-House Workloads

Minimum Cost Plan for	Probability	January	February	March	April	Total Cost ($)
W_1'	0.28	$X_{1,1} = 60$	$X_{1,2} = 60$ $X_{2,2} = 20$	$X_{1,3} = 80$	$X_{1,4} = 60$ $X_{3,4} = 20$	29,600
W_2'	0.42	$X_{1,1} = 60$ $X_{2,1} = 10$	$X_{1,2} = 70$ $X_{2,2} = 20$	$X_{1,3} = 90$ $X_{4,3} = 5.71$	$X_{1,4} = 90$	35,757
W_3'	0.12	$X_{1,1} = 60$ $X_{2,1} = 20$	$X_{1,2} = 80$ $X_{2,2} = 10$	$X_{1,3} = 90$ $X_{4,3} = 5.71$	$X_{1,4} = 70$ $X_{3,4} = 20$	35,757
W_4'	0.18	$X_{1,1} = 60$ $X_{2,1} = 20$ $X_{4,1} = 5.71$	$X_{1,2} = 80$ $X_{2,2} = 20$	$X_{1,3} = 100$ $X_{4,3} = 20$	$X_{1,4} = 100$	43,057

The next procedure is to iterate costs and risks for all feasible values of $X_{1,1}$ and $X_{2,1}$ so that the minimum $R(\cdot)$ can be identified. It may be noted in Table 6 that $X_{1,1}$ is 60 for all four workloads; therefore, there is only one feasible solution for $X_{1,1}$. However, $X_{2,1}$ varies from 0 to 20. The problem then is to determine a value ($X'_{2,1}$) which minimizes $R(\cdot)$ keeping in mind that the objective is to adjust to the Minimum Cost plan in the most expedient and economic manner consistent with the manpower constraints.

The results of the iteration process for each of the four possible workloads is presented in Tables 7 through 10 where $X_{2,1}$ was varied from 0 to 20 in increments of 5 units. Increments of five units each were arbitrarily selected for simplification of calculations in this illustrative problem. Note the heavy line in each of the tables. This line indicates when the level of full time employment reaches the Minimum Cost plan for that particular workload. Total cost for each of the trial solutions is also presented in Tables 7 through 10.

Summary of the expected risk R'_k calculations with the cost data from Tables 7 through 10 is presented in Table 11. The term $R(\cdot)$ may then be calculated by the equation $R(\cdot) = \sum_k R'_k$ for each of the $X'_{2,1}$ solutions. The $R(\cdot)$ data are tabulated in Table 12 where it is shwon that $R(\cdot)$ is minimum when $X'_{2,1}$ is 10 units.

Table 2. [illegible] Risk Plan for Workload Risk at Beginning Contracts

Solution	Period: W	January 50	February 70	March 80	April 60	Cost C_1' ($)
	$X_{1,1}$ $X_{2,1}$					
$S_{1,1}$	60 0	$X_{1,1} = 60$	$X_{1,2} = 60$ $X_{4,2} = 20$	$X_{1,3} = 80$	$X_{1,4} = 60$ $X_{3,4} = 20$	29,600
$S_{1,2}$	60 5	$X_{1,1} = 60$ $X_{2,1} = 5$	$X_{1,2} = 65$ $X_{2,2} = 15$	$X_{1,3} = 80$	$X_{1,4} = 60$ $X_{3,4} = 20$	30,100
$S_{1,3}$	60 10	$X_{1,1} = 60$ $X_{2,1} = 10$	$X_{1,2} = 70$ $X_{2,2} = 10$	$X_{1,3} = 80$	$X_{1,4} = 60$ $X_{3,4} = 20$	30,600
$S_{1,4}$	60 15	$X_{1,1} = 60$ $X_{2,1} = 15$	$X_{1,2} = 75$ $X_{2,2} = 5$	$X_{1,3} = 80$	$X_{1,4} = 60$ $X_{3,4} = 20$	31,100
$S_{1,5}$	60 20	$X_{1,1} = 60$ $X_{2,1} = 20$	$X_{1,2} = 80$	$X_{1,3} = 80$	$X_{1,4} = 60$ $X_{3,4} = 20$	31,600

Table 3. Enhanced Risk Plan if Contract I is Received

Solution	Period: W	January 60	February 80	March 94	April 90	Cost C_2' ($)
	$X_{1,1}$ $X_{2,1}$					
$S_{2,1}$	60 0	$X_{1,1} = 60$	$X_{1,2} = 60$ $X_{2,2} = 20$ $X_{4,2} = 14.3$	$X_{1,3} = 80$ $X_{2,3} = 10$ $X_{4,3} = 12.9$	$X_{1,4} = 90$	36,980
$S_{2,2}$	60 5	$X_{1,1} = 60$ $X_{2,1} = 5$	$X_{1,2} = 65$ $X_{2,2} = 20$ $X_{4,2} = 7.1$	$X_{1,3} = 85$ $X_{2,3} = 5$ $X_{4,3} = 9.3$	$X_{1,4} = 90$	36,360
$S_{2,3}$	60 10	$X_{1,1} = 60$ $X_{2,1} = 10$	$X_{1,2} = 70$ $X_{2,2} = 20$	$X_{1,3} = 90$ $X_{4,3} = 5.7$	$X_{1,4} = 90$	35,757
$S_{2,4}$	60 15	$X_{1,1} = 60$ $X_{2,1} = 15$	$X_{1,2} = 75$ $X_{2,2} = 15$	$X_{1,3} = 90$ $X_{4,3} = 5.7$	$X_{1,4} = 90$	36,255
$S_{2,5}$	60 20	$X_{1,1} = 60$ $X_{2,1} = 20$	$X_{1,2} = 80$ $X_{2,2} = 10$	$X_{1,3} = 90$ $X_{4,3} = 5.7$	$X_{1,4} = 90$	36,755

Table 9. Enumerated Risk Plan if Contract II is Received

Solution	Period W_3'		January 70	February 85	March 94	April 70	Cost C_3' ($)
	$X_{1,1}'$	$X_{2,1}'$					
$S_{3,1}$	60	0	$X_{1,1} = 60$ $X_{4,1} = 14.3$	$X_{1,2} = 60$ $X_{2,2} = 20$ $X_{4,2} = 21.4$	$X_{1,3} = 80$ $X_{4,3} = 20$	$X_{1,4} = 70$ $X_{3,4} = 10$	38,455
$S_{3,2}$	60	5	$X_{1,1} = 60$ $X_{2,1} = 5$ $X_{4,1} = 10.7$	$X_{2,1} = 65$ $X_{2,2} = 20$ $X_{4,2} = 14.3$	$X_{1,3} = 85$ $X_{4,3} = 12.9$	$X_{1,4} = 70$ $X_{3,4} = 15$	37,685
$S_{3,3}$	60	10	$X_{1,1} = 60$ $X_{2,1} = 10$ $X_{4,1} = 7.1$	$X_{1,2} = 70$ $X_{2,2} = 20$ $X_{4,2} = 7.1$	$X_{1,3} = 90$ $X_{4,3} = 5.7$	$X_{1,4} = 70$ $X_{3,4} = 20$	36,917
$S_{3,4}$	60	15	$X_{1,1} = 60$ $X_{2,1} = 15$ $X_{4,1} = 3.6$	$X_{1,2} = 75$ $X_{2,2} = 15$ $X_{4,2} = 3.6$	$X_{1,3} = 90$ $X_{4,3} = 5.7$	$X_{1,4} = 70$ $X_{3,4} = 20$	36,335
$S_{3,5}$	60	20	$X_{1,1} = 60$ $X_{2,1} = 20$	$X_{1,2} = 80$ $X_{2,2} = 10$	$X_{1,3} = 90$ $X_{4,3} = 5.7$	$X_{1,4} = 70$ $X_{3,4} = 20$	35,757

Table 10. Enumerated Risk Plan if Contracts I and II are Received

Solution	Period	January	February	March	April	Cost
	W'_4	74	90	114	100	C'_4
	$X'_{1,1}$ $X'_{2,1}$					($)
$S_{4,1}$	60 0	$X_{1,1} = 60$ $X_{4,1} = 20$	$X_{1,2} = 60$ $X_{2,2} = 20$ $X_{4,2} = 28.6$	$X_{1,3} = 80$ $X_{2,3} = 20$ $X_{4,3} = 30$ $X_{5,3} = 8.6$	$X_{1,4} = 100$	48,280
$S_{4,2}$	60 5	$X_{1,1} = 60$ $X_{2,1} = 5$ $X_{4,1} = 16.4$	$X_{1,2} = 65$ $X_{2,2} = 20$ $X_{4,2} = 21.4$	$X_{1,3} = 85$ $X_{2,3} = 15$ $X_{4,3} = 30$ $X_{5,3} = 1.4$	$X_{1,4} = 100$	46,580
$S_{4,3}$	60 10	$X_{1,1} = 60$ $X_{2,1} = 10$ $X_{4,1} = 12.9$	$X_{1,2} = 70$ $X_{2,2} = 20$ $X_{4,2} = 14.3$	$X_{1,3} = 90$ $X_{2,3} = 10$ $X_{4,3} = 27.1$	$X_{1,4} = 100$	45,345
$S_{4,4}$	60 15	$X_{1,1} = 60$ $X_{2,1} = 15$ $X_{4,1} = 9.3$	$X_{1,2} = 75$ $X_{2,2} = 20$ $X_{4,2} = 7.1$	$X_{1,3} = 95$ $X_{2,3} = 5$ $X_{4,3} = 23.6$	$X_{1,4} = 100$	44,200
$S_{4,5}$	60 20	$X_{1,1} = 60$ $X_{2,1} = 20$ $X_{4,1} = 5.7$	$X_{1,2} = 80$ $X_{2,2} = 20$	$X_{1,3} = 100$ $X_{4,3} = 20$	$X_{1,4} = 100$	43,057

Table 11. Risk Summary for Example Problem

Solution	p_k	C'_k	C''_k	R_k	R_k
$S_{1,4}$	0.28	31,000	29,600	1500	420
$S_{1,5}$	0.28	31,600	29,600	2000	560
$S_{2,1}$	0.42	36,980	35.757	1223	514
$S_{2,2}$	0.42	36,360	35,757	603	253
$S_{2,3}$	0.42	35,757	35,757	0	0
$S_{2,4}$	0.42	36,255	35,757	498	209
$S_{2,5}$	0.42	36,755	35,757	998	419
$S_{3,1}$	0.12	38,455	35,757	2698	324
$S_{3,2}$	0.12	37,685	35,757	1928	231
$S_{3,3}$	0.12	36,917	35,757	1160	139
$S_{3,4}$	0.12	36,335	35,757	578	69
$S_{3,5}$	0.12	35,757	35,757	0	0
$S_{4,1}$	0.18	48,280	43,057	5223	940
$S_{4,2}$	0.18	46,580	43,057	3523	634
$S_{4,3}$	0.18	45,345	43,057	2288	412
$S_{4,4}$	0.18	44,200	43,057	1143	206
$S_{4,5}$	0.18	43,057	43,057	0	0

Table 12. Risk Analysis for Example Problem

$X^1_{1,1}$	$X^1_{2,1}$	ΣS	ΣR_k	$R(\cdot)$
60	0	$S_{1,1} + S_{2,1} + S_{3,1} + S_{4,1}$	0 + 514 + 324 + 940	1778
60	5	$S_{1,2} + S_{2,2} + S_{3,2} + S_{4,2}$	140 + 253 + 231 + 634	1258
60	10	$S_{1,3} + S_{2,3} + S_{3,3} + S_{4,3}$	280 + 0 + 139 + 412	(831)
60	15	$S_{1,4} + S_{2,4} + S_{3,4} + S_{4,4}$	420 + 209 + 69 + 206	904
60	20	$S_{1,5} + S_{2,5} + S_{3,5} + S_{4,5}$	560 + 419 + 0 + 0	979

What does this solution mean? The Minimum Risk plan for this illustrative problem is to retain the 60 experienced employees who will be available for January and, in addition, hire 10 new employees before January so that their services will be available during the first month. If workload I_1 occurs, the corporate plan will be as shown in Table 13. If workloads I_2, I_3, or I_4 occur, the corporate plan will be as shown in Tables 14 through 16 respectively.

Table 13. Minimum Risk Plan Without New Contracts

$X'_{1,1}$	$X'_{2,1}$	January	February	March	April
60	10	$X_{1,1} = 60$ $X_{2,1} = 10$	$X_{1,2} = 70$ $X_{2,2} = 10$	$X_{1,3} = 80$ $X_{6,3} = 12.5$	$X_{1,4} = 60$ $X_{3,4} = 20$

Table 14. Minimum Risk Plan if Contract I is Received

$X'_{1,1}$	$X'_{2,1}$	January	February	March	April
60	10	$X_{1,1} = 60$ $X_{2,1} = 10$	$X_{1,2} = 70$ $X_{2,2} = 20$	$X_{1,3} = 90$ $X_{4,3} = 5.7$ $X_{6,3} = 20$	$X_{1,4} = 90$

Table 15. Minimum Risk Plan if Contract II is Received

$X'_{1,1}$	$X'_{2,1}$	January	February	March	April
60	10	$X_{1,1} = 60$ $X_{2,1} = 10$ $X_{4,1} = 7.1$ $X_{6,1} = 12.5$	$X_{1,2} = 70$ $X_{2,2} = 20$ $X_{4,2} = 7.1$ $X_{6,2} = 6.25$	$X_{1,3} = 90$ $X_{4,3} = 5.7$ $X_{6,3} = 20$	$X_{1,4} = 70$ $X_{3,4} = 20$

Table 16. Minimum Risk Plan if Contracts I and II are Received

$X'_{1,1}$	$X'_{2,1}$	January	February	March	April
60	10	$X_{1,1} = 60$	$X_{1,2} = 70$	$X_{1,3} = 90$	$X_{1,4} = 100$
		$X_{2,1} = 10$	$X_{2,2} = 20$	$X_{2,3} = 10$	
		$X_{4,1} = 12.9$	$X_{4,2} = 14.3$	$X_{4,3} = 27.1$	
		$X_{6,1} = 20$	$X_{6,2} = 12.5$	$X_{6,3} = 20$	

The organization is hereby presented with a strategy for planning the future manpower requirements in the face of uncertainty. A mathematical simulation has been developed which can assist management in understanding the problem and the effects of various plans available to them. It is felt that this approach can be computerized and provide management with a rapid assessment of the situation at any given time.

ANALYSIS OF FACTORIAL EXPERIMENTS IN NON-CONNECTED BLOCK DESIGNS*

B. M. Ku kjian and R. C. Woodall
Harr Diamond Laboratories
Washington, D. C.

1. Introduction

The object of the analysis is to estimate the effect of v treatments on the response of a device in the presence of b extraneous side effects (blocks), whose effects are removed from the treatment effects. The treatments may be simple treatments involving only one factor, or they may be treatment-combinations involving several factors applied simultaneously to the device. In the latter case, the effect of each factor and the effects of interaction between factors are also estimated. The results presented here require no restrictions on the experimental design. That is, the design may be unbalanced, treatments may be missing, there may be an unequal number of observations per cell, etc.

2. Model

The model is the fixed effects model:

$$y_{ijk} = \mu + t_i + b_j + \epsilon_{ijk} \qquad (1)$$

where μ - overall constant

t_i - i^{th} treatment effect, $i = 1, \ldots, v$

b_j - j^{th} block effect, $j = 1, \ldots, b$

y_{ijk} - k^{th} observation of treatment i in block j, $k = 1, \ldots, \ell_{ij}$

ϵ_{ijk} - experimental error in y_{ijk} assumed to be $N(0,\sigma^2)$

3. Incidence Matrix

The design of the experiment is characterized by the incidence matrix, N, which is a vxb matrix whose elements n_{ij} are one if treatment i is applied in block j, and are zero otherwise. For example, if v = 3 and b = 3, the incidence matrix might be:

Treatments	Blocks 1	2	3	
1	1	1	0	t_1 is applied in blocks 1 and 2
2	0	1	1	t_2 is applied in blocks 2 and 3
3	1	0	1	t_3 is applied in blocks 1 and 3

Let ℓ_{ij} be the number of observations on treatment i in block j. Then

$$L = ((\ell_{ij})) = ((n_{ij}\ \ell_{ij}))$$

*This is a condensed version of a paper which is to appear in a national journal. This article has been reproduced photographically from the author's manuscript.

Preceding page blank

4. Definition of Connected Designs and Sets of Connected Blocks

A connected design is one in which all the blocks are connected by a chain of treatments. In the example, p. 1, blocks 1 and 2 are connected by treatment 1 and blocks 2 and 3 are connected by treatment 2 - hence all blocks are connected and the design is said to be a connected design. The number of sets of connected blocks is one, and the set consists of $\{b_1, b_2, b_3\}$.

In the example below, blocks 1 and 2 are connected by treatment 1, but there is no treatment which connects either block 1 or block 2 to block 3 - hence the design is said to be non-connected. There are two sets of connected blocks $\{b_1, b_2\}$ and $\{b_3\}$.

$$N = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

5. Review of Solution for Connected Designs with No Missing Treatments

The least-squares solution obtained by minimizing

$$S = \sum_{i=1}^{v} \sum_{j=1}^{b} \sum_{k=1}^{\ell_{ij}} n_{ij} (y_{ijk} - \mu - t_i - b_j)^2$$

with respect to μ, t_i, and b_j, and eliminating μ and b_j, gives the reduced normal equations: $C \hat{t} = Q$, where

$\hat{t}' = (\hat{t}_1, \hat{t}_2, \ldots \hat{t}_v)$, the vector of treatment effects

$$C = R - L K^{-1} L' \qquad (2)$$

$$Q = T - L K^{-1} B$$

and: R is a diagonal matrix with diagonal elements equal to the number of observations on t_i, i.e. $r_i = \sum_{j=1}^{b} \ell_{ij}$, $i = 1, \ldots, v$

All $r_i > 0$ in the no missing treatments case

K is a diagonal matrix with diagonal elements equal to the number of observations in block j, i.e. $k_j = \sum_{i=1}^{v} \ell_{ij} > 0$, $j = 1, \ldots, b$

L is the matrix: $L = ((\ell_{ij} n_{ij}))$

$T' = (T_1, T_2, \ldots, T_v)$ is a vector of treatment totals, i.e.

$$T_i = \sum_{j=1}^{b} \sum_{k=1}^{\ell_{ij}} n_{ij} y_{ijk}, \quad i = 1, \ldots, v$$

$B' = (B_1, B_2, \ldots, B_b)$ is a vector of block totals, i.e.

$$B_j = \sum_{i=1}^{v} \sum_{k=1}^{\ell_{ij}} n_{ij} y_{ijk}, \quad j = 1, \ldots, b$$

The solution $\hat{t} = C^+ Q$, where C^+ is the generalized inverse of C, gives the minimum-variance, unbiased estimates of the t_i, i = 1, ..., v, subject to the constraint $\sum_{i=1}^{v} t_i = 0$. The rank of the matrix C gives the number of linearly independent treatment effects which can be estimated, hence the degrees of freedom associated with treatments. Thus for the connected design, no missing treatments case, we have the following:

Treatment effects $\hat{t} = C^+ Q$

Variance-covariance matrix $V(\hat{t}) = C^+ \sigma^2$

Sums of squares due to treatments $SS(t) = \hat{t}' Q = \hat{t}' C \hat{t}$

Degrees of freedom d.f. (t) = Rank C = v - 1

Let: w be the total number of observations in the experiment,

Y, the vector of observations, $((n_{ij} y_{ijk}))_{w \times 1}$

$G = \sum_{i=1}^{v} \sum_{j=1}^{b} \sum_{k=1}^{\ell_{ij}} n_{ij} y_{ijk}$, the grand total of the observations.

Then the analysis of variance table is given by:

Source of Variation	Sums of Squares	Degrees of Freedom
Treatments (adjusted for blocks)	$\hat{t}' Q$	Rank C = v - 1
Blocks (unadjusted)	$B' K^{-1} B - G^2/w$	b - 1
Error	$Y' Y - \hat{t}' Q - B' K^{-1} B$	w - Rank C - b
Total	$Y' Y - G^2/w$	w - 1

(3)

For the factorial case, the notation developed by Kurkjian and Zelen in "A Calculus for Factorial Arrangements", Annals of Math Stat, Vol. 33, No. 2, June 1962, will be used. Two operators, ⊛, symbolic direct product (SDP), and x, direct product (DP) are needed.

The SDP is used to order the combinations of levels of the various factors, illustrated by the following example: Assume three factors in the experiments, two at two levels and one at three levels. Let $\theta_s' = (1, 2, \ldots, m_s)$ be a vector designating the levels of the s^{th} factor, which has m_s levels. Then:

$$\theta_1 \otimes \theta_2 \otimes \theta_3 = \begin{bmatrix}1\\2\end{bmatrix} \otimes \begin{bmatrix}1\\2\end{bmatrix} \otimes \begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}111\\112\\113\\121\\122\\123\\211\\212\\213\\221\\222\\223\end{bmatrix}$$

The final vector gives a particular order of the combinations of the levels of the three factors, and is obtained by setting the first two factors at level 1 and running through all levels of the third factor; setting the second factor at level 2, running through all levels of the third factor again; and finally setting the first factor at its second level and repeating the sequence on the second and third factors again. The procedure can easily be generalized to any number of factors at various levels.

The DP is the matrix multiplication defined as follows:

$$A_{m \times n} \times B_{p \times q} = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1}B & a_{m2}B & & a_{mn}B \end{bmatrix}_{mp \times nq}$$

That is, each element of A is multiplied times the entire matrix B, by ordinary multiplication of a scalar times a matrix.

Now let $A_1, A_2, \ldots, A_n$ be n factors in the experiment

at $m_1, m_2, \ldots, m_n$ levels, respectively.

The number of treatments (or treatment-combinations) resulting from applying all the factors simultaneously at all combinations of their levels is $v = \prod_{i=1}^{n} m_i$.

Let $(i_1, i_2, \ldots, i_n)$ be the i^{th} treatment-combination where factor A_1 is at level i_1, factor A_2 is at level i_2, etc., and order the treatments by the SDP of the levels of the factors.

For example, if $n = 2$, $m_1 = 2$, $m_2 = 3$, then:

$$\theta_1 \oplus \theta_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \oplus \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 11 \\ 12 \\ 13 \\ 21 \\ 22 \\ 23 \end{bmatrix}; \quad \text{Thus} \quad t = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \end{bmatrix} = \begin{bmatrix} t_{11} \\ t_{12} \\ t_{13} \\ t_{21} \\ t_{22} \\ t_{23} \end{bmatrix}$$

Thus treatment 1 is the combination with both factors at level 1, treatment 2 is the combination with the first factor at level 1 and the second factor at level 2, etc.

Let $a_s(i_s)$ be the main effect of factor A_s at the i_s level,

$a_{rs}(i_s, i_r)$ be the second-order interaction effect between factors A_r and A_s at levels i_r and i_s, respectively,

$\vdots$

$a_{12 \ldots n}(i_1, i_2, \ldots, i_n)$ be the n^{th}-order interaction effect between the n factors at levels $i_1, i_2, \ldots, i_n$, respectively.

Then the i^{th} treatment expressed in terms of the main and interaction effects of the factors is:

$$t_i = t(i_1, i_2, \ldots, i_n) =$$

$$\sum_{s=1}^{n} a_s(i_s) + \sum_{1 \le s < r \le n}^{s} \sum_{r} a_{rs}(i_r, i_s) + \ldots + a_{1,2 \ldots n}(i_1, i_2, \ldots i_n) \qquad (4)$$

For the previous example, we get the relationships:

$$\begin{aligned} t_1 &= t_{11} = a_1(1) + a_2(1) + a_{12}(11) \\ t_2 &= t_{12} = a_1(1) + a_2(2) + a_{12}(12) \\ &\vdots \qquad\quad \vdots \qquad\qquad \vdots \\ t_6 &= t_{23} = a_1(2) + a_2(3) + a_{12}(23) \end{aligned} \qquad (5)$$

Let a_X represent a general interaction term vector of effects where $X = (x_1, x_2, \ldots, x_n)$ such that

$x_i = 1$ if factor A_i is present in the interaction term and

$x_i = 0$ if not.

The number of elements in the vector is $\prod_{i=1}^{n} m_i^{x_i}$, again the elements are assumed to be in the order defined by the SDP of the levels of the factors involved, and the order of the interaction is given by $p = \sum_{i=1}^{n} x_i$.

The number of interaction terms for an n-factor experiment is $2^n - 1$, which is the number of combinations of zeroes and ones in the vector X, excluding all zeroes.

Continuing with the example, there are $2^n - 1 = 3$ interaction terms as follows:

$X = (10)$ denotes the main effect (first order) term for A_1 and

$$a_X = a_1 = \begin{bmatrix} a_1(1) \\ a_1(2) \end{bmatrix}$$

$X = (01)$ denotes the main effect term for A_2 and

$$a_X = a_2 = \begin{bmatrix} a_2(1) \\ a_2(2) \\ a_2(3) \end{bmatrix}$$

$X = (11)$ denotes the second-order interaction term A_1A_2 and

$$a_X = a_{12} = \begin{bmatrix} a_{12}(11) \\ a_{12}(12) \\ a_{12}(13) \\ a_{12}(21) \\ a_{12}(22) \\ a_{12}(23) \end{bmatrix}$$

Adding the constraints that the sum of the effects in an interaction term over all levels of any one factor is zero, i.e.

$$\sum_{i_s=1}^{m_s} a_s(i_s) = 0 \qquad s = 1, \ldots, n \tag{6}$$

$$\sum_{i_r=1}^{m_r} a_{rs}(i_r, i_s) = \sum_{i_s=1}^{m_s} a_{rs}(i_r, i_s) = 0 \qquad \begin{matrix} r = 1, \ldots, n \\ s = 1, \ldots, n \end{matrix} \quad r \neq s$$

etc.,

the relationships in (4) can be solved uniquely for the interaction effects in terms of the treatment effects, giving:

$$\hat{a}_X = \frac{1}{v} M_X \hat{t} \qquad \text{where} \tag{7}$$

$$M_X = M_1^{x_1} \times M_2^{x_2} \times \ldots M_n^{x_n} \quad \text{(Using DP multiplication)}$$

$$\text{and } M_i^{x_i} = M_i = m_i I_i - J_i, \text{ if } x_i = 1$$

$$= 1_i' = (1, 1, \ldots, 1)_{1 \times m_i}, \text{ if } x_i = 0$$

where I_i is the identity matrix of order m_i, and

J_i is a matrix of all ones of order m_i.

Thus for the example, the constraints are:

$$\sum_{i=1}^{2} a_1(i) = \sum_{i=1}^{3} a_2(i) = \sum_{i=1}^{2} a_{12}(i1) = \sum_{i=1}^{2} a_{12}(i2) = \sum_{i=1}^{2} a_{12}(i3)$$

$$= \sum_{i=1}^{3} a_{12}(1i) = \sum_{i=1}^{3} a_{12}(2i) = 0$$

And the interaction effects expressed in terms of the treatment effects are given by:

$$X = (10): \hat{a}_X = \begin{bmatrix} \hat{a}_1(1) \\ \hat{a}_1(2) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 & 1 & 1 \end{bmatrix} \hat{t}$$

$$X = (01): \hat{a}_X = \begin{bmatrix} \hat{a}_2(1) \\ \hat{a}_2(2) \\ \hat{a}_2(3) \end{bmatrix} = \begin{bmatrix} 2 & -1 & -1 & 2 & -1 & -1 \\ -1 & 2 & -1 & -1 & 2 & -1 \\ -1 & -1 & 2 & -1 & -1 & 2 \end{bmatrix} \hat{t}$$

$$X = (11): \hat{a}_X = \begin{bmatrix} \hat{a}_{12}(11) \\ \hat{a}_{12}(12) \\ \hat{a}_{12}(13) \\ \hat{a}_{12}(21) \\ \hat{a}_{12}(22) \\ \hat{a}_{12}(23) \end{bmatrix} = \begin{bmatrix} 2 & -1 & -1 & -2 & 1 & 1 \\ -1 & 2 & -1 & 1 & -2 & 1 \\ -1 & -1 & 2 & 1 & 1 & -2 \\ -2 & 1 & 1 & 2 & -1 & -1 \\ 1 & -2 & 1 & -1 & 2 & -1 \\ 1 & 1 & -2 & -1 & -1 & 2 \end{bmatrix} \hat{t}$$

Having the interaction effects expressed in terms of the treatment effects, then the following equations hold for the connected block, no missing treatments case:

$$\hat{a}_X = \frac{1}{v} M_X \hat{t} = \frac{1}{v} M_X C^+ Q$$

$$\text{Var}(\hat{a}_X) = \frac{\sigma^2}{v^2} M_X C^+ M_X' = \sigma^2 \sum_X$$

$$\text{Cov}(\hat{a}_{X_i}, \hat{a}_{X_j}) = \frac{\sigma^2}{v^2} M_{X_i} C^+ M_{X_j}', \quad i \neq j \tag{8}$$

Sums of Squares due to $a_X = SS(a_X) = \hat{a}_X' \sum_X^+ \hat{a}_X$

$SS(a_X)/\sigma^2$ is chi-square distributed with

$f_X = \text{Rank}(\sum_X) = \prod_{i=1}^{n} (m_i - 1)^{x_i}$ degrees of freedom.

If the design is connected, and orthogonal (i.e. $\text{cov}(\hat{a}_{X_i}, \hat{a}_{X_j}) = 0$ for all i,j, $i \neq j$)

then: $\sum_{i=1}^{\prod_{i=1}^{n}(m_i - 1)} SS(a_{X_i}) = \hat{t}' Q = SS(\text{treatments})$.

If the design is not orthogonal, the $SS(a_X)$ are not additive, but each $SS(a_X)/\sigma^2$ is statistically independent of the error term, so that F-tests are valid.

The analysis of variance table is given by:

Source of Variation	Sums of Squares	Degrees of Freedom
X_1	$SS(a_{X_1})$	Rank Γ_{X_1}
X_2 (Adjusted for Blocks)	$SS(a_{X_2})$	Rank Γ_{X_2}
$\vdots$	$\vdots$	$\vdots$
X_{2^n-1}	$SS(a_{X_{2^n-1}})$	Rank $\Gamma_{X_{2^n-1}}$
Blocks (Unadjusted)	$B' K^{-1} B - G^2/w$	$b - 1$
Error	$Y' Y - \hat{t}' Q - B' K^{-1} B$	$w - \text{Rank } C - b$
Total	$Y' Y - G^2/w$	$w - 1$

(9)

From a computational viewpoint, the calculations involved in (8) can be greatly reduced by eliminating the elements of each a_X which are linearly dependent because of the constraints in (6). The total number of elements in all the a_X vectors is $\prod_{i=1}^{n} (m_i + 1) - 1$, while only (v-1) are linearly independent in the connected design case. If all elements involving any factor at its highest level (each of which can always be expressed in terms of other elements in that term using the relationships in (6)) are eliminated, there will result (v-1) independent elements.

Then let $\bar{a}_X$ be a vector containing the linearly independent elements of $\hat{a}_X$ (selected as above) and let $\bar{M}_X$ be the corresponding rows of M_X. Then the equations corresponding to those in (8) become:

$$\begin{aligned}
&\bar{\hat{a}}_X = \frac{1}{v} \bar{M}_X \hat{t} = \frac{1}{v} \bar{M}_X C^+ Q \\
&\text{Var } (\bar{\hat{a}}_X) = \frac{\sigma^2}{v^2} \bar{M}_X C^+ \bar{M}_X' = \sigma^2 \bar{\Gamma}_X \\
&\text{Cov } (\bar{\hat{a}}_{X_i}, \bar{\hat{a}}_{X_j}) = \frac{\sigma^2}{v^2} \bar{M}_{X_i} C^+ \bar{M}_{X_j}' \quad i \neq j \\
&SS (\bar{a}_X) = SS(a_X) = \bar{\hat{a}}_X' \bar{\Gamma}_X^{-1} \bar{\hat{a}}_X \\
&\bar{f}_X = f_X = \text{Rank } \bar{\Gamma}_X = \prod_{i=1}^{n} (m_i - 1)^{x_i}
\end{aligned} \tag{10}$$

The dimensions of the var and cov matrices are reduced from $\prod_{i=1}^{n} m_i \; x_1$ to $\prod_{i=1}^{n}(m_i-1)^{x_1}$, and the matrix $\overline{\sum}_X$ is non-singular, so SS can be computed using the regular inverse.

The analysis of variance table (9) remains the same, since the sums of squares and degrees of freedom are equal.

6. Non-connected Designs and Missing Treatment Solutions

If the design is not connected, then additional constraints are needed to find a unique solution to the reduced normal equations $C\hat{t} = Q$.

Let z_1 be the number of sets of connected blocks and let S_i be the i^{th} set, $i = 1, 2, \ldots, z_1$. Let z_2 be the number of missing treatments (i.e. the number of $r_i = 0$). Then there must be $z_1 + z_2$ constraints to find a unique solution to the reduced normal equations.

If the constraints are taken to be:

$$\sum_{s_j} t_i = 0 \qquad j = 1, 2, \ldots, z_1, \text{ and}$$

$$t_i = 0 \qquad \text{for each } i \text{ such that } r_i = 0,$$

that is, if the sum of the treatments associated with each set of connected blocks is zero, and each treatment that is missing is assumed to be zero, then the solution $\hat{t} = C^{*} Q$ with C and Q as previously defined, satisfies the constraints. The analysis of variance table (3) remains the same except that the degrees of freedom for treatments, Rank $C = v - z_1 - z_2$.

Now, in the factorial case, the problem is to find the relationships resulting from the additional constraints on the t_i's, select a set of $(v - z_1 - z_2)$ independent a_X's, and compute the corresponding sums of squares and degrees of freedom for the analysis of variance table.

To determine the relationships on the a's in addition to those in (6) let

$$\overline{\hat{a}} = \begin{bmatrix} \overline{\hat{a}}_{X_1} \\ \overline{\hat{a}}_{X_2} \\ \vdots \\ \overline{\hat{a}}_{X_2 n-1} \end{bmatrix} \quad v-1 \times 1 \qquad \overline{M} = \begin{bmatrix} \overline{M}_{X_1} \\ \overline{M}_{X_2} \\ \vdots \\ \overline{M}_{X_2 n-1} \end{bmatrix}$$

That is, $\bar{\hat{a}}$ consists of all the elements in the $\hat{a}_\gamma$ vectors, as defined in (10) and $\bar{M}$ consists of the corresponding rows of each M_X. Then the system of equations to calculate the (v-1) $\bar{\hat{a}}$ elements is:

$$\bar{\hat{a}} = \frac{1}{v} \bar{M} \hat{t} = \frac{1}{v} \bar{M} C^+ Q \tag{11}$$

Linear relationships among the rows of $\bar{M} C^+$, and hence among the elements of $\bar{\hat{a}}$, can be determined by numerical techniques. If the rows of $\bar{M} C^+$ are arranged so that rows corresponding to elements of main effects are first, then those of second-order interaction terms, then third-order, etc. and a pivotal-method is used in which rows are interchanged only when necessary to remove a zero element from the diagonal, elements of terms of lowest order possible can be selected for the independent elements, and the remainder expressed in terms of those elements.

The linear relationships so determined can be used to categorize terms involving the dependent elements as being aliased with independent terms for which the coefficients are non-zero, or unestimable if all coefficients are zero.

Having so determined a set of linearly independent elements, reduce $\bar{\hat{a}}$ by eliminating the dependent elements, getting

$$\bar{\bar{\hat{a}}} = \begin{bmatrix} \bar{\hat{a}}_{X_1} \\ \vdots \\ \bar{\hat{a}}_{X_k} \end{bmatrix} = \frac{1}{v} \bar{\bar{M}} C^+ Q$$

where $\bar{\bar{M}}$ contains the rows of $\bar{M}$ corresponding to the elements in $\bar{\bar{\hat{a}}}$, and where the X_i, $i = 1, \ldots, k$, represent those interaction terms for which at least one element is among the final set of independent elements. Some terms may not appear at all (if all elements associated with that term have been eliminated), while others may have degrees of freedom less than $\prod_{i=1}^{n} (m_i - 1)^{x_i}$ (if only part of the elements have been eliminated). Then the following relationships hold;

$$\begin{aligned}
&\bar{\bar{\hat{a}}}_{X_i} = \frac{1}{v} \bar{\bar{M}} \hat{t} = \frac{1}{v} \bar{\bar{M}} C^+ Q \qquad i=1, \ldots, k \\
&\mathrm{Var}(\bar{\bar{\hat{a}}}_{X_i}) = \frac{\sigma^2}{v^2} \bar{\bar{M}}_{X_i} C^+ \bar{\bar{M}}'_{X_i} = \sigma^2 \bar{\bar{\Sigma}}_{X_i} \\
&\mathrm{Cov}(\bar{\bar{\hat{a}}}_{X_i}, \bar{\bar{\hat{a}}}_{X_j}) = \frac{\sigma^2}{v^2} \bar{\bar{M}}_{X_i} C^+ \bar{\bar{M}}'_{X_j} \qquad i \neq j \\
&SS(a_{X_i}) = \bar{\bar{\hat{a}}}'_{X_i} \bar{\bar{\Sigma}}^{-1}_{X_i} \bar{\bar{\hat{a}}} \\
&f_{X_i} = \mathrm{Rank} \ \bar{\bar{\Sigma}}_{X_i}
\end{aligned} \tag{12}$$

If the design is orthogonal, then $\sum_{i=1}^{k} SS(a_{X_i}) = \hat{t}'Q = SS(\text{treatments})$.

The analysis of variance table is given by:

Source of Variation	Sums of Squares	Degrees of Freedom
X_1	$SS(a_{X_1})$	Rank $\overline{\overline{\sum}}_{X_1}$
X_2 (adjusted for blocks)	$SS(a_{X_2})$	Rank $\overline{\overline{\sum}}_{X_2}$
$\vdots$	$\vdots$	$\vdots$
X_k	$SS(a_{X_k})$	Rank $\overline{\overline{\sum}}_{X_k}$
Blocks (unadjusted)	$B'\frac{1}{k}B - G^2/w$	b-1
Error	$Y'Y - \hat{t}'Q - B' K^{-1} B$	w - Rank C - b
Total	$Y'Y - G^2/w$	w - 1

(12)

DESIGN OF FIELD TEST PROGRAMS AND STATISTICAL TECHNIQUES FOR ANALYSIS OF THE PERFORMANCE OF NAVIGATION AND POSITIONING SYSTEMS*

Emil H. Jebe
University of Michigan

and

Ralph A. King
University of Wisconsin

GLOSSARY AND DEFINITIONS

1. Analysis of Variance - A statistical technique based on a linear model and the application of least squares for subdivision of the total variability in a sample into components specified by the model.

2. ACE(S) - Along course error of the system position.

3. b_{EO} - Slope of the orthogonal regression line describing the system's path based upon external position data.

4. CRD - Completely randomized design, the simplest type of experimental design or pattern of experimentation. The treatment combinations are randomly assigned to the entire set of experimental units.

5. CCE(S) - Cross-course error of the system position.

6. CCE(SP) - Cross-course error based upon the system's estimate of its own position.

7. CCE(E) - Cross-course error (external) = length of a perpendicular from point (X_E, Y_E) to the programmed path as determined by external measurements.

8. Correlation - A measure of linear association between two random variables: $\rho = \sigma_{xy}/\sigma_x \sigma_y$, i.e., a ratio of the covariance to the product of the standard deviations. Sample estimator, $r = s_{xy}/s_x s_y$.

9. Chi Square Distribution - The probability distribution of the square of a standard normal variable. Let z_i be N(0,1), then z_i^2 has Chi Square distribution with one degree of freedom.

---

*This paper also was presented by the senior author at the "Technical Symposium on Navigation and Positioning," 23-25 September 1969, USAECOM, Fort Monmouth, New Jersey.

10. $\Delta x = X_S - X_E$ = deviation of the system's indicated position (X coordinate from an external measure of system's location.

11. $\Delta y = Y_S - Y_E$ = deviation of the system's indicated position (Y coordinate) from an external measure of system's location.

12. d = radial error = $[(\Delta x)^2 + (\Delta y)^2]^{1/2}$ = straight line distance from system's indicated position to position determined by external measuring equipment. Note that if Δx and Δy are normal random variables, then $d^2\sigma^2$ is distributed as Chi Square with two degrees of freedom.

13. Duplicate = A subsample of an experimental unit; one of two measures of system performance for the same experimental unit.

14. Degrees of freedom = Formally, a parameter of the Chi Square probability distribution. In application, the number of independent deviations available for estimating a variance or mean square.

15. Experiment = Study of system performance over a set of experimental units.

16. Experimental error = A mean square or quadratic measure of system variability about its average performance measured over a set of homogeneous experimental units.

17. Experimental unit = A period or segment of system operation for which an independent measure of system performance can be obtained.

18. F ratio = A ratio of two independent estimates of variance for which under the "null hypothesis" both numerator and denominator are distributed as (Chi Square) (σ^2) with degrees of freedom, say f_1 and f_2.

19. Interaction - A situation in which the observed results for the simultaneous application of two or more factors cannot be explained by addition of the direct effects of each factor separately estimated.

20. Local Control - (Blocking) a subdivision of the total set of available experimental units into relatively homogeneous subsets. Each subset is called a block. A complete block contains one experimental unit for each treatment combination. Two or more such blocks then comprise a RCB.

21. Median - A value which divides a population or a sample into two equal parts.

22. Orthogonal regression line - A least squares fitted line such that the sum of squares of normal distances from points to the line is minimized.

23. Position error (refer text)

24. Quartile - The quartiles are values that divide a population or a sample into four equal parts. The second quartile is the median.

25. Radial error = d (see 12 above).

26. RCB - Randomized complete block design (see 20 above).

27. Response - A measure of system performance. May be univariate but is often multivariate.

28. Replication - In a simple measurement situation a single independent observation of system performance. Otherwise, one replicate comprises one observation of system performance for each treatment combination of the entire set of treatment combinations being investigated.

29. Regression mean square - The mean square of deviations of points from a regression line (based upon division by the degrees of freedom).

30. Sample size - The number of observations on each treatment combination or the number of complete replicates. Note: this is <u>not</u> the total number of experimental units.

31. Standard deviation - Square root of the variance; a measure of variability about the average of observations from a homogeneous set of experimental units.

32. Structure of a Test Program or Experiment - The overall arrangement of a test program which includes the treatment combinations to be investigated, the environments and locations in which the system is to be operated and the experimental design imposed.

33. Treatments (and Treatment Combination) - If a system is to be tested at altitudes, say Low and High, we say that altitude is a factor at two levels. We also refer to altitude as a treatment imposed on the system. Suppose we also wish to test the system over land and over water. Then water and low altitude and water and high altitude are two different treatment combinations. Two factors each at two levels provide a total of four treatment combinations.

Field Test programs are fraught with many difficulties. Developmental equipment just never seems to perform as well as desired by its producers or as hoped for by the Army. Characteristics of the field environment may not have been adequately anticipated by the development engineers. Often an extensive shakedown period is required before a system is really ready to be entered into a field test program.

Even before the shakedown trials are started a complete TEST PLAN must be developed for the field test program. The field test envisaged may comprise several parts with each part designed to exercise the system in a different way. When this is the case, a specific TEST PLAN should be developed for each part.

It has been our experience that field test programs are often inadequate or incomplete in several respects. Therefore, we need to consider the question, "What are the GENERAL FEATURES OF A TEST PROGRAM?" These features are set out as a list of ten items (prepared by the senior author at a time when he first came in contact with the study of navigation and positioning systems) [1].

GENERAL FEATURES OF A TEST PROGRAM

1. Careful delineation of the problem and thorough understanding of the system or systems to be examined.

2. Definition of the phenomena to be studied. (Including "What are the requirements?")

3. Selection of the response (i.e., performance characteristics) and the technique of measurement for each response. Know the standards that should be applied.

4. Determination of a suitable experimental unit.

5. Selection of treatments to be studied (i.e., equipment parameters to be varied.

6. Selection of environmental conditions or parameters to be varied.

7. Choice of a pattern of experimentation (suitable combinations of experimental units, treatments and environment). Result is an experimental plan or design that includes the randomization procedures, adequate local controls and sufficient replication.

8. Complete layout of the plan for analysis of the responses or measurements to be obtained (before the data are taken).

9. Interpretations to be made from all possible experimental results.

10. What is the next experiment that may be relevant after the currently proposed one is completed?

Let us consider these ten items in turn. Item 1, we leave to the engineers although much questioning is often required to obtain a clear statement of the problem. Item 2, we also regard largely as an engineering area. Spelling out what is expected of the system in realistic and useful terms is a major step. Later, the question is to be asked and answered, "Does the system fulfill the requirements?" To a considerable extent the answer will depend on the data acquired and our analysis of these data. Examples of requirements might be, "Take off from Dulles International for Paris; make landfall in France with cross-course deviation less than 5 miles with respect to a designated point." Or, "Take off from Field A; fly over point X, Y with an average radial error not to exceed 20 meters; land at Field B."

Succeeding items on the list lead us more into the statistical and experimental design problems. Determination and definition of the relevant response (Item 3) for judging the performance of the system is basic to all that follows. Yet many "test programs" have been written without having the performance measures for the system quantified and the methods of measurement clearly stated. Related to the performance measure is the selection of the standard for assessment of that performance. With respect to navigation and positioning systems we may ask,

(1) Do we need photo-theodolite data?, or

(2) Is a radar network required?, or

(3) Will the measurements from a single radar such as the FPS-16 be sufficient?, or

(4) Can we rely on a higher resolution non-radar electronic network?, or

(5) Will cruder methods, simple photographs or visual observation, be sufficient?

Depending on the stated requirements, we may select one or more of these alternatives.

After much though about item 4, we have reached the conclusion that for an electronic system mounted in a land vehicle, a ship or an A/C, the <u>entire mission</u> on a given day must be regarded as the <u>experimental unit</u>. In this mission we include -

starting up the system
warm-up
check out
calibrations
departure from base
calibrations enroute
traversing selected courses
return to base
checking calibrations
shutting down the system
complete return to ambient conditions.

This view of the experimental unit means that any repetition maneuvers performed by the system within the same mission must be regarded as duplicates and not as replicates. Of course, we are interested in the variation among duplicates but major interest centers on the replicates, that is, the repeated performance of the system over a set of experimental units that we regard as similar or sufficiently homogeneous for the problem under study. By definition, experimental error is the failure of a system to produce identical responses over a set of independent trials (or experimental units). The key word here is independent; we believe that repeated maneuvers in any one mission are likely to be highly correlated. Therefore, we insist that an independent trial for an electronic system include the complete sequence given above from "starting up the system" through "return to ambient conditions."

A system may have several "modes" of operation, threshold settings may be required and variation of dial settings may affect the performance of the system. All these equipment parameter variations we include under the set of treatments that may be investigated (Item 5). Further, we usually extend our concept of the treatments of interest to include the variations external to the system which may or may not affect (hopefully not) the performance of the system. Under Item 6, we include weather, altitude, day or night operation, electromagnetic disturbances (natural or man-made), terrain, direction over a course, etc.

The result of considering Items 4, 5, and 6 leads us to selection of a pattern or program for the system test. The structure of the test program is determined by the factors (conditions and parameter settings) which we wish to investigate. The simpler this structure can be made, the easier it will be to:

(1) Cope with the inevitable modifications of the test program that arise due to revision of test objectives, unexpected equipment limitations, or failure to obtain adequate data for some courses; and,

(2) Analyze the data.

The simplest form of experimental pattern is called a Completely Randomized Design (CRD). This design is preferred when it is feasible. A simple description is that we write down on slips of paper each combination of conditions and parameter settings that is to be included in the test program. Then we put the slips into a hat, mix thoroughly, draw them out one at a time and write out a complete list of the consecutive drawings. Suppose altitudes of one thousand and 12,000 feet were included in the test program for an airborne system. If any part of the consecutive sequence of drawings came out with altitudes (in thousands of feet) 12, 1, 12, 1, 12 for the sequence of courses to be flown the pilots would object; hence, we regard a CRD as not feasible for such a situation. Therefore, split-plot structures or nested designs must be worked out when some of the treatments cannot be submitted to complete randomizations.

Performance of systems tends to vary with time, or for a development item prototype the performance is even likely to deteriorate with time. Such results are to be expected when the "bugs" are not all ironed out, and the test program covers a 3 to 6 month period. Because of this time variability in performance, it is highly desirable to introduce a "blocking" with respect to time. Such blocking is a form of what is generally known in experimental design as "local control." This local control permits the removal of (or elimination) of time variation so that any two treatment combinations (choice of parameter settings) can be compared without time bias. What this means in practice is that if two particular combinations are run in, say, the second week of the test program, and if one or the other is scheduled again for the 7th and 13th weeks of the test program, then if both are run in the 7th and 13th week, then the time differences (if any) among the 2nd, 7th, and 13th weeks can be removed in making the desired comparison. The balancing of the experimental program against time or some other possible source of undesirable variability is accomplished by setting up a Randomized Complete Block design. We regard the use of local control by blocking as a necessary requirement in the study of complex systems used for navigation and position determination. Here, we have assumed one week as comprising a block.

It is to be noted that each block as just described forms one complete replicate for a set of treatment combinations. The time period included in the block can be any reasonably short period of homogeneous test conditions, say, one day, three days, or one week. Thus, the number of blocks completed determines the total number of replicates for this set of treatment combinations. The number of blocks completed then determines the sample size so the natural question is, "How many blocks do we need?"

Two considerations enter into the determination of the desired sample size. First is the requirement of obtaining a stable estimate of the experimental error. It is our experience that an estimate with 10 to 20 degrees of freedom may often be adequate for development test programs. Such an estimate can be obtained with as few as three blocks when eight

or more treatment combinations are to be investigated in each block. Larger blocks, however, may introduce other problems; e.g., lack of homogeneity of experimental units. The second consideration is the magnitude of real differences in system performance that may be associated with environmental and/or parametric changes for the system. Again from experience we have found that system developers and system users have limited information on the magnitudes of these differences. It can be shown from theory that for a specified probability a "large" sample is required to detect "small" differences, but that a "modest" sample may detect easily "large" differences. These vague words (large, small, modest) can be given numerical values only when we are able to insert in the available formulae actual values for (1) the standard deviation of our experiment the (experimental error)$^{1/2}$ previously described); and (2) the magnitude of the difference to be detected.

The discussion of Item 7 of the "General Features" has been rather lengthy, but we have tied together in this discussion the preceding Items 4, 5, and 6 with Item 7. In this discussion we have covered some aspects of the choice of experimental pattern and its associated randomization, local control by blocking on time, and the choice of sample size.

Item 8 follows quite easily if we have done our homework well in covering Items 3 through 7. Perhaps, we should note that it is easy only in principle. We recall a paragraph from our abstract as follows:

> When these 'GENERAL FEATURES' have been closely adhered to, then the work of summarization and analysis of data and the final interpretation of results becomes much simpler. An experimental design for the field test program has associated with it a mathematical model; the two together determine the analytical procedures. One of the most useful and severe disciplines to impose on the military personnel and the development contractor is to require that a set of tables be prepared before the field test is started. This set of tables should include the detailed format of the summary data on which the performance of the system is to be judged. Further, the parties should agree that the performance is to be judged on these criteria.

The last two Items, 9 and 10, are essentially self-explanatory. It is usually salutary to give them some consideration, however, before the first experiment is begun. As the test program proceeds, other considerations will appear or come to bear on the problem. Thoughts about 9 and 10 will then take new directions. Without the pre-first-experiment considerations well though out and written down, the new directions may turn out to be undesirable tangents. The "whole forest needs to be kept in view rather than the interesting trees that appear as we walk in the woods." A remark on the use of the term <u>experiment</u> may be added here. Physical scientists often think of an experiment as a <u>single trial</u> under carefully specified

conditions. From the analysis point of view, which must be taken by the statistician, an experiment comprises all replicates of a set of treatment combinations among which comparisons are to be made. The TEST PLAN (or test program) for a given system may consist of only one, or two or more experiments.*

We now turn to the consideration of the second area of our paper as indicated by our title. Analysis of the performance of a navigation and positioning system must describe this performance quantitatively in terms of precision and accuracy [5]. Various statistical techniques may be required to describe this performance. In order to give concreteness to this section of the paper we shall base our discussion upon the analysis of the performance of an airborne navigation system in which we were engaged several years ago [6].

The field test program for this system included a requirement that the system depart from a base, fly over a calibration check point, and then proceed to maneuver the A/C over a series of six parallel flight paths whose end points were defined by specified longitude and latitude coordinates (see Figure 1). In Figure 1, we show two series of six parallel lines, sets 1A and 1B. The set 1B was actually laid over the same ground area as set 1A. Each line of a set of six we refer to as a LEG, so that the total flight course comprised six LEG's. Starting with LEG 1 in series 1A as shown we refer to this pattern as a Zero Degree flight forward over the course (A0,F). Beginning with LEG 6 and reversing direction over each LEG is called (A0, R). Using series 1B in the direction shown starting with LEG 1 is designated as 90 degrees forward (B90, F). Similarly, reversing course beginning with LEG 6 is designated as (B90, R). Other designations are possible such as starting at other end of LEG 1 in each series, which gives (A180, F) and (B270, F).

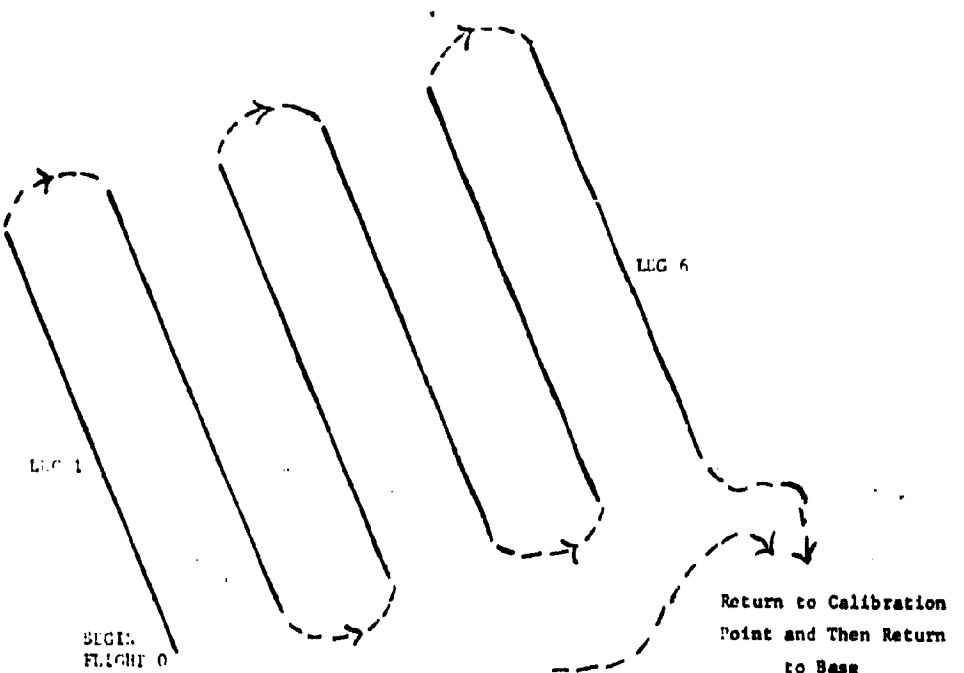
With this view of the flight area pattern we may approach the details of describing the system performance. Assessment of the system performance will be based largely on a position error; i.e., the difference in location of the system as determined by an external measuring system and the system's own indication of its location (at a given time). This position error information is to be analyzed by averaging and/or decomposition to provide descriptions of system performance. Among these descriptions are:

(1) The difference between the average location of the system over a number of repetitions under essentially similar conditions for a programmed flight over a point or a course and the desired point or course is a measure of system accuracy [7]. This accuracy, however, may vary over the flight area (1A & 1B) for a variety of reasons. Thus, it may be useful to speak of the system's predictability or reproducibility for a group of points or LEG's in the assessment of accuracy,

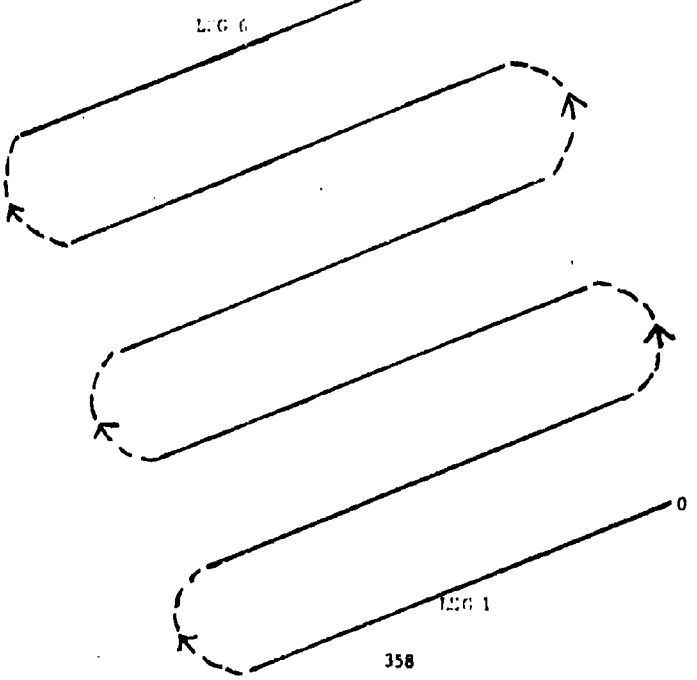
*Appropriate references for this first section of the paper are [2], [3], and [4].

FIGURE 1

1A



1B



(2) For precision or repeatability assessment we may describe the performance along a specific LEG or segment of a course which was programmed for the system, or

(3) We may describe the repeatability of the system in flying over the same programmed course a number of times, each time appearing on a different day.

Thus, it is seen from (2) and (3) that we can describe precision over experimental units (replicates), which is of greatest interest, and also in terms of within replicates over a segment of a LEG, a whole LEG or the set of LEG's. Replications of any LEG or part of a LEG on the same day may be regarded as duplicates from the viewpoint of sampling the system performance. We note that within a single programmed flight on the same day all individual position determinations made by the system must be regarded as inherently correlated to some greater or lesser but unknown degree. This point of view is conceptually correct in regarding the output of a single programmed flight as one "realization" in the sense of the theory of stochastic processes. The degree of correlation, of course, depends on the time and/or distance separation between any two position determinations. The actual magnitude or form and shape of this correlation function may be quite relevant for system design but need not be of major concern for evaluation of system performance. The realization of its presence, however, requires the definition of a single trial or experimental unit in the way already described and then it guides our analysis.

The discussion thus far has been general in the evaluation of system performance. It will be helpful to list some of the actual variables measured in relation to the determination of position error. These random variables were:

(1) $\Delta x = X_S - X_E$

(2) $\Delta y = Y_S - Y_E$

(3) CCE(S) = Cross-course error for the system

(4) ACE(S) = Along course error for the system

(5) $d = [(\Delta x)^2 + (\Delta y)^2]^{1/2}$ = Radial error for the system.

A rectangular grid system was laid out over the area indicated in Figure 1 with the point (0,0) arbitrarily selected. At time t_i, (S_{si}, Y_{Si}) was the system's indicated position while (X_{Ei}, Y_{Ei}) was the actual position of the system as determined by an external means. Thus, d_i was the radial error at time t_i. The time interval from t_i to $t_i + 1$ was five seconds.

The assessment of repeatability is most easily begun by examining the performance within a single LEG. For each LEG a number of summary statistics were computed for the variables just listed. These statistics included:

(1) Average value for the variable.

(2) Mean square deviation of the individual values from the average. Note that this quantity although calculated like a sample variance does not have the usual Chi Square distribution with n - 1 degrees of freedom because of the correlation of data points within a given LEG (as already discussed).

(3) Minimum value.

(4) Maximum value.

(5) Median value.

(6) First and third quartiles.

In this paper we can illustrate only a few analyses of these many statistics. A mere tabular summary, of course, gives some description of repeatability. A further analysis considers the behavior of these LEG statistics from LEG to LEG, from (programmed) flight to flight, at different altitudes, orientations (or direction of flight), and even over different areas. The statistical technique used for this further analysis is known as the analysis of variance. This technique has been well described by Kempthorne and Scheffe in its application to the analysis of experimental data [8, 9]. Briefly, the technique may be described as a procedure for evaluating the variation of averages and the variation of individual observations. These evaluations, called mean squares, may be compared by forming Snedecor's F ratio in order to make inferences about the magnitude of the variations of the averages. Specific assumptions, of course, are made in the application of the technique. Currently, most attention is given to these assumptions: (1) the specified linear model adequately represents the experimental structure; and, (2) independence, i.e., the data comprise a random sample from the universe of interest.

A simplified example will illustrate the application of the analysis of variance to a possible set of data from the flight program described above. Let us suppose a series of flights made over the area of Figure 1 with variations in altitude and heading. The series of flights is carried out in a completely randomized design with the results obtained as in Table 1. There are two replicates of each combination of conditions. Note that only average results for each entire flight are presented.

The analysis of variance appears in Table 2.

TABLE 1

ILLUSTRATIVE EXAMPLE OF POSSIBLE RESULTS
FOR EIGHT FLIGHTS FOR COMBINATIONS OF
TWO ALTITUDES AND TWO HEADINGS

Flight No.	Altitude	Heading (Degrees)	Average Radial Error (Meters)
1	7,000	0	80
2	7,000	90	90
3	15,000	0	50
4	7,000	0	90
5	15,000	90	40
6	7,000	90	100
7	15,000	0	50
8	15,000	90	60
			560

TABLE 2

Analysis of Variance of Average Radial Error

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square
Total	8	42800	----
Average	1	39200	----
Altitude	1	3200	3200
Heading	1	50	50
Altitude by Heading Interaction	1	50	50
Remainder	4	300	75

The model upon which this analysis is based is written as -

$$Y_{ijk} = \mu + A_i + H_j + (AH)_{ij} + \varepsilon_{ijk}$$

where Y_{ijk} is an average radial error as shown in Table 1 and the terms on the right in order are -

a general mean,
an altitude effect,
a heading effect,
an altitude-heading interaction, and
a random component associated with the ijk th experimental unit (flight).

It is the variation among these eight averages given in Table 1 which is to be subdivided into parts associated with the sources of variation present. We note that the altitude means are: 1/4 (80+90+90+100) = 90 at 7,000', and 1/4 (50+40+50+60) = 50 at 15,000'. Similarly, the Heading means are: 67.5 at 0 degrees and 72.5 at 90 degrees.

Thus, the 2 x 2 table of means for average radial error is

	Altitudes		
Headings	7	15	Averages
0	85	50	67.5
90	95	50	72.5
Averages	90	50	70

Details of the calculations, the assumptions underlying the analysis and interpretation of the results are given in most modern texts on statistical theory or techniques [10, 11]. We cannot consider these matters further here, but we point out two aspects of this hypothetical example: (1) The "Remainder" with 4 degrees of freedom is an appropriate estimate of experimental error, so that $(75)^{1/2} = 8.66$, is a standard deviation that estimates the repeatability of the System over repeated flights; and, (2) that the Mean Square for Altitude, 3200, when compared with the Remainder Mean Square provides a basis for assessing the effect of Altitude. If Altitude variation did not affect the System, we would expect these Mean Squares to be about equal. From Table 2 we would conclude by looking at the interaction component (F ratio = 50/75<1) that the Altitude effect does not vary with Heading. The Heading effect

appears negligible (F ratio = 50/75 1). Finally, we would conclude that performance differs with altitude (F ratio = 3200/75 = 43) ($P < <0.01$). From the averages, we see that the radial error is much smaller at the higher altitude.

In reference to the description given above for the analysis of variance as a technique for studying the variation of averages in contrast to the variation of individual observations, there is a point to be noted in relation to the hypothetical example just given. In the example, the individual values analyzed are themselves averages. Thus, there is a further component of variation associated with individual observations or points within LEGS, that has been suppressed in the example. Generally, in analyzing data for studying the system we followed this same procedure of studying averages. Thus, a simple LEG average provided a single datum and we analyzed the variation of these averages in relation to other factors.

There are several reasons for following this procedure. First, this approach, of course, has simplified some problems in analysis due to unequal numbers of observations within LEGS. Second, even though numbers of observations on a given LEG varied from as low as 80 to around 200, there was no reason for giving more weight to one flight over a given LEG than another if a reasonable set of data were obtained to represent that flight over that LEG. Thus, using averages and giving each average equal weight seemed a proper procedure for assessing the overall performance. Third, the use of averages, even though each average is computed from data with considerable correlation, will provide values of a random variable which more closely approach the assumptions of the analysis of variance technique. In analyzing the repeatability within LEGS as measured by the variances of designated random variables (Mean Square Deviations from Average or from an Orthogonal Regression Line), these variances may also be considered as "averages." Because of the greater apparent dispersion of these variances, it seemed desirable to analyze the natural logarithms of these quantities to obtain a transformed variable more suitable for the analysis of variance technique. Fourth, and last, this approach in terms of further analysis of original statistics (averages, variances, slopes of regression lines and deviations from such lines) is in keeping with the spirit of Professor John W. Tukey's suggestions [12].

The preceding example was made small in order to be easy to follow. The conclusions stated relate only to the hypothetical data of Table 1 as if they were real data. We now present in Table 3 some real data for six flights over the area represented by Figure 1. These flights were flown at three altitudes with zero degree heading (i.e., AO, F as noted above). Table 3 gives averages of Δx, Δy and d = radial error for each LEG of each flight. Hence, 36 averages are shown for each variable. The

Table 3

Tabulation of Leg Means for Selected Variables
From Six Flights Over the Figure 1 Area
(Units are Meters)

Flight Number	(Altitude/1000)	Leg Number 1	2	3	4	5	6	
		Variable: $\Delta x = X_S - X_H$						
5	7.5	143	32	54	109	120	116	
6	7.5	103	32	51	91	108	115	
1	11	30	-105	-9	51	71	86	
7	11	95	70	74	101	124	139	
3	15	121	81	107	88	125	116	
4	15	91	89	59	119	97	140	
		Variable: $\Delta y = Y_S - Y_H$						
5	7.5	-70	-155	-335	-341	-144	+11	
6	7.5	-75	-101	-233	-257	-80	+45	
1	11	+35	-339	-75	-130	+18	+185	
7	11	-92	-230	-341	-373	-172	+3	
3	15	-109	-159	-345	-324	-95	+3	
4	15	-103	-275	-331	-391	-122	-3	
		Variable: Radial Error $= [(\Delta x)^2 + (\Delta y)^2]^{1/2}$						
5	7.5	319	187	349	370	219	159	267*
6	7.5	277	139	233	288	195	142	216
1	11	278	378	176	169	168	212	230
7	11	262	253	356	402	231	175	280
3	15	264	186	370	344	208	138	252
4	15	260	298	357	412	209	196	287
		277**	240	310	331	204	170	
								255***

*Flight Averages;
**Leg Averages;
***Overall Average.

Table 4

Analysis of Variance of the Leg Averages
for the Variable: Radial Error

Source of Varation	Degrees of Freedom	Sum of Squares	Mean Square
Total	36	2582140	
Average	1	2346156	
Altitudes	2	4778	2389
Flights at Same Altitude	3	19099	6374
(Pooled Variation for Flights)	(5)	(23870)	(4774)
Legs Over All Flights	5	115968	23194
Legs x Altutudes Interactions	10	40300	4030
Legs x Flights at Same Altitudes	15	55847	3723
(Grouped Legs x Flights)	(25)	(16147)	(3846)

analysis of variance for one of the variables, radial error, is given in Table 4. Therefore, Table 3 also shows the marginal averages for this variable, that is, over all LEG's of the same flight, and over all flights for the same LEG.

The model for the analysis of Table 4 is written as -

$$Y_{ijk} = \mu + A_i + \epsilon_{ij} + L_k + (AL)_{ik} + \delta_{ijk}$$

where Y_{ijk} is average radial error as given in Table 3 and the terms on the right are -

a general mean,
an altitude effect,
an error component for flights at same altitude,
a LEG effect,
an ALTITUDE x LEG interaction, and
a residual which measures failure to obtain same results for a LEG when a repeated flight is made at the same Altutude.

Major interest in Table 4 first centers on the Altitude Comparison. The mean square for Altitudes is 2389 while the mean square for repeated flights made at the same altitude is 6364. The latter is our measure of experimental error for Altitudes; hence, the F ratio is $2389/6364 < 1$. We conclude that altitude variation did not affect the performance of the system over the range of altitudes selected (our choice of altitudes was limited by the performance capability of the A/C carrying the navigation system).

Next, we examine the variation within flights or between LEGs. The "LEGs over all flights" mean square is 23194, a large value relative to all other mean squares in Table 4. Thus, we are inclined to conclude that there are large differences among the six LEGs of the programmed flight pattern. The remaining two mean squares, LEGs x Altitudes = 4030 and LEGs x Flights at same Altitude = 3723, indicate the consistency of these large LEG differences. Pooling of the last two sources of variation yields a mean square of 3846 with 25 degrees of freedom. An approximate F ratio for comparing LEGs could be formed by $F = 23194/3846 \doteq 6$. We regard this ratio as an approximate F in distribution because of the correlation of LEGs within the same flight although this may be small because of the apparently large LEG differences. Perhaps, a multivariate test could be devised for comparing LEGs; we have not considered this approach. In view of the consistency of these LEG differences over different days throughout the test program it seemed reasonable to us to conclude that natural electromagnetic field variations over the six LEGs affected the system performance.

We shall not endeavor in this paper to summarize our conclusions about the repeatability or precision of the system which produced the data given in Table 3. It is our purpose here merely to present statistical methods and techniques for securing such information about any navigation system. It will be useful to give one more table, however, to show another aspect of the repeatability. In Table 5, we give the mean square deviation of the radial errors from the average radial error (given in Table 3) arranged by Flights and LEGs as in Table 3. We shall not give the analysis of variance for the data in Table 5 but we note that natural logarithms of these mean square deviations were taken before computing the analysis of variance. This log transformation is usually applied before analyzing variances of observations.

Although we have given only a small sample of the large amount of repeatability information obtained for the system we have been using for our discussion, we turn now to the system accuracy. If the system exhibits accurate performance we may say that it has predictability or reproducibility. In addition to the variables listed above, which, were used for examining the system repeatability, we also obtained the cross course error of the system location from the external measurements, CCE(E), which was the distance of the point (X_E, Y_E) from the programmed path (lines shown in Figure 1). An average of these values would show the bias or systematic error of the system in flying the programmed course. If this bias were negligible over all flights we would regard the system as accurate or that its performance is reproducible.

Further, from the (X_{Ei}, Y_{Ei}) data we obtained a derived quantity, the slope of the orthogonal regression line, b_{EO}, through the points traversed by the system. The slope of this line for each LEG of the programmed path then could be compared with the actual slope, β, of the programmed path in terms of the arbitrary X, Y coordinate system imposed on the area of Figure 1. Departures of the observed slopes, b_{EO} from the desired slope, β, then give further information on the system predictability.

Before presenting actual results it will be helpful to discuss briefly the use of the orthogonal regression line. From the above, it is clear that both X_E and Y_E are random variables. The usual regression models consider X_E to be an independent variable observed or measured with no error or negligible error. Natrella in "Experimental Statistics" gives a good discussion of regression analysis for functional and statistical relations [13]. It is clear that none of the standard models apply to this navigation system analysis. After considerable study, we concluded that fitting the orthogonal regression line would best describe the system performance. Derivation of the normal equations for fitting this line was given by Coleman in 1932 [14]. To our

TABLE 5

Mean Square Deviations by LEGs for Six Flights
Variable: Radial Error

Flight Number	Altitude	LEG 1	2	3	4	5	6
5	7.5	53240	15630	12740	5215	12720	6883
6	7.5	41050	7854	26780	5716	7835	1017
1	11	22700	155100	16280	4622	3332	6394
7	11	39490	27780	33510	6991	12890	4270
3	15	36400	23090	26200	9361	5372	1227
4	15	34170	28880	25420	3196	7804	8589

TABLE 6

Tabulation of Altitude Averages and General Average for Variables Which Describe System Predictability (based on LEG AVERAGES weighted equally)

(units are meters)

Variable	General Average*	Altitudes/(1000)(feet)** 7.5	11	15
		AVERAGES		
CCF(F)	-20.2	-7.4	-22.5	-30.8
CCF(SP)	5.4	10.0	7.0	2.2
Slope, b_{FO}***	-1.3001	-1.2993	-1.3030	-1.2999
Reg. Mean Square	5800	3500	7600	6200

* Based on average of 36 values; six flights of six LEGs.

** Based on average of 12 values; two flights of six LEGs.

*** Note: [illegible] = -1.3032

NOT REPRODUCIBLE

TABLE 7

Tabulation of Flight Averages for Variables Which Describe System Predictability

(based on Leg Averages weighted equally)

(units are meters)

Flights	Altitudes 7500 5	6	11,000 1	2	15,000 3	4	Standard Deviation**
Variables							
CCF-(F)	-13.9	-0.9	-9.4	-35.7	-10.4	-50.8	15.9
CCF(SP)	16.9	4.1	7.1	4.9	9.4	-5.9	6.9
Slope, b_{FO}	-1.2990	-1.3073	-1.3075	-1.2985	-1.3000	-1.2980	0.0038
Reg. M. Sq.	3900	3100	11500	3700	5500	6900	3100

* Based on average over six Legs within each flight.

** This standard deviation is derived from the flight mean square in the analysis of variance, i.e., $\sqrt{(\text{Flight Mean Square}/6)}$. The divisor six arises from the averaging over six legs within each flight.

knowledge, however, the sampling theory for the slope of this orthogonal regression line has never been presented (approximations could be obtained, perhaps). In a replicated experiment, this lack of adequate sampling theory does not create an impasse. Independent estimates obtained from repeated flights will permit direct estimation of the variability of the orthogonal regression slopes.

Along with computation of the orthogonal regression line we also present the regression mean square for deviations from the regression line. The magnitude of this mean square indicates the scatter of the X_E, Y_E points about the fitted line. The system whose data we have been presenting also provided an estimate of its own position which we designate as X_S, Y_S. From this data series we calculated a CCE(SP) = Cross Course Error of the system's indicated position. Table 6 presents average values of these four statistics for the six flights over the area of Figure 1.

As averages, these numbers in Table 6 speak for themselves. With respect to cross course error, if the system actually was on the left side of the programmed path, the deviation was designated as negative. Thus, we see that the system generally directed the flight slightly to the left of the programmed path. On the other hand, the system's indication of its position on the average was an even smaller deviation but to the right of the programmed path. For reference, the slope of the parallel lines comprising the programmed path was -1.3032. Thus, the average slopes shown in Table 6 agree quite well with the desired direction.

Overall averages, however, do not tell the whole story. Hence, we present average values for the six individual flights for these same variables in Table 7. The right hand column in Table 7 shows the standard deviations of these averages as obtained from the analyses of variance for these four variables. Again, these data need little explanation. We note that for Flight No. 4 the average of the systems' indicated position also was to the left of the programmed path. Furthermore, the largest CCE(E) was observed for this flight. We have discussed the estimation of sampling error for the regression orthogonal slope. Here we see the application of this estimation even though we have no direct sampling theory for b_{EO}. The values shown for each flight are based on average slopes for all six LEGs. The estimated pooled standard deviation for these slopes is only 0.0038.

SUMMARY. In this paper we have considered the assessment of the performance of navigation and positioning systems. Such assessment comprises two parts: (1) the development of a comprehensive TEST PLAN; and, (2) adequate statistical analyses of the data collected. We emphasize that no more information can be extracted from data than has been built into the structure of a test program [15]. This structure is created by the TEST PLAN.

Our discussion of the TEST PLAN has been developed from an Outline of the "General Features of a Test Program." Among the ten items included in this outline we have directed particular attention to the following:

(1) Selection of the system response or performance measures.

(2) Definition of an experimental unit.

(3) Selection of the treatment combinations.

(4) Determination of the pattern of experimentation or choice of experimental design.

(5) Blocking of the test program against time or other sources of variation in the test program, and,

(6) The sample size or how many experimental units should be completed.

For analysis of the test data we have considered the assessment of both precision and accuracy. There are many ways of presenting data summaries to provide information on both of these characteristics of system performance. We have illustrated the application of the analysis of variance in different ways. Generally, we prefer this approach because of the ability to subdivide the total experimental variation into sources associated with the structure of the TEST PLAN. In using some results obtained from the flight test program for a navigation system we have been able to give only a small sample of the many analyses performed for measuring both precision and accuracy. The latter we note also has been referred to as: (1) predictability; and, (2) reproducibility. For one measure of predictability, the slope or direction of a flight path, we showed how to measure directly the variability of the slope estimates.

REFERENCES

1. Emil H. Jebe, "General Features of a Test Program," unpublished, dated 5 December 1962 (circulated for various purposes since that date).

2. E. B. Wilson, Jr., "Intro. to Scientific Research," N.Y., McGraw-Hill (1952). (Now available in paperback).

3. D. J. Finney, "The Theory of Experimental Design," Chicago, University Press (1960) (in paperback).

4. D. R. Cox, "Planning of Experiments," N.Y., J. Wiley and Sons (1958.

5. ASTM Designation: E-177-68T, "Use of the Terms Precision and Accuracy as Applied to Measurement of a Property of a Material," Phila., ASTM (1968).

6. James O'Day, et al., "Field Test of a Navigation System," Vol. I, Rpt. of Project Michigan, December 1966, Willow Run Laboratories, IST, The University of Michigan, Tech. Rpt. ECOM-00013-58(I), USAECOM, Contract DA-28-043 AMC-00013(E).

7. Henry W. Bigelow, USN Oceanographic Office, "Electronic Positioning Systems," Undersea Technology, April 1964 (pp. 24-28).

8. O. Kempthorne, "Design and Analysis of Experiments," N. Y., J. Wiley & Sons, (1952).

9. H. Scheffe, "The Analysis of Variance," N. Y., J. Wiley & Sons (1959).

10. W. J. Dixon & F. J. Massey, Jr., "Intro. to Stat. Analysis," N.Y., McGraw-Hill, 3rd ed. (1969).

11. G. W. Snedecor & W. G. Cochran, "Stat. Methods," Ames, Ia., University Press, 6th ed. (1967).

12. J. W. Tukey, "The Future of Data Analysis," Anals of Math. Stat. 33, 1 (1962).

13. Mary Natrella, "Experimental Statistics," NBS Hdbk 91, USGO (1963) (also ORDP 20-110--20-114 and AMCP-706-110-114).

14. J. B. Coleman, "A Coefficient of Linear Correlation Based on the Method of Least Squares and the Line of Best Fit," Annals of Math. Stat. 3, 79 (1932).

15. C. R. Rao, "Adv. Stat. Methods in Biom. Research," N.Y., J. Wiley & Sons, (1952) (see quotation from R. A. Fisher, preceding the Preface).

A UNIFIED PROCEDURE FOR SELECTING ALTERNATE EXPERIMENTAL DESIGNS

Edwin M. Bartee
Vanderbilt University
Nashville, Tennessee

Considerable knowledge has been developed in the literature that provides for the more effective design of experiments using, primarily, certain statistical techniques for analysis purposes. Such methods are concerned, for the most part, with analyzing the degree of dependence between the variables. These techniques have exerted a significant influence upon the amount of precision and accuracy that is realized in many experiments.

Additional impact on the optimization of experiments is potentially possible through the application of modeling techniques in the synthesis of experiments. Such techniques are concerned with the design of the experimental model, providing a basis for systematic optimization of the design criteria.

Design Criteria

As in any engineering design problem, the ultimate character of the final design is dictated by certain design criteria. Some typical criteria for an experimental design are as follows:

1. The number of factors to be varied
2. The number of levels to be measured for each factor
   (a) Are levels qualitative or quantitative?
   (b) Are nonlinear effects to be measured?
   (c) Are deviations to be measured from a nominal?
   (d) Are all factors to be set at an equal number of levels?
3. The number of measurements of the response variable to be taken
   (a) Are interactions to be measured?
   (b) Are there any physical limitations on the number of measurements in the experiment?
   (c) What precision is required for measuring experimental error?

This article has been reproduced photographically from the author's manuscript.

The synthesis of an experimental model will be discussed in three steps:

1. Design of the structural model
2. Design of the functional model
3. Design of the experimental model

The first two criteria are important in the determination of the structural model. Criteria 3 (a) and 3 (b) are important in the design of the functional model. Criterion 3 (c) is the major consideration in the design of the experimental model. The ultimate experimental model is the objective of the design process discussed in this paper. Alternate standard experimental designs are compared to the developed experimental model so that a design choice can be made that will optimize compliance with design criteria.

Such an optimization effort differs with the traditional type in statistical design of experiments. This traditional optimization process is typically concerned with a trade-off between the cost of experimentation and the statistical decision. Such optimization would provide the design criterion in 3 (c); i.e., a determination of the number of measurements required to provide a certain precision in estimating experimental error so that certain risk and/or cost requirements can be met. Optimization of the experimental design in this paper is concerned with the selection of the design that will best meet the design criteria established for the experiment. One of these design criteria usually consists of the number of measurements to be made in order to optimize certain statistical and cost requirements.

The Structural Model

The structural model of an experiment is described by

$$N_s = k_1 \cdot k_2 \cdot k_3 \cdots k_m \cdots k_p \qquad (1)$$

N_s = number of cells (defined below) in the experiment

k = number of levels for a factor or independent variable

$j = 1, 2, 3, \ldots, k$

m = the m^{th} factor or independent variable; $1, 2, \ldots, p$

p = the total number of factors in an experiment

The simplest form of an experiment is the case of one factor, for example x_1, at one level, so that

$$p = 1, \qquad k_1 = 1$$

and thus, from Eq. (1), the structural model becomes

$$N_s = k_1 = 1$$

This model is called a <u>cell</u>, the basic structural unit of an experiment. The next form of an experiment is the case of one factor at two or more levels, so that $p = 1, \quad j_1 = 2, 3, \ldots, k_1$

and thus, the structural model becomes

$$N_s = 2, 3, \ldots, k_1 \text{ cells}$$

The next form or level of an experiment is illustrated by a case in which there are two factors at two or more levels. Thus

$$p = 2, \quad j_1 = 2, 3, \ldots, k_1, \quad j_2 = 2, 3, \ldots k_2$$

and the structural model becomes

$$N_s = k_1 \cdot k_2 \text{ cells}$$

Consider another example. A three-factor experiment is described by

$$p = 3, \quad k_1 = 2, \quad k_2 = 3, \quad k_3 = 3$$

and

$$N_s = k_1 \cdot k_2 \cdot k_3 = 2 \cdot 3 \cdot 3 = 18 \text{ cells}$$

A special case of the structural model occurs when the experiment is symmetrical, meaning that all factors have an equal number of levels. Therefore, when

$$k_1 = k_2 = k_3 = \cdots = k_m = \cdots = k_p$$

Eq. (1) becomes

$$N_s = k_1 \cdot k_2 \cdots k_m \cdots k_p = k^p \qquad (2)$$

To illustrate, let us consider an experiment with two factors, each at two levels, described by

$$p = 2, \qquad k_1 = k_2 = k = 2$$

Thus,

$$N_s = k^p = 2^2 = 4 \text{ cells}$$

For another example consider a case of the symmetrical model with three factors, each at two levels.

$$p = 3, \qquad k = 2$$

$$N_s = N_f = 2^3 = 8$$

Thus, Eq. (2) determines the number of cells for any symmetrical model with p factors each, at an equal number of k levels.

The design criteria that are described by the structural model are:

1. The number of factors
2. The number of levels per factor

These criteria are determined by the objectives of the experiment, the measurability of the factors, the interest in nonlinear effects, etc. They should not be dictated by any limitations upon the total number of measurements that can be made of the response variable. Such limitations, or lack of them, is the concern of the functional model.

The Functional Model

The functional model determines how many cells in the structural model will contain a response measurement. Such functional models are either complete or incomplete. A functional model is considered to be complete when all cells contain a response. A functional model is incomplete when the number of responses are systematically limited, so that the number of responses is less than the number of cells. Each of these basic types of functional models will now be discussed.

The necessary and sufficient conditions for a complete functional model are:

$$N_f = N_s = k_1 \cdot k_2 \cdots k_m \cdots k_p \qquad (3)$$

where:

N_s = the number of cells in the experiment.

N_f = the number of responses in the experiment.

k = the number of factor levels ≥ 2.

p = the total number of factors ≥ 1.

$m = 1, 2, 3, \ldots, p$.

For the special case of symmetry where

$$k_1 = k_2 = k_3 = \cdots = k_p$$

Equation (3) can be written as

$$N_f = N_s = k^p \qquad (4)$$

In both Eq. (3) and (4), it can be observed that the number of cells in the structural model (N_s) and the number of responses in the functional model (N_f) are equal. This equality is the basic characteristic of a complete model. In other words, for every cell there is a response, or

$$N_s = N_f$$

For example, given the experiment with two factors, x_1 and x_2, one at two levels and the other at three levels, we have

$$p = 2, \qquad k_1 = 2, \qquad k_2 = 3$$
$$N_s = N_f = k_1 \cdot k_2 = 2 \cdot 3 = 6$$

A functional model is incomplete when

$$N_f < N_s$$

or when the number of responses in the experimental model are determined, in some systematic manner, to be less than the number of cells. Our concern at this point is to consider the fundamental methods that are involved in designing such an incomplete model.

Functional models can be made incomplete in three fundamental ways. The first of these is the restriction of responses exponentially, so that the number of excluded responses are determined by restriction with the factors in the model. The second method for designing incomplete models is to restrict the responses linearly, so that the number of excluded responses in a model are determined by restriction with a certain number of levels of a single factor in the model. The third method consists of a combination of the first two, in which case the restriction of responses is accomplished by both exponential and linear methods. Each of these methods will now be discussed.

From Eq. (3), the necessary and sufficient conditions for an incomplete functional model whose responses are restricted with factors are

$$N_f = \frac{k_1 \cdot k_2 \cdot \cdots k_m \cdots k_p}{k_1 \cdot k_2 \cdots k_l \cdots k_q} \qquad (5)$$

where: q = the number of factors restricting the number of responses in the model. $l = 0, 1, 2, \ldots q$ (Non-negative integers.)

$m = 1, 2, \ldots, p$

When q is equal to zero, no restriction on responses exists. Consider the case of a structural model with three factors, x_1, x_2, and x_3, with

$$N_s = k_1 \cdot k_2 \cdot k_3 = 2 \cdot 4 \cdot 2 = 16$$

in which the number of responses is to be restricted by one factor, for example x_2. Therefore, we have one restricting factor, making

$$q = 1$$

and, thus from Eq. (5)

$$N_f = \frac{k_1 \cdot k_2 \cdot k_3}{k_2} = \frac{2 \cdot 4 \cdot 2}{4} = 4$$

giving four responses that are contained in the sixteen cells of the structural model.

Consider another case. Suppose that a structural model contained four factors x_1, x_2, x_3, x_4, with

$$N_s = k_1 \cdot k_2 \cdot k_3 \cdot k_4 = 4 \cdot 5 \cdot 4 \cdot 4 = 320$$

Suppose that the number of responses in the functional model is to be restricted by the two factors, x_2 and x_3. Thus, we have the l^{th} factors ($l = 2, 3$) restricting, so that

$$p = 4, \quad q = 2; \qquad k_2 = 5, \quad k_3 = 4$$

and from Eq. (5)

$$N_f = \frac{k_1 \cdot k_2 \cdot k_3 \cdot k_4}{k_2 \cdot k_3} = \frac{4 \cdot 5 \cdot 4 \cdot 4}{5 \cdot 4} = 16$$

giving that 16 responses will be contained in the 320 cells.

For the symmetrical functional model, the exponential characteristic of this restriction method becomes more apparent. From Eq. (5), when

$$k_1 = k_2 = k_m = k_p$$

and

$$k_1 = k_2 = k_1 = k_q$$

then:

$$N_f = \frac{k_1 \cdot k_2 \cdots k_m \cdots k_p}{k_1 \cdot k_2 \cdots k_1 \cdots k_q} = k^{p-q} \qquad (6)$$

where: $q < p$ and is a non-negative integer.

The q restriction becomes a negative exponent of the number of equal factor levels. An example is a case in which a symmetrical model contains three factors, p = 3, each at two levels, k = 2. The structural model is

$$N_s = k^p = 2^3 = 8$$

Suppose that the functional model is to be incomplete by restricting the number of responses with one factor, so that, from Eq. (6)

$$q = 1$$

$$N_f = k^{p-q} = 2^{3-1} = 4$$

giving the functional model containing a total of four responses in eight cells.

Consider another example. Suppose that for the structural model

$$N_s = k^p = 3^4 = 81$$

it is desirable to limit the number of responses in the functional model to nine. The value for q to accomplish this is determined as follows:

$$N_f = k^{p-q} = 3^{4-q} = 9 = \frac{3^4}{3^q}$$

$$3^q = \frac{81}{9} = 9$$

$$q = 2$$

The second method for restricting the responses in an incomplete functional model limits the responses within the levels of a particular factor rather than with q number of factors. This is done by subtracting the total number of blank cells for a particular factor from the total number of cells in the structural model. Thus, from Eq. (3), we have

$$N_f = k_1 \cdot k_2 \cdot \cdots k_m \cdots k_p - c_m k_m \qquad (7)$$

where: c_m = the number of blank cells in each level of the m^{th} factor. (a non-negative integer)

k_m = the number of levels of the m^{th} factor.

$c_m < k_m$

The m^{th} factor can be any one of the p factors in the model. For example, a $p = 2$ model can be systematically limited by arbitrarily determining the number of blank cells to exist in each level of one of the two factors, x_1 and x_2. This is c. The number of responses, N_f, is then calculated from Eq. (7). Consider an example in which the levels for the first factor are six, $k_1 = 6$, and the levels for the second factor are three, $k_2 = 3$. If we choose to restrict the first factor, x_1, so that each factor level has one blank cell, then,

$$k_1 = 6, \quad k_2 = 3, \quad m = 1, \quad c_1 = 1$$

The number of cells are

$$N_s = k_1 \cdot k_2 = 6 \cdot 3 = 18$$

and the number of responses are, from Eq. (7)

$$N_f = k_1 \cdot k_2 - c_1 k_1 = k_1 (k_2 - c_1) = 6 (3 - 1) = 12$$

In the case of a symmetrical model, we determine the incomplete functional model from Eq. (7) to be

$$N_f = k_1 \cdot k_2 \cdots k_m \cdots k_p - c_m k_m$$

and since: $k_1 = k_2 = k_m = k_p$

$$N_f = k^p - ck^{p-1} \qquad (8)$$

where: c = the number of blank cells in the level of any p factor.

Eq. (8) gives emphasis to the <u>linear</u> feature of this method. In the case of the model

$$N_s = k^p = 3^4 = 81$$

we could limit the number of responses by creating blank cells in the factors. For example, with c = 1, we can calculate from Eq. (8)

$$N_f = k^p - ck^{p-1}$$
$$= 3^4 - (1)\, 3^{4-1} = 81 - 27 = 54$$

The model can be used in a different, and more useful, way from a design standpoint. As an example, what value of c is required to reduce the model

$$N_s = k^p = 5^5 = 3125$$

to the functional model of

$$N_f = 625$$

This is determined from Eq. (8) thus:

$$N_f = k^p - ck^{p-1} = 625$$

$$5^5 - c5^4 = 625$$

$$c = \frac{3125 - 625}{625} = 4$$

Therefore, the functional model can be restricted to 625 responses by providing for four blank cells in each factor level.

In order to further increase the possible combinations of N_f values, the third method utilizes both the q and c criteria. This can be accomplished by Eq. (6) and Eq. (8) to become

$$N_f = k^{p-q} - c\,k^{p-q-1} \qquad (9)$$

so that the number of restricting factors, q, and the number of blank cells per factor level, c, can be used to determine a particular number of responses for a given model. The application of Eq. (9) will be illustrated by an example. Suppose that it is desirable to restrict the number of responses for the model

$$N_s = k^p = 4^3 = 64$$

to eight responses. This can be done by using Eq. (9), and following a systematic procedure. First, assume c = 0, and q = 1

$$N_f = k^{p-q} - ck^{p-q-1}$$
$$= 4^{3-1} - 0 = 16$$

which is greater than the desired number. Next, keep c = 0 and assume q = 2

$$N_f = 4^{3-2} - 0 = 4$$

which is less than the desired number. Therefore, hold q = 1, and assume c = 1

$$N_f = 4^{3-1} - 4^{3-1-1} = 16 - 4 = 12$$

which is more than desired. Next, hold q = 1 and set c = 2

$$N_f = 4^{3-1} - (2)\,4^{3-1-1} = 16 - 8 = 8$$

which is the desired number of responses.

Suppose that we wanted to determine how to design a functional model with nine responses for the model

$$N_s = k^p = 3^6 = 729$$

with k = 3, the desired number of responses is $k^2 = 9$. It can be seen that such a value for N_f is possible in two ways. First, $N_f = k^2$ for the case when

$$q = p - 3 = 6 - 3 = 3$$
$$c = k - 1 = 3 - 1 = 2$$

Therefore, in this problem

$$q = 3, \quad c = 2$$
$$N_f = k^{p-q} - c\,k^{p-q-1}$$
$$= 3^{6-3} - (2)\,3^{6-3-1}$$
$$= 27 - 18 = 9$$

The same number of responses can be obtained with a different combination of q and c. $N_f = k^2$ is possible with

$$q = p - 2 = 6 - 2 = 4$$
$$c = 0$$

Therefore, the model becomes

$$N_f = k^{p-q} = 3^{6-4} = 3^2 = 9$$

A complete functional model is the same as a factorial experiment, with a single response in each cell. An incomplete functional model is desirable when there is no interest in interaction effects and the total number of measurements required is less than N_s. An incomplete model is necessary when the total possible number of measurements is less than N_s. More specifically, the design of the functional model is made to meet the following design criteria:

1. The total number of possible measurements is equal to or less than N_e.
2. The total number of measurements is limited to some number less than N_s because of some physical limitation of the experimental situation or equipment.

The Experimental Model

The final step in the synthesis of an experiment is to design the experimental model. The experimental model is described by

$$N = n N_f \qquad (10)$$

where:

n = the number of replications of the experiment

N_f = the number of responses in the functional model

N = the total number of responses in the experiment

From Eq. (9) and Eq.(10) we get

$$N = n N_f = n (k^{p-q} - c k^{p-q-1}) \qquad (11)$$

which provides a general expression for a symmetrical experimental model. Eq. (10) thus defines the experimental model as follows:

> The total number of responses in an experiment is a function of the number of factors (p), the number of factor levels (k), the number of factor restrictions (q), the number of cell restrictions (c), and the number of replications (n).

Given the number of factors and factor levels, the number of possible values for N can be determined by certain combinations of values for n, q, and c. For example, if it is desirable to design an experimental model with 54 responses of the type

$$N_s = k^p = 3^4 = 81$$

we can set q = 0, c = 1, and n = 1 and get

$$N = n\,(k^{p-q} - c\,k^{p-q-1})$$
$$= 1\;3^{4-0} - (1)\,3^{4-0-1}$$
$$= 81 - 27 = 54$$

Table 1 provides a general tabulation of the experimental model in Eq. (11)

TABLE 1. Values of N for All Values of q, c, and n in a Symmetrical Experimental Model

q	c					
	k	k - 1	k - 2	k - 3	. . .	0
p	0	n/k	2 n/k	3 n/k	. . .	n
p - 1	0	n	2 n	3 n	. . .	nk
p - 2	0	nk	2 nk	3 nk	. . .	nk^2
p - 3	0	nk^2	2 nk^2	3 nk^2	. . .	nk^3
.	.	.	.	.		.
.	.	.	.	.		.
.	.	.	.	.		.
0	0	nk^{p-1}	2 nk^{p-1}	3 nk^{p-1}	. . .	nk^p

The number of responses, N, for a symmetrical experimental model can be determined if given the values for p, k, q, c, and n. As an

example of its use, suppose that we have a model with p = 5 factors and each factor has k = 5 levels so that

$$N_s = 5^5 = 3{,}125$$

Assume that the experimental model is to contain forty-five responses. The responses are first limited by

$$q = p - 2 = 5 - 2 = 3 \text{ factors}$$

Also, the responses are further restricted by

$$c = k - 3 = 5 - 3 = 2 \text{ blank cells per factor}$$

When $q = p - 2$ and $c = k - 3$

$$N = 3nk$$

and with $k = 5$

$$N = 3n(5) = 15n = 45$$

$$n = 3$$

Selection of Optimal Alternate Designs

The experimental model provides the specifications necessary for the final experimental design to meet the established design criteria, as to total number of responses. Such a selection is not concerned with the problems of balancing the responses in the cells or randomizing the arrangement of the responses. These are considerations made in certain standard designs with which the subject design procedure is not concerned.

The synthesis of any experiment can be described by its experimental model. For example, a complete factorial experiment is described by the following necessary and sufficient conditions:

$$n = 1, 2, 3, \ldots, \quad k = 2, 3, 4, \ldots, \quad p = 2, 3, 4, \ldots, \quad q = 0, \quad c = 0$$

A specific case of such a factorial experiment is

$$n = 1, \quad p = 3, \quad k = 4, \quad q = 0, \quad c = 0$$

and from Eq. (11)

$$N = n\,(k^{p-q} - c\,k^{p-q-1}) = (1)\,(4^3) = 64$$

Examples of other models for certain traditional experiments are listed as follows:

1. A <u>one-way classification experiment</u> with five responses in each of four columns of a single factor is described by

$$n = 5, \quad p = 1, \quad k = 4, \quad q = 0, \quad c = 0$$

$$N = n\,(k^{p-q} - c\,k^{p-q-1})$$

$$= (5)\,(4^{1-0}) = 20$$

Such a model thus explains the one-way classification experiment as a single factor experiment that is replicated.

2. Consider a <u>nested</u> experiment with three factors: x_1, with two levels, x_2, with four levels, and x_3 with two levels. Factor x_2 is such that only half of its levels are crossed with each of the two levels of x_1. Thus:

$$p = 3, \quad k_1 = 2, \quad k_2 = 4, \quad k_3 = 2$$

From Eq. (1)

$$N_s = k_1 \cdot k_2 \cdot k_3 = 2 \times 4 \times 2 = 16$$

Since factor x_1 restricts the number of responses in the experiment

$$q = 1, \quad k_1 = k_1 = 2$$

and, thus, from Eq. (5)

$$N_f = \frac{k_1 \cdot k_2 \cdot k_3}{k_1} = \frac{2 \cdot 4 \cdot 2}{2} = 8$$

with only one replicate

$$N = n\,N_f = (1)(8) = 8$$

A hierarchical layout becomes

Factor X_1															
1								2							
Factor X_2								Factor X_2							
1		2		3		4		1		2		3		4	
X_3		X_3		X_3		X_3		X_3		X_3		X_3		X_3	
1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
X	X	X	X									X	X	X	X

A matrix layout becomes

		X_1			
		1	2		
		X_2			
		1	2	3	4
X_3	1	X	X	X	X
	2	X	X	X	X

3. Consider a Latin Square experiment. The necessary and sufficient conditions for this symmetrical restricted model are

$$p = 3, \quad q = p - 2, \quad k \geq p + 1, \quad c = 0$$

The minimum case occurs when there are three factors, each at two levels, with

$$p = 3, \quad k = 2, \quad q = p - 2 = 1$$

and the number of cells and responses are

$$N_s = k^p = 2^3 = 8$$

$$N_r = k^{p-q} = 2^2 = 4$$

so the experiment contains a total of four responses in eight cells.

A hierarchical layout becomes

X_1							
1				2			
X_2				X_2			
1		2		1		2	
X_3		X_3		X_3		X_3	
a	b	a	b	a	b	a	b
X			X		X	X	

A matrix layout becomes

		X_1	
		1	2
X_2	1	a	b
	2	b	a

4. A Graeco-Latin Square experiment is described by the following necessary and sufficient conditions

$$c = 0, \quad k \geq 3, \quad p = 4, \quad q = p - 2 = 2$$

For the case in which $k = 3$, the number of cells in the structural model would be

$$N_s = k^p = 3^4 = 81$$

and the number of responses would be

$$N_f = k^{p-q} = 3^{4-2} = 9$$

Only one replicate is taken. Thus,

$$N = n N_f = (1)(9) = 9$$

5. An incomplete block experiment is represented by the incomplete functional model whose necessary and sufficient conditions are

$$N = n(k_1 \cdot k_2 \cdots k_m \cdots k_p - c_m k_m)$$

where:

$$p = 2, \quad k_1 \geq 3, \quad k_2 \geq 3, \quad n = 1$$

Thus:

$$N = k_1 \cdot k_2 - c_m k_m$$

where: $m = 1$ or 2

c_m = the number of black cells in the j_m^{th} level.

One of the two factors is a block.

An example would be a model as follows:

k_1 (blocks) = 6 $\qquad$ k_2 (treatments) = 3

$$c_1 = 1, \quad c_2 = \frac{k_1}{k_2} c_1 = \frac{6}{3}(1) = 2$$

Therefore:

$$N_s = k_1 \cdot k_2 = 6 \cdot 3 = 18 \text{ cells}$$

$$n = 1$$

$$N = k_1 \cdot k_2 - c_2 k_2 = 6 \cdot 3 - 2(3) = 12$$

so the twelve responses are to be balanced in the eighteen cells of the incomplete block design.

6. A <u>symmetrical incomplete block</u> experiment is described by the necessary and sufficient conditions from Eq. (8)

$$k \geq 3, \quad p = 2, \quad c = 1, 2, \cdots, k-1$$

$$n = 1$$

$$N = k^2 - ck$$

where: One of the two factors is a block.

7. A <u>Youden Square</u> experiment is described by the following necessary and sufficient conditions from Eq. (9)

$$N = n\,(k^{p-q} - c\,k^{p-q-1})$$

$k \geq 2, \quad p = 3, \quad q = 1, \quad c = 1, 2, \cdots, k-1$

$$n = 1$$

A specific example is a case in which

$$p = 3, \quad k = 4, \quad q = 1, \quad c = 1, \quad n = 1$$

$$N_s = k^p = 4^3 = 64$$

$$N = k^{p-q} - c\,k^{p-q-1}$$

$$= k^{p-q}\left(1 - \frac{c}{k}\right) = 4^{3-1}\left(1 - \frac{1}{4}\right) = 12$$

8. A <u>lattice square experiment</u> is described as a type of incomplete block (See 6 above) that is replicated. Thus

$$N = n\,(k_1 \cdot k_2 - c_2 k_2)$$

where: $2 \leq n \leq t + 1; \quad t \geq 3$

$k_1 \geq m\,t$ (blocks)

$k_2 = t^2$ (treatments)

$c_2 = k_1 - 1$

Example:

$t = 3, \quad n = 4, \quad k_1 = 12, \quad k_2 = 9, \quad c_2 = 11$

$N_s = 12 \cdot 9 = 108$

$N = 4\,(12 \cdot 9 - 11 \cdot 9) = 36$

Another example of a lattice square will demonstrate the relationship between the structural, functional, and experimental model more clearly. A 13 x 13 balanced lattice square will be used to illustrate. ($t = 13, n = 7$.)

Structurally speaking, the experiment consists of two factors: blocks at $k_1 = 21$ levels and treatments at $k_2 = 169$ levels. Thus

$$N_s = k_1 \cdot k_2 = 21 \cdot 169 = 3{,}549 \text{ cells}$$

The functional model is incomplete and its characteristic feature is that only one treatment can occur in each replicate. Thus, the functional model is restricted so that for each treatment all cells are empty except one. Thus $c_2 = 20$, giving

$$N_f = k_1 \cdot k_2 - c_2 k_2 = 3,549 - 20(169) = 169$$

The experiment is replicated 7 times, thus the experimental model becomes

$$N = n N_f = 7 (169) = 1,183 \text{ responses}$$

which is the total number of responses to be balanced.

9. Following are a number of miscellaneous incomplete block designs with their corresponding structural, functional, and experimental models.

BALANCED DESIGN FOR 9 TREATMENTS IN BLOCKS OF 3 UNITS

Block	Rep. I		Rep. II		Rep. III		Rep. IV	
(1)	1 2 3	(4)	1 4 7	(7)	1 5 9	(10)	1 8 6	
(2)	4 5 6	(5)	2 5 8	(8)	7 2 6	(11)	4 2 9	
(3)	7 8 9	(6)	3 6 9	(9)	4 8 3	(12)	7 5 3	

$$p = 2, \quad k_1 = 9, \quad k_2 = 12$$

$$N_s = k_1 \cdot k_2 = 9 \cdot 12 = 108$$

$$c_1 = 11$$

$$N_f = N_s - c_1 k_1 = 108 - (11)\,9 = 9$$

$$n = 4$$

$$N = n N_f = 4 (9) = 36$$

BALANCED DESIGN FOR 7 TREATMENTS IN BLOCKS OF 3 UNITS

Block							
(1)	1 2 4	(3)	3 4 6	(5)	1 5 6	(7)	1 3 7
(2)	2 3 5	(4)	4 5 7	(6)	2 6 7		

$p = 2, \qquad k = 7$

$N_s = k^p = 7^2 = 49$

$c = 4$

$N_f = N_s - c\,k^{p-1} = 49 - (4)\,7 = 21$

$n = 1$

$N = n\,N_f = (1)\,21 = 21$

BALANCED DESIGN FOR 9 TREATMENTS IN 4 LATTICE SQUARES

	Rep. I		Rep. II		Rep. III		Rep. IV
				Columns			
Rows	(1)(2)(3)		(4)(5)(6)		(7) (8) (9)		(10)(11)(12)
(1)	1 2 3	(4)	1 4 7	(7)	1 6 8	(10)	1 9 5
(2)	4 5 6	(5)	2 5 8	(8)	9 2 4	(11)	6 2 7
(3)	7 8 9	(6)	3 6 9	(9)	5 7 3	(12)	8 4 3

$p = 3, \quad k_1 = 9, \quad k_2 = 12, \quad k_3 = 12$

$N_s = k_1 \cdot k_2 \cdot k_3 = 9 \cdot 12 \cdot 12 = 1{,}296$

$c_1 = 143$

$N_f = N_s - c_1 k_1 = 1{,}296 - 143\,(9) = 9$

$n = 4$

$N = n\,N_f = 4\,(9) = 36$

BALANCED DESIGN FOR 7 TREATMENTS IN AN INCOMPLETE LATIN SQUARE

	Columns (Blocks)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rows							
(1)	1	2	3	4	5	6	7
(2)	2	3	4	5	6	7	1
(3)	4	5	6	7	1	2	3

$p = 3, \quad k = 7$

$N_s = k^p = 7^3 = 343$

$q = 1, \quad c = 4$

$N_f = k^{p-q} - c\,k^{p-q-1} = 7^{3-1} - (4)\,7^{3-1-1} = 21$

$n = 1$

$N = n N_f = 21$

10. A composite design is used to estimate the regression coefficients for a second degree ploynomial. These designs are traditionally constructed by adding further treatment combinations to those obtained from a 2^p factorial. Such designs are described here as replicated incomplete models with $N_s = 3^p$. Such an approach recognizes the necessity for three factor levels to measure second degree effects. Thus

$q = p - 1, \quad c = 2$

$N_f = 3^{p-p+1} - (2)\,(3)^{p-p} = 3 - 2 = 1$

$n = 2^p + 2p + 1$ (central design)

$n = 2^p + 2p + p$ (noncentral design)

$N = n N_f = 2^p + 2p + 1$ (central)

$N = n N_f = 2^p + 2p + p$ (noncentral)

Examples

1. Design Criteria. A flight vehicle trajectory is to be designed so that a multistage rocket may place a payload into a circular orbit about the earth. An experiment is to be designed to determine how the first stage booster thrust program affects the amount of mass which is injected into a circular orbit with an altitude of 100 miles. The thrust program of the booster consists of programmed adjustments in the angle of attack for a given time period. The length of the first stage burn, t_b, is determined by the propellant loading of the first stage. The parameters that control the thrust program are (1) the initial rate of increase of the angle of attack, R, and (2) the length of time this rate is flown, t_1.

The factor, R, is to be set at six different rates of increase and t_1 is set at four time levels. The flight is to be simulated on a digital computer. Previous experience in similar studies indicates that each run (response) requires about 3/4 of a minute of computer time. The computer "turn around" time is very slow. It is necessary to obtain the maximum priority time required on the computer, which is 10 minutes. Therefore, a maximum of 13 runs can be made. There is no interest in interaction effects.

Model Synthesis. The structural model, according to the above design criteria, is determined to be

$$p = 2 \text{ factors}, \quad k_1 = 6 \text{ levels}, \quad k_2 = 4 \text{ levels}$$

$$N_s = k_1 \cdot k_2 = 6 \cdot 4 = 24 \text{ cells}$$

The maximum number of responses requires an incomplete functional model restricted by the c criteria. Thus

$$N_f = k_1 \cdot k_2 - c_1 k_1$$

Setting $N_f = 13$ (maximum) we calculate

$$c_1 = \frac{6 \cdot 4 - 13}{6} = 11/6$$

which is not an integer. Using $N_f = 12$ we get

$$c_1 = 2$$

Therefore, the number of blank cells per level of factor R is 2, providing a total of 11 degrees of freedom. With the two main effects requiring 5 and 3 degrees of freedom, respectively, the experimental error is estimated with three degrees of freedom since the experiment is not replicated. Thus

$$n = 1$$

$$N = n N_f = (1)(12) = 12$$

2. Design Criteria. Two different analogue-to-digital converters are contained in test stations used in checking out a particular instrument unit. An experiment is designed to determine the causes of variation in the digital output of these converters. The response variable is the difference between input voltage and output voltage. The variables to be measured are (1) input voltage, (2) converter units, and (3) adjustments. The input voltage is to be set at two levels, -10 volts and +10 volts. The number of converters are limited to two. The adjustments consist of gain and balance settings as specified by the manufacturer. Four different adjustments will be made. The adjustments are unique with each unit and, therefore, they cannot be duplicated between the two converters. Thus, the first two adjustments will be unique with the first unit and the second two adjustments will be unique with the second unit. All possible interactions are to be measured. The optimal degrees of freedom for the error estimate, considering cost of experimentation and desired decision confidence levels, has been determined to be 16 in a previous study.

Model Synthesis. The structural model is

$$p = 3, \quad k_1 = 2, \quad k_2 = 2, \quad k_3 = 4$$

$$N_s = k_1 \cdot k_2 \cdot k_3 = 2 \cdot 2 \cdot 4 = 16$$

The converter factor restricts the adjustment factor thus providing conditions for an incomplete functional model, restricted by the q criteria. Thus

$$c = 0, \quad q = 1 \text{ (converters, } k_2 = 2)$$

$$N_f = \frac{k_1 \cdot k_2 \cdot k_3}{k_2} = \frac{2 \cdot 2 \cdot 4}{2} = 8$$

The optimal degrees of freedom for the error estimate, 16, is provided by replicating the functional model. Thus

$$n = 3$$

$$N = nN_f = 3(8) = 24$$

The degrees of freedom are partitioned as follows:

Source of Variation	Degrees of Freedom
Converters (C)	1
Adjustments (A)	2
Voltage (V)	1
C V	1
A V	2
Error	16
Total	23

A layout of the selected experiment, which is a nested factorial, is shown in Table 2.

	Converters			
	1		2	
	Adjustments		Adjustments	
Voltage	1	2	3	4
-10	X	X	X	X
	X	X	X	X
	X	X	X	X
+10	X	X	X	X
	X	X	X	X
	X	X	X	X

Table 2. Layout of Analogue-to-Digital Converter Experiment

3. Design Criteria. An electronic manufacturer has designed a component board using four capacitors to establish a time base. He wishes to test five different brands of the capacitors in the component boards. Four capacitors are placed in parallel and then connected through a resistor to an input plug where a fixed voltage may be applied. The voltage across the capacitors is connected to an output jack. The test is made by applying a fixed voltage to the plug at the input of the component board. The output jack is monitored with an oscilloscope to measure the time required for the output voltage to rise to a specified amplitude.

The response variable (T) is the time required for the output of the component board to rise to a specified amplitude upon application of a fixed input voltage. The factors of interest are capacitor

brands (C) and component boards (B). Since both factors are qualitative, nonlinear effects are not applicable. Also, past experience in tests of this type has shown negligible interaction between the capacitors and component boards. Since information is desired on the capacitors only, the same resistor will be used for each test. A set of terminals allows the resistor to be plugged in or removed from the component board. Five different capacitor brands are, therefore, to be tested in a circuit that is limited to four capacitors. A minimum of 10 degrees of freedom is required to make an error estimate.

Model Synthesis. Since five brands are being tested, it would seem reasonable to test these brands in five different component boards. We, therefore, have a symmetrical model. The structural model is

$$p = 2, \qquad k = 5$$

$$N_s = k^p = 5^2 = 25$$

Since there are four capacitors in the circuit but five different brands we will have one missing value in each level of capacitor brand. Thus, the functional model is incomplete with

$$q = 0, \quad k = 5, \quad p = 2, \quad c = 1$$

$$N_f = k^p - c\,k^{p-1} = 5^2 - (1)5^{2-1} = 20$$

The degrees of freedom are

Source of Variation	Degrees of Freedom
Capacitors (C)	4
Boards (B)	4
Error	12
Total	20

Only one replicate is required since the minimum of 10 degrees of freedom is met. The experimental model is

$$n = 1$$

$$N = n N_f = (1)(20) = 20$$

A balanced layout of the experiment is shown in Table 3, as an incomplete block design.

Component Boards	Capacitor Brands				
	1	2	3	4	5
I	X		X	X	X
II	X	X	X	X	
III	X	X		X	X
IV		X	X	X	X
V	X	X	X		X

TABLE 3. Incomplete Block Design for Capacitor Experiment

Summary

The modeling of experiments has been described as a three-phase process, namely

1. Designing the structural model
2. Designing the functional model
3. Designing the experimental model

The structural model determines the number of cells in the experiment as a function of the number of factors and the levels for each factor. For the symmetrical case the structural model is

$$N_s = k^p$$

The functional model determines the number of responses to be taken in the structural model. A complete symmetrical functional model is expressed as

$$N_f = N_s = k^p$$

A functional model can be incomplete in three ways. First, if the responses are restricted by q number of factors, the symmetrical functional model becomes

$$N_f = k^{p-q}$$

Second, if the responses are restricted by c cells within a factor, the symmetrical functional model becomes

$$N_f = k^p - c\,k^{p-1}$$
$$= k^p \left(1 - \frac{c}{k}\right)$$

Third, if the responses are restricted by both q and c the symmetrical functional model becomes

$$N_f = k^{p-q} - c\,k^{p-q-1}$$
$$= k^{p-q} \left(1 - \frac{c}{k}\right)$$

The final experimental model is defined as

$$N = n\,N_f$$

for the symmetrical case, where n is the number of replications. All types of matrix experiments can be described by such models.

The unified procedure for selecting alternate experimental designs can be summarized as

1. Determine experimental design criteria
2. Synthesize the experimental model
3. Compare model to standard experimental designs and choose the optimal design.

A PROBLEM IN CONTINUOUS SAMPLING VERIFICATION

Mary E. Blome

U.S. Army Ammunition Procurement and Supply Agency

Joliet, Illinois

There are basically two types of sampling inspection procedures in use today. These are lot-by-lot and continuous sampling procedures. In addition to the two types of sampling inspection, there are also two different methods of inspection, namely, by attributes and by variables. Inspection by attributes is on a go-no-go basis. That is, a unit of product is inspected and determined to be either satisfactory or unsatisfactory with respect to the characteristic under consideration. Under inspection by variables, the actual value of the measurement of a measurable characteristic is recorded. Several of these measurements might then be used together to estimate some parameter upon which a lot of product may be judged relative to its conformity to specification requirements. Our discussion shall be limited to inspection by attributes.

Under inspection by attributes, the inspection can be performed on a lot-by-lot basis or continuously. Let us first consider the lot-by-lot case. The units of product are divided into identifiable lots, and a lot is judged either conforming or nonconforming on the basis of the number of defective units found in a sample from the lot.

One of the most widely used Military Standards listing sampling plans for this type of inspection is MIL-STD-105D, "Sampling Procedures and Tables for Inspection by Attributes," 29 April 1963. When using this Standard, a sampling plan is determined by the following:

This article has been reproduced photographically from the author's manuscript.

a. the size of the lot,

b. the specified acceptable quality level (AQL),

c. the specified Inspection Level (when none is specified, Inspection Level II is used), and

d. the type of plan specified or approved for use (single, double, or multiple).

The size of the lot may be stated in the specifications, or it may be up to the supplier, subject to approval by the consumer, to determine a suitable lot size. The AQL is the maximum percent defective of product which can be considered satisfactory for the process. For example, in this Standard possible AQL values are .010%, 1.0% and 10%. Once a plan has been determined, the plan parameters (sample sizes and acceptance and rejection numbers) can be found.

As an example of a lot-by-lot plan, consider a single sampling plan where the lot size is 1000, the sample size is 100, the acceptance number is 3 and the rejection number is 4. Then, under this plan a random sample of 100 units would be selected from the lot. The number of defective units would be counted, and if the number were 3 or less the lot could be submitted to the consumer for acceptance. If, however, the number of defective units in this sample were 4 or more, then the lot could not be submitted to the consumer for acceptance, and it must be rejected.

Let us now turn our attention to continuous sampling inspection. A limited Standard which defines various types of these sampling plans is MIL-STD-1235(ORD), "Single and Multilevel Continuous Sampling Procedures and Tables for Inspection by Attributes," 17 July 1962. It is a limited Standard

in that it is applicable only to the Army. This Standard is a composite of Inspection and Quality Control Handbooks (Interim) H106, "Multi-level Continuous Sampling Procedures and Tables for Inspection by Attributes," 31 October 1958 and H107, "Single-level Continuous Sampling Procedures and Tables for Inspection by Attributes," 30 April 1959.

In order to use these plans the following criteria must be met:

a. the units of product must be moving, which means that they must pass by the inspection station by means of a conveyor belt or some other conveyance, such as a tote box or skid,

b. the process must produce homogeneous material or be capable of producing homogeneous material,

c. there must be relative ease of inspection, and

d. there must be ample physical facilities for rapid 100% inspection.

All continuous sampling plans are characterized by periods of screening and sampling. The simplest CSP plan is designated CSP-1 and was developed by Dodge (See Annals of Mathematical Statistics, Sept., 1943). Under this plan, 100% inspection (screening) is performed until i consecutive good units have passed inspection. The prescribed value "i" may be some value between 4 and 2000, depending upon the specific plan being used. After i consecutive good units have passed inspection, sampling is begun at a certain prescribed frequency, f. The value of f may be some value between 1/2 and 1/200, again depending upon the specific plan being used. Since each unit of product should have an equal chance of being selected, the interval between the sampled units

should vary somewhat. Sampling is continued until a defective unit is found. When this occurs, screening (100% inspection) begins and continues until i consecutive good units have passed inspection, at which time sampling will again be introduced.

A sampling plan under MIL-STD-1235(ORD) is also determined by the following factors:

a. the number of units in a production interval,

b. the specified AQL, and

c. the specified Inspection Level (when none is specified, Inspection Level II is used) and

d. the type of continuous sampling plan specified or approved for use (CSP-1 or one of the other types of plans provided in the Standard).

The production interval is that period of time, usually a day or shift, during which conditions of manufacture can reasonably be expected to remain stable. Of the four continuous sampling procedures provided in MIL-STD-1235(ORD) CSP-1 is the simplest. It will be the only one considered here.

As an example of a CSP-1 plan, consider one in which i=20 and f=1/10. Screening would be performed until 20 consecutive good units had passed inspection. When this had been accomplished, sampling could begin at the rate 1 in 10. This means that the sampling inspector would select 1 out of 10 units but would vary the interval between these selected units to give each unit of product an equal chance of being included in the sample. Sampling would continue until a defective unit is found. At that time screening would again be instituted, and it would be necessary to screen 20 consecutive good units before sampling could be resumed again.

Verification of the supplier's inspection records is advantageous to the consumer because he would like to ascertain that the supplier is following the inspection plan and classifying inspected units properly. That is, inspected units which are defective should be classified defective and inspected units which are non-defective should be classified non-defective. In order to achieve this aim, AMSMU-P-715-503, "Army Ammunition Plant Quality Assurance Procedures," December, 1966, describes the appropriate procedures to be used by Army Ammunition Plants for verification purposes. This document is designed to be used in conjunction with either lot-by-lot or continuous sampling inspection, and can therefore be used with MIL-STD-105D or MIL-STD-1235(ORD). In the lot-by-lot case, it is a relatively easy matter to perform verification. First, the supplier selects a random sample from the lot in question and counts the number of defective units in this sample. He then compares the number of defective units to the acceptance number for his specified sampling plan from MIL-STD-105D. If the number of defectives is equal to or less than the acceptance number, the lot may be submitted to the consumer for acceptance. The consumer takes a sample from the lot, and counts the number of defective units. The consumer is then ready to compare his results with those of the contractor using Table I of Quality Control and Reliability Handbook (Interim) H109, "Statistical Procedures for Determining Validity of Suppliers' Attributes Inspection," 6 May 1960. For purposes of this comparison, it is assumed that the consumer has classified all of his sample units properly. The H-109 comparison is in effect a test of significance between the number of defectives found by the supplier and the number of defectives found by the consumer, given

a certain value r, which is the ratio of the supplier's sample size to the consumer's sample size. Rejection under this test will cause the supplier's data to be considered invalid.

Verification of inspection results when the sampling inspection is done by continuous sampling procedures is more complicated. Under the provisions of MIL-STD-1235(ORD) and AMSMU-P-715-503, the supplier performs checking inspection at rate f during all periods of screening, in order to ascertain that the screening crew is doing an efficient job. The units inspected during this checking inspection plus the units inspected by the supplier's sampling inspector form the supplier's sample for comparison purposes, where the period under consideration is a production interval.

Concurrently with the inspection by the supplier described above, the consumer is performing verification inspection at rate (1/r)f, where r is the ratio of comparison sample sizes described previously and f is the prescribed sampling frequency. The method of determining the particular value of r (1, 2, 3, 5 or 8) to be used is outlined in AMSMU-P-715-503 and is not important to our discussion here, since we will only concern ourselves with the case r=8.

The various types of inspection described above are summarized in Table I. Reviewing the Table, and from the preceding discussion, it can be noted that only one type of inspection is performed by the consumer, namely, verification inspection, and this is done at a definite sampling frequency which is proportional to that used by the supplier (in the case to be considered here, the proportion is one-eighth). The units inspected in this manner constitute the consumer's verification sample which is used for comparison purposes with

that compiled by the supplier. Again, the sample, which for convenience shall be called the comparison sample hereafter, is composed of units which may have come from the screening or sampling phase with the proportion of units from any phase for a production interval dependent upon the amount of time spent on this phase by the supplier. The consumer usually has no knowledge as to which units came from which phase since verification inspection might be performed at a place far removed from the inspection conducted by the supplier.

Let us consider how these inspections function. Since we are considering only continuous operations under CSP-1 procedures, the units of product will be moving past the various inspection stations via conveyor belts, tote boxes or some other conveyance. Let us first consider the supplier's function. As the operation begins, the product is inspected 100% to remove any defective units and to see if i consecutive good units can be found. Concurrently with this initial product inspection is checking inspection which is performed at a rate f (the specified sampling frequency) and is a means of checking the effectiveness of the screening operation. The units sampled during this checking inspection will form part of the supplier's comparison sample. Once i consecutive good units have been found, sampling inspection of the product is initiated. This sampling of the units of product is done in a random manner at some specified sampling frequency, f. The units sampled form the remainder of the supplier's comparison sample.

Let us now review the consumer's inspection function. As can be seen from Table I, there is only one type of inspection which the consumer performs, namely, verification inspection. This inspection is done concurrently with

the supplier's inspection. The point at which the consumer conducts this inspection may be far removed from the site of the supplier's inspection operations. Since the units of product are not marked or designated as to which units came from which phase, the consumer generally is ignorant of this information. The consumer samples the units in a random manner at a sampling frequency which is proportional to the sampling frequency used by the supplier. This value of the sampling frequency is $[(1/r)(f)]$, where $1/r$ is the proportional factor (one-eighth for purposes of discussion here) and f is the prescribed sampling frequency. Because the sampling is done in a random manner without requiring a certain number or percentage of the inspected units to be from any one phase, there might be a considerable difference in the proportion of units from one of the phases for the consumer and supplier during the production interval.

To use Table I of H-109 to compare d_s ($= d_{s,100} + d_{s,f}$) with d_c ($= d_{c,100} + d_{c,f/8}$), the probability of accepting the hypothesis of validity should remain the same as reflected on the O.C. curves (See Figure I, extracted from H-109) for the test to be of the level α and probability of acceptance over the parameter space as shown on the O.C. curves. By way of explanation the parameter under consideration is the ratio of fractions defective, p_c/p_s, which can be thought of as

$$\frac{\text{Prob (defective inspected unit will be classified defective by consumer)}}{\text{Prob (defective inspected unit will be classified defective by supplier)}}$$

This, then, is our problem: To show that the probability of accepting the hypothesis of validity over the parameter space is approximately the same as that shown on the O.C. curves.

To simplify the remainder of the discussion and the problem definition, the notation below shall be used.

Let

$n_{s,100}$ = number of units in supplier's comparison sample coming from the screening phase,

$n_{s,f}$ = number of units in supplier's comparison sample coming from the sampling phase,

$n_{c,100}$ = number of units in consumer's comparison sample coming from the screening phase,

$n_{c,f/8}$ = number of units in consumer's comparison sample coming from the sampling phase.

Let d, subscripted as above, refer to the number of defective units found in the portion of units identified by the subscripts.

Let us now reflect on some aspects of the problem.

Since there are two phases, namely, the screening phase and the sampling phase, from which the verification sample as well as the supplier's comparison sample can come, there is a possibility of considerable variation between the two in the proportion of units from any one phase. That is, for example,

$\frac{n_{s,100}}{n_{s,100} + n_{s,f}}$ might be considerably

different from $\frac{n_{c,100}}{n_{c,100} + n_{c,f/8}}$.

Let us now consider only one value of the parameter space, $p_c/p_s = 1$, which is equivalent to saying that the supplier has perfect inspection

efficiency. Then no defectives should be found in the samples from the screening phase since these should have been removed during the screening phase of product inspection. Hence, any defective which would be found in either of these samples would come from the sampling phase.

As reflection will show, if the units comprising the samples were selected completely independently of order or position in the production interval, we would have a situation equivalent to a lot-comparison situation, and the O.C. curves would be exactly as defined for H-109. Further, if the proportions described previously were exactly the same, that is, the fraction of the supplier's comparison sample coming from the screening phase were exactly the same as the fraction of the consumer's phase, we would have essentially a stratified sampling problem, and again the O.C. curves would be exactly as defined in H-109.

Since the prescribed method of sampling, however, is to take about one out of every 1/f units, allowing the interval between inspected units to vary somewhat, we have neither of the situations described above. This brings us to the reason why we are only considering the case r=8. It is reasonable to assume that the greatest variation from the O.C. curves of H-109 is possible for the largest value of r. Therefore, if this variation is insignificant for r=8 it should be insignificant for the lower value of r. Let us now consider a specific example.

Since screening need only be done at the initiation of production, and thereafter only when a defect is found during a period of sampling inspection, it is not necessary in our example to assume that screening is initiated at

the start of the production interval, but for sake of discussion let us assume that it does. Suppose the supplier is sampling at a frequency of 1/10, and the consumer is using a ratio of r=8. Therefore, the consumer would be sampling at a frequency of 1/80. First, the supplier's screening crew inspects all units of product until the appropriate number of consecutive good units has been cleared. At the same time, the checking inspector is selecting one unit out of ten in a random manner to see if the screening crew is doing its job properly. After the necessary number of consecutive good units has been cleared, sampling inspection is begun whereby one out of ten units is selected for inspection. There is no checking inspection during this phase.

During the entire production interval, the consumer's verification inspector selects one out of eighty units in a random manner. At the completion of the production interval, the supplier's and consumer's comparison sample inspection results can be compared. The supplier's sample consists of those units inspected by the checking inspector during the screening phase plus the units inspected by the supplier during the sampling phase. The consumer's sample consists of all units inspected by the verification inspector, whether these units came from the screening or sampling phase.

Let us assume that the production interval encompasses 80 units and 76 of these units were subjected to screening while the remaining 4 units were part of the sampling phase. Let us suppose the sampling frequencies are as above, namely, f=1/10 for the supplier and f=1/80 for the consumer.

Reflection will show that there are many possible variations in the values of $n_{s,100}$, $n_{s,f}$, $n_{c,100}$, and $n_{c,f/8}$. It is possible, for example, that all of the units for the supplier's sample came from the screening phase while the single unit composing the consumer's sample came from the sampling phase. In this case, the proportion of units in the supplier's sample from the screening phase is 1.0 whereas the corresponding proportion of units in the consumer's sample from the same phase is 0.

Since the probability of each possible variation is not known, since strict probabilistic sampling is not performed, the effective O.C. curve cannot be determined simply.

Ideally then, a mathematical model describing the O.C. curves would be desirable.

In lieu of such a mathematical model, we conducted a Monte Carlo simulation of the process. Twenty different simulations of various CSP-1 and CSP-2 plans were considered. A few selected AQL's ranging from 0.01% to 4.0% were used, with production intervals ranging from 70 units to 1000 units. The value of p was set equal to the AQL in each case on these first attempts. Ten production intervals were considered for each simulation. Finally, it was assumed that the screening crew was 100% efficient, i.e., all defective units were removed during the screening phase.

Random numbers were used to designate the defective units. Once this had been determined, the inspection processes could be simulated. First, the units from the initial screening phase were identified, and then random numbers were used to select the first unit to be sampled by the supplier.

For convenience on these first attempts, a systematic sample followed the random selection of the first unit. When all ten production intervals had been completed in this manner, the units inspected by the checking inspector during the screening phase needed to be specified. Random numbers were again employed to designate the initial units sampled during these phases and systematic sampling ensued. When this had been completed, the proportion of units from the screening phase for each production interval and for the ten production intervals as a whole could be calculated. Then, the consumer's inspection had to be simulated. Since the sampling was done at a specified sampling frequency without regard as to which phase the supplier was on, a random number was used to indicate the first unit of the sample, and a systematic sampling followed for the duration of the ten production intervals. Upon the completion of the ten production intervals, the proportion of units from the screening phase for each production interval and for the ten production intervals as a whole could be tallied. These proportions could then be compared to the corresponding one for the supplier. Table 2 shows the results of one of these simulations. For this simulation, the production size was 70; the AQL was 2.5%; the i value was 25; the supplier's sampling frequency was 1/5, and the consumer's sampling frequency was 1/40. There does not appear to be too much difference between the proportions except for the seventh production interval where the supplier's proportion was .357, and the consumer's proportion was 0.

In order to use the O.C. curves from H-109, some calculations needed to be performed. The fractions defective for the supplier and consumer as

well as the expected number of defective units in the supplier's sample needed to be specified. Since it was assumed that the screening crew was 100% efficient, theoretically no defective units should have appeared in either the supplier's or the consumer's sample from the screening phase. Therefore, the fraction defective for either the supplier or consumer is the proportion of units from the screening phase times the appropriate AQL (since p was set equal to the AQL, as mentioned previously). Then, the ratio of the consumer's fraction defective to the supplier's fraction defective was calculated. Finally, the expected number of defective units in the supplier's sample was estimated by the number of units in the production interval times the fraction defective described above. The results of these computations for each of the ten production intervals and for the ten production intervals as a whole are summarized in Table 3. The last two columns are of more interest. It will be noted that most of the ratios are around 1.0 except for production interval #7 where the ratio is 2.6040.

Note that all of the expected number of defective units in the supplier's sample for our example are considerably less than the smallest value, indexing the H-109 curves (see figure at end of paper), namely, 0.75. Hence, the O.C. curves for these figures would be above that for 0.75. Also, some of our ratios are less than 1.0 which is the smallest ratio given on the chart. This means that the probability of acceptance for these ratios would be even greater than 0.95 which is the corresponding value when the ratio is 1.0.

While we were unable to develop a suitable model to determine whether the probability of acceptance over the long run would be of any important difference from that yielded by the H-109 O.C. curve formula described by Ellner (see Technometrics, February 1963, pp. 23-46) it seemed reasonable to assume that if the variation of individual simulation results from the H-109 value were small, the probability of acceptance under the continuous sampling verification method could be adequately described by the Ellner formula.

To simplify our work, we arbitrarily decided to concern ourselves only with the frequency of simulation for which the probability of acceptance was less than .90. This would allow us to get a quick picture of the results without having to compute an O.C. curve point for each simulation.

If we consider all of the production intervals, it can easily be seen that they meet the criterion of having a probability of acceptance of greater than .90. Therefore, in this example, it seems reasonable to assume that the O.C. curve under the continuous sampling assumption is probably close to the range of values (94%-96%) provided by the Ellner formula.

Thus, it is possible to study this problem using simulation methods. However, it obviously would be preferable to have a mathematical model. Therefore, to reiterate the problem: a mathematical model describing the operating characteristic of the procedure described is desired.

TABLE I

INSPECTION REQUIRED UNDER CSP-1 AND ASSOCIATED CHECKING AND VERIFICATION INSPECTION

TYPE OF INSPECTION	PERFORMED BY	PHASES WHEN PERFORMED	FREQUENCY	COMPARISON SAMPLE OF	
				SUPPLIER	CONSUMER
1. Product screening	Supplier	Screening	100%		
2. Product sampling	Supplier	Sampling	f	X	
3. Checking	Supplier	Screening	f	X	
4. Verification	Consumer	Screening and Sampling	(1/r)f		X

TABLE II

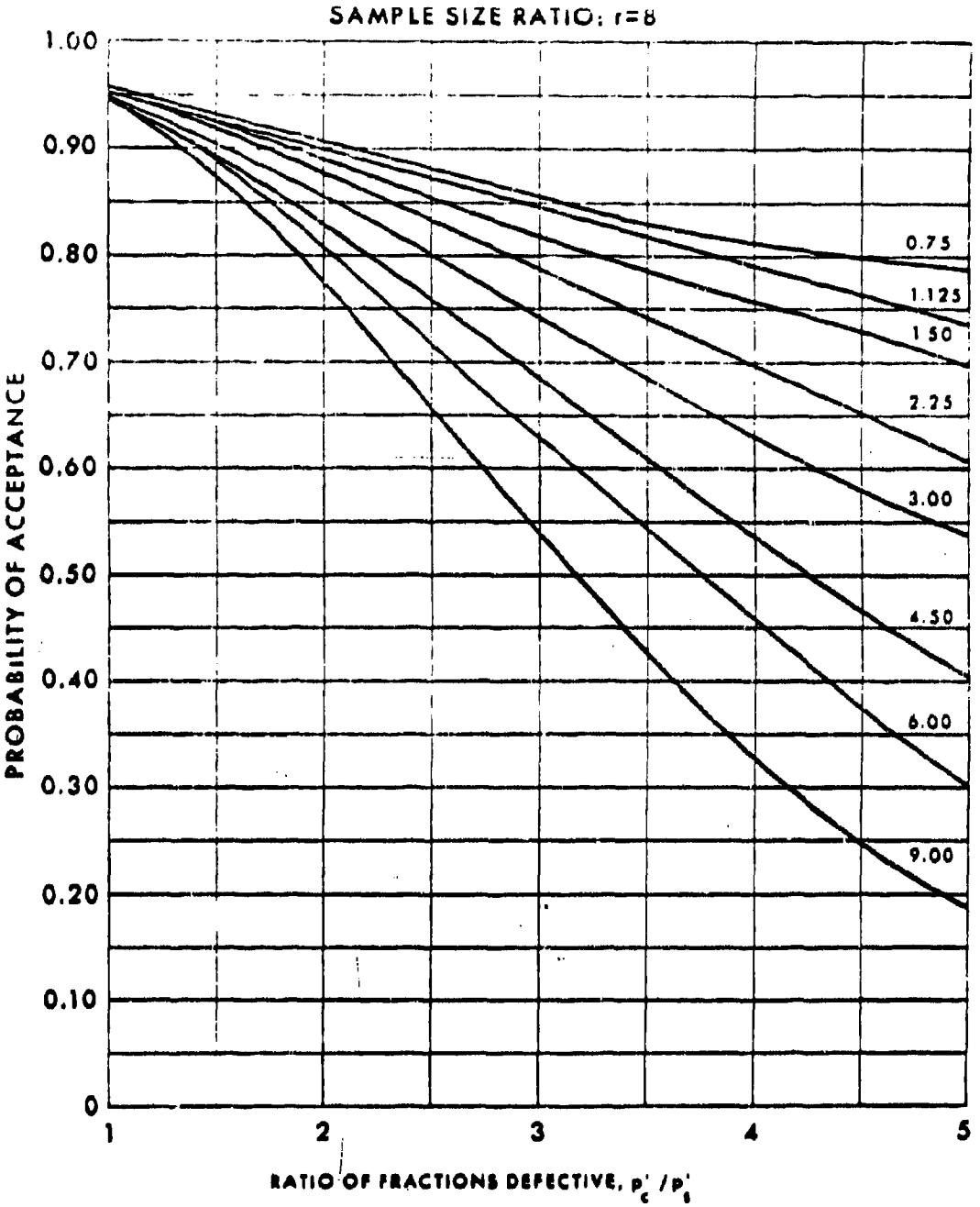
PROPORTION OF UNITS SUBJECTED TO 100% INSPECTION

Production Interval	Screening Phase	
	Supplier	Consumer
1	0.357	0.500
2	0.000	0.000
3	0.000	0.000
4	0.000	0.000
5	0.000	0.000
6	0.000	0.000
7	0.357	0.000
8	0.357	0.500
9	0.067	0.000
10	0.308	0.500
Cumulative	0.143	0.150

TABLE III

Production Interval	P_{eff_c}	P_{eff_s}	$\frac{P_c}{P_s}$	$n_s P_s$ (expected number of defectives in supplier's sample)
1	.0125	.0161	.7760	.2250
2	.0250	.0250	1.0000	.3500
3	.0250	.0250	1.0000	.3500
4	.0250	.0250	1.0000	.3500
5	.0250	.0250	1.0000	.3500
6	.0250	.0250	1.0000	.3500
7	.0250	.0096	2.6040	.1340
8	.0125	.0096	1.3020	.1340
9	.0250	.0230	1.0900	.3450
10	.0125	.0173	.7225	.2250
Cumulative	.021	.021	1.00	.294

OPERATING CHARACTERISTIC CURVES OF TWO-SAMPLE TEST FOR HOMOGENEITY



NOTE:
Figures on curves are the expected numbers of defectives (defects) in the supplier's sample

☆ U.S. GOVERNMENT PRINTING OFFICE 1988 O—55-705

TOWARD A STOCHASTIC MODEL OF TERRAIN

R. H. Peterson, Methodology and Cost Effectiveness Office
Army Materiel Systems Analysis Agency
US Army Aberdeen Research and Development Center
Aberdeen Proving Ground, Maryland

and

William Clare Taylor, Applied Mathematics Division
Ballistic Research Laboratories
US Army Aberdeen Research and Development Center
Aberdeen Proving Ground, Maryland

ABSTRACT

We present an account of an attempt to find useful random models of terrain. Measurements have shown that the distribution of slopes is what has been called the bilateral exponential distribution, definitely not normal. The problem is to find a convenient random function of geographical positions of two real variables which has this distribution for slopes and fits, in some approximations, the dependence of slopes in various directions at neighboring points. A family of random functions, the probability distributions in function space which are spherically symmetric in a Hilbert norm suitable to the purposes of the study, was introduced with an enormous latitude in the choice of parametric functionals. We felt sure that random functions with the required properties must be included. Sad to relate further mathematical developments which we deem intrinsically interesting have shown it not to be so. We know not how to proceed. Help!

This article has been reproduced photographically from the authors' manuscript.

Preceding page blank

We have found it easier to separate our contributions to this exposition although they are interdependent. Peterson has written the first paragraphs and Taylor the later ones as is indicated in the text.

Terrain, being the medium of ground combat, has been the subject of many investigations by analysts in the field of weapon systems analysis. Most of these studies have been focused on the particular role played by terrain in the particular problem at hand. Others have been more general in nature with a goal of giving more insight into the quantitative aspects of describing terrain.

I would like to indicate a sample of the type of problems that arise involving terrain and its influence on the outcome of combat that have received attention. In order to lend some semblance of order to such a listing I have attempted this simple two way classification of these roles. (Figure 1) For lack of better terms I have labeled them scales and mechanisms. For scales I have fallen back on the vernacular of micro and macro, micro generally referring to distances of up to a few meters and macro from there on out to perhaps several kilometers. Mechanisms I have broken into a clear dichotomy of contact and non-contact. By contact I mean that the terrain is actually supporting the objects whether they be vehicles or other pieces of equipment being considered. By non-contact I mean we are concerned with the existence of a line-of-sight. I have listed those roles of terrain which can be fairly well categorized but I will also try to indicate problem areas where there is not a clear distinction or there are strong interactions.

Under contact with the micro structure of terrain I have listed V_{RIDE} and posture. V_{RIDE} has come to be used as an indication of the speed of a vehicle that is tolerable to both the occupants and to the vehicle itself due to roughness of terrain. It is concerned with dynamics of the vehicle over the terrain.

FIGURE 1

ROLES OF TERRAIN IN COMBAT

Mech. \ Scale	Micro	Macro
Contact	V_{RIDE} Posture	Routes
Non-Contact	Fragment Shielding	Ranges of Engagement

Posture refers more to the static role of the vehicle and is concerned with the capability of a weapon system. As an example, if a tank is canted and the gun is elevated, this elevation will introduce a horizontal component of error into the aim of the gun.

Under contact with the macro characteristics of terrain I have listed simply routes. Factors other than slopes influence the routes taken by a vehicle, of course.

Under non-contact in the micro regime I have indicated the fragment shielding which has been quantified in terms of "cover functions". The non-contact aspects of the macro relief are closely tied up with the ranges of engagement. A defender may choose fields of fire to get his opponent out into the open and yet there may be draws and gullies which can allow the attacker to approach under cover.

As an example of the multiple interaction of all of these roles, we might consider the case of a tank hastily taking up a firing position. The tank is advancing along some preplanned axis -- his route has been established. The enemy is encountered -- the approximate range of engagement has been established. The tank may stop or head for a nearby rise in the terrain to get into hull defilade -- V_{RIDE} and shielding come into play. How the tank is canted in position may influence his accuracy -- hence the role of posture.

We see then that there are a number of properties of terrain that are of concern to the military OR analyst and, as I have mentioned at the outset, there are a number of ways that terrain has been categorized, measured, stored in machine memories for retrieval, generated by Monte Carlo means, etc. In order to state the problem which we bring to this clinical session I'd like to discuss two observations concerning the nature of terrain which we feel have not been exploited to their fullest in dealing with this problem area. One concerns the results of a statistical study of terrain slope. The other concerns the underlying geometry of the nature of terrain.

The statistical study to which I refer was conducted to determine the posture of tanks as measured by the pitch and cant of the trunnion after taking up simulated firing positions. A sampling of widely varying terrain types was obtained in that the test was run at Fort Knox, Fort Bragg, Fort Hood, Camp Pickett and Camp Erwin. The pattern that emerged indicated that the distributions of slopes in these firing positions were not normal but seemed to be much more like the bilateral exponential distribution. (Figure 2) Moreover the mean absolute slope varied greatly from one test site to another. In order to check out the possibility that this non-normal characteristic of these distributions was due to the selection of the firing positions sample profiles of each of the test sites were constructed from maps of each of the installations and the distributions of slopes measured over 200 yard intervals were obtained. Here again, the bilateral exponential distribution seemed to be the natural means of describing these slopes.

A detailed map study of the type mentioned above was made of the region around Houffalize, Belgium (based on a map we happened to have available). It showed that the distributions of north-south slopes and of east-west slopes both seemed to fit the bilateral exponential. The inadequacy of the normal distribution for generating profiles from which lines-of-sights can be determined was demonstrated some 20 years ago by people in the U. K. (personal communication from Mr. Eddie Benn then at the Armament Research and Development Establishment). This finding has seemed to influence their subsequent investigations along this line. (See Forbes, "The Generation of Terrain on an Electronic Computer," A.R.D.E. Memorandum (B) 75/60).

In several of the studies mentioned above, attempts were made at establishing distributions of the height or elevation of terrain itself. The results were erratic and no pattern was observed. Such behavior is probably due to general trends which can be attributed to near-zero frequency components in the spectrum.

FIGURE 2

THE FREQUENCY FUNCTION OF THE BILATERAL EXPONENTIAL DISTRIBUTION

$$f(s) = \frac{1}{2a} e^{-|s|/a}$$

where

a = mean absolute deviation

$$a^2 = \sigma^2/2$$

characteristic function

$$\phi(\alpha) = [1 + \sigma^2\alpha^2/2]^{-1}$$

In addition to the non-normal nature of terrain there is the problem of dimensionality. Many of the existing schemes for generating random terrain profiles proceed in the same manner that one would treat a time series. This approach cannot be used to generate a surface, as two neighboring rays say emanating from a point, will be completely independent. Put in terms of statistically describing terrain rather than generating it we must think in terms of the gradient of a surface rather than the slope of a curve. We know from vector analysis that the curl of a gradient is zero. In other words there are constraints between the two perpendicular components of the gradient at a point. In the one dimensional case, as typified by a time series, the random function or stochastic process is readily expressed in terms of Fourier series, i.e., sines and cosines. In the two dimensional case the functions which replace the trigonometric functions in a natural way are the Bessel functions. Other areas of endeavor on which reference to two dimensional random functions have been found include windblown waves, agricultural productivity and images both photographic and video. The household term of snow as applied to a television picture is just an adoption of the television engineer's term "white snow" which is his extension to two dimensions of the concept of "white noise" in the one dimensional process. (We might add in passing that the most well known application of three dimensional random functions is in the field of turbulence.)

We have briefly stated two characteristics of terrain which we believe to be pertinent to the statistical description of terrain. One based on data analysis that, whereas terrain height itself does not seem to have any pattern to its distribution, its difference field as measured over a few meters or a few hundred meters has a common non-normal distribution which can be expressed in terms of a single parameter. The other based on geometrical reasoning indicates that the tools developed for one dimensional processes are not adequate for describing a two dimensional random surface.

We are now at a point of being able to state the problem which has plagued us for a number of years. Is it possible to construct a meaningful stochastic model of terrain which embodies these two considerations?

Questions for which we would like to get more insight, include the following: 1) are what we categorize as rough, rolling or flat in fact simply manifestations of the same basic model with different scale factors in the horizontal and vertical directions, 2) to what extent can we use easily obtained information for a region and infer the details from the model and/or 3) can we build a composite model from which we can infer both the micro and macro characteristics of a given terrain type?

In closing my part of this presentation I want to stress that we are not posing the general question as to how to statistically categorize terrain but as to what extent the theory of two dimensional random functions can contribute to our basic understanding of the statistical properties of terrain.

Dr. Taylor will now describe one approach we have taken to this problem along with its triumphs and pitfalls.

A Class of Random Functions

After careful consideration of some requirements on a random function that it be eligible for consideration as a random terrain, Peterson was led to propose the following wide class of random functions as candidates for investigation. Let

$$P: x = (x_1, x_2) \tag{1.1}$$

be rectangular coordinates of a point P in a horizontal datum plane. Let u(x) be the height of a terrain above the datum plane at the geographical point P. For our purpose u(x) is a complete description of the terrain. We are concerned with a random function U(x), a probability distribution on certain subsets of a set, say B, of functions u(x). We consider a linear set B′ of linear functionals $\ell(u(\cdot))$ and suppose that the expection

$$E(\ell(U)) = 0 \tag{1.2}$$

for all ℓ of the set. By the variance of ℓ we mean

$$E(\ell(U)^2) = \text{Var } \ell \tag{1.3}$$

and by the characteristic functional of ℓ we mean the expectation of the exponential

$$E(\exp i\ell(U)) = Ch\ \ell \tag{1.4}$$

The proposal is to limit our discussion to those random functions for which there exists a complex valued function of a real positive argument g(z) such that, for all ℓ,

$$Ch\ell = g(Var\ \ell) \tag{1.5}$$

Example: For a gaussian random function U,

$$E(\exp i\ \ell(U)) = \exp(-\tfrac{1}{2} E(|\ell(U)|^2)),$$

since $E(\ell(U)) = 0$.

Spherical Symmetry

We may introduce also the inner product

$$\langle \ell_1, \ell_2 \rangle = E(\ell_1(U)\ell_2(U)) \tag{2.1}$$

and

$$\begin{aligned} \ell^2 &= \langle \ell, \ell \rangle \\ &= E(\ell(U)^2) \\ &= Var\ \ell. \end{aligned} \tag{2.2}$$

It is but a small step to extend our discussion to the Hilbert space, H, of linear functionals and to suppose further that this space is sufficient in the following sense: For any u(x) under discussion

$$\ell(u(x)) = 0 \text{ for all } \ell \in H \tag{2.3}$$

implies $u(x) \equiv 0$. This is not necessary for the rather loose discussion we are presenting but it may ease the reader's way. Now a function u(x) defines on H a linear functional whose value at the element ℓ is $\ell(u)$. Whether every linear functional in B' is thus represented by some function u(x) is of no importance to our discussion. What is very important is to

realize that the linear functional defined by u(x) need not be in any fixed sense a bounded linear functional and indeed, for a given $\ell \in H$, $\ell(U)$ need be defined only with probability one.

The preceding discussion of Hilbert spaces has been principally only for orientation. We need at first be concerned only with finite dimensional subspaces defined as follows: disregard all but a finite set of the linear functionals, along with their linear combinations. We define the projection of the measure space, and the measure, into this finite dimensional space by identifying all functions u(x) which agree in the values taken for them by each of this finite set of linear functionals. These finite dimensional spaces are euclidean with the inner product we have introduced. The characteristic functional and the variance of each of these finite dimensional projections of the probability measure will have the same values as when they were considered to be defined on the infinite dimensional space and the characteristic functional defined on the conjugate space will thus be a function only of the distance from the origin. That is to say that it will be spherically symmetric. It follows immediately that the n dimensional measure is spherically symmetric and must be described by a spherically symmetric density -- at least if we assume it to be described by a density at all, and we do. Even though no spheres nor radii are defined on our infinite dimensional space (at least not with positive probability) we may nonetheless define spherical symmetry of the measure: A measure is spherically symmetric if all its projections into finite dimensional subspaces are spherically symmetric.

Characterization of Spherically Symmetric Measures on Infinite Dimensional Spaces*

In each finite dimensional projection of a spherically symmetric measure the density, if supposed to exist, must be the same function of the distance from the center as in any other projection of the same dimension. In n dimensions, let the density at distance r from the center

*We are indebted to J. Feldman and R. M. Dudley for the information that this result concerning spherically symmetric measures in infinite dimensional spaces is not new. It was published in 1962 by Umemura, who obtained it in a more recondite context.

be $p_n(r^2)$. Then, considering the projection of the measure from n + 2 dimensions to an n dimensional subspace, an easy argument shows that

$$p_n'(r^2) = -\pi\, p_{n+2}(r^2) \tag{3.1}$$

From this it follows that the derivatives of each of the p's alternate in sign. Such functions are called completely monotone. There is a theorem of S. Bernstein [see e.g., Feller, Th. of Probability, Vol II, p 415] which states that a completely monotone function $p(z)$, $0 \leq z \leq \infty$, with $p(\infty) = 0$ can be expressed as a linear aggregate of decreasing exponentials with positive coefficients:

$$p(z) = \int_0^\infty e^{-\lambda z}\, d\phi(\lambda),\ 0 < z < \infty, \tag{3.2}$$

with $d\phi(\lambda) \leq 0$.

Setting $z = r^2$, $\lambda = 1/2\sigma^2$ and re-defining the measure $d\phi(\lambda)$, we may then write

$$p_n(r^2) = \int_0^\infty p_{n\sigma}(r^2)\, d\phi(\sigma) \tag{3.3}$$

where

$$p_{n,\sigma}(r^2) = (\sigma\sqrt{(2\pi)})^{-n}\, e^{-r^2/2\sigma^2} \tag{3.4}$$

is the n dimensional gaussian density. This formula, once obtained for any value of n, implies the same formula for all lower dimensional densities, as is seen by successive integration with respect to each of an orthogonal set of coordinates. The integrals are all absolutely convergent and may be integrated freely in any order. The same statement is then true for all n. Further the corresponding statement may be asserted expressing the given measure m similarly in terms of the gaussian measures m_σ:

$$m = \int_0^\infty m_\sigma\, d\phi(\sigma) \tag{3.5}$$

Adjusting the Parameters of the Model

The procedure we are to follow is now quite clear. Whatever may be the distribution of the individual linear functionals, we shall adjust the density $p_1(r^2)$ to it by choosing the weights $d\phi(\sigma)$ in (3.3). A necessary condition is of course that the density be a completely monotone function of r^2. But, as Peterson has pointed out above, it lies very near an exponential function e^{-ar}, which, fortunately, satisfies this condition. We shall need only to be firm with the small residue, if there be any, and its derivatives, and insist that it conform. There will then be the task of fitting the remaining free element, the variance of a linear functional. Here there is a great deal more freedom. There is a functional to be adjusted to approximate as best we can the statistical interdependence of the values of U(x) at neighboring values x. (We want them to become independent at distant points.) But this is just the same problem to be faced in fitting a gaussian random function. For any $\ell(u)$ we need only go to the samples we wish to fit and estimate $E(\ell(U)^2)$, or what is simpler to tabulate, for some linear basis $\ell_1,\ell_2,\ldots$ of the linear functionals we estimate $E(\ell_i(U)\ell_j(U))$ from the samples for all pairs i,j. There is no arbitrary decision left to be made. It's just a question of whether it works or not! or how well it works!

Sad to say, it doesn't work at all. We shall see this without any further examination of samples. The reason lies in an additional significant difference between the finite and infinite dimensional cases.

Lack of Ergodicity

We shall see that (3.5) is, in a reasonable sense, an orthogonal representation of the measure.* For this purpose it is convenient (and

*For the source of the train of thought which led to this analysis, we are indebted to Jacob Feldman for a lucid and provocative briefing on relatively singular measures, a briefing which grew out of a discussion some years ago of the application of information theory to empirical functions. But the simple case with which alone we need be concerned here was known to us as well as many other people long ago. It appears, for example, in a paper of W. T. Martin and R. H. Cameron in the 1940's.

perhaps something equivalent is also necessary) to introduce a sequence of linearly independent bounded linear functionals $\ell_i(u)$, i=1,2,..., which we can then as well suppose to have been replaced by an orthonormal sequence, so that

$$E(\ell_i(u)\ell_j(u)) = \delta_{ij}$$

$$= \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases} \tag{5.1}$$

The question of when and in what sense does a sequence of numbers w_i, i=1,2,..., represent a function u such that

$$w_i = \ell_i(u) \tag{5.2}$$

will not be discussed.

The random variables

$$W_i = \ell_i(U), \quad i=1,2..., \tag{5.3}$$

are uncorrelated but not necessarily independent. However, for any one of the gaussian measures, m_σ, calling the random function U_σ, the random variables

$$W_{\sigma i} = \ell_i(U_\sigma) \tag{5.4}$$

are uncorrelated gaussian variables and hence independent. Since

$$E_{m_\sigma}(W_{\sigma i}^2) = \sigma^2, \tag{5.6}$$

we have, with probability 1,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} W_{\sigma i}{}^2 = \sigma^2. \tag{5.7}$$

The measure m_σ is not essentially altered if we trim its space to the set A_σ of sequences $w_1, w_2, \ldots$ for which (5.7) is true and to those functions u(x) which give rise to such sequences. We restrict our measure, supposed to exist and to be given, to the set

$$A = \sum A_\sigma \tag{5.8}$$

of functions u(x). We do not discuss which subsets of A have probability and how it is defined since this subject seems either too trivial or too difficult.

Suppose now that we test the distribution of slopes in A in the same fashion that was described above. That is, we draw a single sample function and measure slopes at many points on it. Further, for simplicity, suppose these points are far enough apart that we may ignore statistical dependence of the slopes. Each sample function from A is, for some σ, taken from A_σ. Slopes at distant points on it are then independent, identical, gaussian variables and the sample values of a large number of them will characterize their common distribution as gaussian with whatever assurance their number permits. But haven't we brought this about by artificial tampering with the ensemble? No. We have only turned a statement true with probability one into a true statement. Devise a statistical test for the normality of the distribution from which a sample is taken, using statistics whose distribution is independent on the variance of the ensemble. The result of the test will (at least at any specified stage) depend on only a finite sample. A finite set of linear functionals has the same distribution in A as in the original probability measure, on the space we have called B, and the distribution of the statistics of the test will thus have the same distribution in B as in A and as in a gaussian ensemble.

In short then, these random functions fail to represent a random terrain since an orthonormal sequence of linear functionals read off any one sample function have values distributed like independent samplings from a univariate gaussian ensemble. We demand of our model of terrain on the contrary that slopes read at widely separated points have a different distribution, approximately the one described by a bilateral exponential density. More generally, in order to make sense our random function model must have the ergodic property: independent identical functionals (such as slopes at widely separated points) must show the same distribution whether read from a single sample or each from a different randomly chosen one.

A SUGGESTED PROCEDURE FOR ANALYZING MISSILE PERFORMANCE BY A LEAST SQUARES FIT TO A GENERALIZED LINEAR STATISTICAL MODEL AND A QUICK CHECK FOR NORMALITY OF THE DATA

N. R. RICH
Systems Evaluation Branch
Advanced Systems Laboratory
Research and Engineering Directorate
Redstone Arsenal, Huntsville, Alabama

ABSTRACT

The data taken in a series of missile tests are often in the form of a variable of interest (such as radial miss distance from a given target) and several dependent variables (e.g., range, temperature, type of missile modification) for each test made. In such cases, it may be possible to construct a linear statistical model relating the main variable, y, to the others, x_1 through x_k. The coefficients of this model can be estimated by a least squares procedure.

The difference between each measured y and the y predicted by the linear model is called a residual. If the set of residuals is normally distributed, several well-known tests of statistical hypotheses and methods of setting confidence intervals are applicable. A procedure for graphically validating the normality of the residuals has also been developed.

1. INTRODUCTION

In the summer of 1968 the Systems Evaluation Branch* had the task of determining which, if any, of three modifications of a certain missile was "best." A modification was considered "best" if the average radial miss distance measured from the center of a target of fixed size was significantly less for the modification than for the other two modifications.

There was no lack of data for this project; in fact, data had been recorded for over 1000 firings of the missile. For each firing, the following had been recorded: radial miss distance (y), target altitude at intercept (v_1), range of the target at launch (v_2), range of the target at intercept (v_3), target closing velocity at intercept (v_4), missile modification (v_5), target type (v_6), and radar power (v_7).

The data were sorted for duplications and missing values. There remained data on over 900 firings. Of these, approximately 6 percent were Mod 1 firings, 15 percent were Mod 2 and 79 percent were Mod 3. For this paper, 100 firings were chosen from the total; 6 of Mod 1, 15 of Mod 2 and 79 of Mod 3. Since the original data were classified, the values were coded or transformed to nonstandard, undefined "units." The coded data are shown in Table I.

The following simple procedure was considered: divide the data into three groups according to modification. Calculate the sample average and sample variance of the radial miss distances for each group. Test these values for equality using the F and t statistical tests. This procedure was rejected for the following reason: the testing procedure was not planned in advance to insure sets of comparable conditions for each modification. For example, most of the firings for Mod 1 were with the second target type ($v_6 = 2$). Thus, if the above test procedure had been used, the effect on the radial miss distance of the modification would have been confounded with the effect of the target type. The conclusions would then be questionable at best.

---

* Advanced Systems Laboratory, Research and Engineering Directorate, U. S. Army Missile Command, Redstone Arsenal, Alabama.

TABLE I. CODED MISSILE DATA, EXAMPLE I

y	v_1	v_2	v_3	v_4	v_5	v_6	v_7
3.8	1.11	27.1	24.3	110.0	3.0	1.0	1.0
4.7	1.88	27.1	20.1	108.2	3.0	1.0	1.0
5.0	2.99	31.3	24.3	187.4	2.0	2.0	2.0
5.0	1.24	24.3	20.1	103.7	2.0	1.0	2.0
5.0	1.67	15.9	13.1	127.1	3.0	1.0	1.0
5.0	1.55	29.9	24.3	101.0	3.0	1.0	2.0
5.6	2.32	45.3	35.5	160.4	3.0	2.0	1.0
6.2	6.06	42.5	27.1	191.0	3.0	2.0	2.0
6.8	6.06	36.9	27.1	204.5	3.0	2.0	2.0
7.1	7.05	46.7	32.7	174.8	2.0	2.0	2.0
7.4	1.55	32.7	27.1	108.2	3.0	1.0	1.0
7.4	1.33	29.9	25.7	110.0	3.0	1.0	2.0
7.7	1.73	34.1	25.7	128.0	1.0	2.0	2.0
7.7	3.64	31.3	25.7	123.5	3.0	1.0	1.0
8.0	2.34	45.3	35.5	112.7	2.0	1.0	2.0
8.0	2.32	29.9	21.5	108.2	3.0	1.0	2.0
8.3	1.11	31.3	25.7	107.3	3.0	1.0	1.0
8.3	10.02	43.9	31.3	191.9	3.0	2.0	1.0
8.9	2.21	24.3	20.1	123.5	3.0	1.0	1.0
9.2	1.66	38.3	29.9	108.2	2.0	1.0	2.0
9.2	1.88	28.5	20.1	110.0	3.0	1.0	1.0
9.2	1.22	28.5	22.9	108.2	3.1	1.0	2.0
9.5	2.87	31.3	25.7	108.2	3.0	1.0	1.0
10.1	1.34	28.5	21.5	182.0	3.0	2.0	1.0
10.1	1.11	29.9	20.1	114.5	3.0	1.0	1.0
10.1	1.67	35.5	25.7	188.3	3.0	2.0	2.0
10.4	6.02	29.9	24.3	123.5	2.0	1.0	2.0
10.7	1.88	22.9	17.3	174.8	1.0	2.0	2.0
10.7	1.66	28.5	24.3	110.0	3.0	1.0	2.0
11.0	1.55	24.3	20.1	108.2	3.0	1.0	2.0
11.0	1.11	38.3	32.7	114.5	3.0	1.0	1.0
11.3	1.55	32.7	27.1	108.2	3.0	1.0	1.0
11.6	2.32	25.7	21.5	110.0	3.0	1.0	2.0
11.9	2.98	24.3	21.5	107.3	3.0	1.0	1.0
11.9	1.33	29.9	22.9	108.2	3.0	1.0	2.0
12.2	1.24	35.5	28.5	108.2	2.0	1.0	2.0
12.2	2.21	27.1	20.1	98.3	3.0	1.0	2.0
12.5	1.66	28.5	24.3	110.9	3.0	1.0	1.0
12.8	1.88	32.7	25.7	114.5	3.0	1.0	1.0
13.1	1.50	31.3	25.7	107.3	2.0	1.0	2.0

TABLE I. CODED MISSILE DATA, EXAMPLE I (Continued)

y	v_1	v_2	v_3	v_4	v_5	v_6	v_7
13.1	2.98	27.1	21.5	95.6	3.0	1.0	1.0
13.4	2.10	45.3	32.7	164.0	3.0	2.0	2.0
13.7	1.67	50.9	36.9	161.3	1.0	2.0	2.0
13.7	2.65	25.7	22.9	121.7	3.0	1.0	1.0
14.0	2.98	34.1	29.9	105.5	2.0	1.0	2.0
14.0	1.11	28.5	21.5	123.5	3.0	1.0	1.0
14.3	1.33	25.7	21.5	117.2	3.0	1.0	2.0
14.9	6.06	49.5	29.9	107.3	3.0	2.0	2.0
14.9	1.88	17.3	14.5	122.6	3.0	1.0	1.0
15.5	1.67	41.1	28.5	209.0	2.0	2.0	2.0
15.5	1.77	31.3	27.1	108.2	3.0	1.0	1.0
16.1	3.42	28.5	18.7	144.2	3.0	1.0	1.0
16.1	4.30	49.5	35.5	211.7	3.0	2.0	1.0
16.7	1.88	28.5	20.1	117.2	3.0	1.0	1.0
17.0	1.55	27.1	22.9	101.0	3.0	1.0	2.0
17.6	3.64	49.5	32.7	200.9	2.0	2.0	2.0
17.6	2.10	25.7	22.9	108.2	3.0	1.0	1.0
17.9	2.00	25.7	21.5	81.2	3.0	1.0	2.0
18.2	1.67	28.5	24.3	85.7	3.0	1.0	2.0
18.5	1.22	28.5	24.3	108.2	3.0	1.0	1.0
18.5	1.55	29.9	24.3	108.2	3.0	1.0	1.0
18.8	1.33	35.5	25.7	198.2	3.0	2.0	2.0
19.4	2.65	28.5	18.7	121.7	3.0	1.0	1.0
19.4	1.55	28.5	24.3	110.0	3.0	1.0	2.0
20.0	1.99	24.3	14.5	316.1	1.0	2.0	2.0
20.3	1.77	21.5	0.5	108.2	3.0	1.0	1.0
20.3	5.95	31.3	25.7	114.5	3.0	1.0	2.0
20.6	1.25	22.9	18.7	112.7	2.0	1.0	2.0
21.2	2.65	21.5	17.3	101.0	3.0	1.0	1.0
21.2	1.33	28.5	25.7	108.2	3.0	1.0	2.0

TABLE I. CODED MISSILE DATA, EXAMPLE I (Concluded)

y	v_1	v_2	v_3	v_4	v_5	v_6	v_7
21.8	2.32	31.3	25.7	96.5	3.0	1.0	1.0
22.1	5.40	34.1	27.1	114.5	3.0	2.0	1.0
22.4	1.77	31.3	25.7	108.2	3.0	1.0	1.0
23.0	1.11	28.5	21.5	114.5	3.0	1.0	1.0
23.3	2.32	35.5	28.5	115.4	3.0	1.0	2.0
23.9	1.11	27.1	18.7	110.9	3.0	1.0	1.0
23.9	1.88	20.1	17.3	103.7	3.0	1.0	1.0
24.2	1.34	22.9	18.7	120.8	2.0	1.0	2.0
24.2	1.99	25.7	18.7	114.5	3.0	1.0	1.0
24.8	2.98	24.3	18.7	107.3	3.0	1.0	1.0
24.8	1.22	29.9	25.7	101.0	3.0	1.0	2.0
26.9	1.99	34.1	25.7	183.8	3.0	2.0	1.0
28.4	4.08	22.9	17.3	137.0	3.0	1.0	1.0
29.3	1.11	28.5	24.3	107.3	3.0	1.0	1.0
30.2	2.00	45.3	31.3	181.1	1.0	2.0	2.0
30.2	5.07	42.5	28.5	225.2	3.0	2.0	1.0
31.1	1.55	27.1	21.5	108.2	3.0	1.0	2.0
31.4	1.29	34.1	28.5	101.9	2.0	1.0	2.0
32.3	1.77	32.7	27.1	108.2	3.0	1.0	2.0
34.1	1.33	43.9	29.9	210.8	3.0	2.0	2.0
35.0	3.20	43.9	31.3	184.7	3.0	2.0	1.0
37.1	1.55	21.5	18.7	110.0	3.0	1.0	2.0
38.6	1.88	20.1	17.3	101.0	3.0	1.0	2.0
41.3	1.88	28.5	24.3	119.0	3.0	1.0	1.0
41.6	1.11	46.7	27.1	374.6	1.0	2.0	2.0
45.8	5.73	32.7	24.3	141.5	3.0	2.0	1.0
48.5	1.99	32.7	28.5	110.0	3.0	1.0	2.0
57.5	1.13	56.5	31.3	386.3	3.0	2.0	2.0
66.5	1.22	24.3	18.7	174.8	3.0	2.0	2.0
69.5	6.02	27.1	22.9	108.2	2.0	1.0	2.0

2. THE LINEAR STATISTICAL MODEL FOR THIS MISSILE

It was decided to set up a linear statistical model [1] relating the radial miss distance, y, to functions of the 7 other variables, v_1 through v_7. Through engineering considerations, the model chosen was:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 + b_8x_8 + b_9x_9 + b_{10}x_{10} + e$$

$x_1 = v_1$ = target altitude at intercept

$x_2 = v_2$ = range of target at launch

$x_3 = x_2^2 = v_2^2$

$x_4 = v_3$ = range of target at intercept

$x_5 = x_4^2 = v_3^2$

$x_6 = v_4$ = target closing velocity at intercept

$$x_7 = \begin{cases} -0.5 \text{ if Mod 1} \\ 0.5 \text{ if Mod 2} \\ 0.0 \text{ if Mod 3} \end{cases}$$

$$x_8 = \begin{cases} 0.0 \text{ if Mod 1} \\ -0.5 \text{ if Mod 2} \\ 0.5 \text{ if Mod 3} \end{cases}$$

$$x_9 = \begin{cases} -0.5 \text{ if target type 1} \\ 0.5 \text{ if target type 2} \end{cases}$$

$$x_{10} = \begin{cases} -0.5 \text{ if low intensity radar} \\ 0.5 \text{ if high intensity radar} \end{cases}$$

e = random error

3. GENERAL LINEAR STATISTICAL MODELS

Frequently the results of experiments or measurements are given as a set of independent variables and as associated result or dependent variable. The data discussed above provides one example. As another example, the velocity of the vehicle could be measured at various time points.

The result or observation, y, is considered as a function of the independent variables, $v_1, v_2, \ldots, v_m$, and random noise e and written:

$$y = y\ (v_1, v_2, \ldots v_m, e)\ .$$

The observation noise or measurement noise e is a result of the inaccuracy of the measuring devices and of variables which are not included in the model but which do affect the observation. If the model is correct, e is the random fluctuation of y for the fixed values of v_1 through v_m.

The most convenient and frequently used model is the linear statistical model:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k + e\ .$$

Here the x_i's are functions of the basic variables v_i, e.g., $x_1 = v_1$, $x_2 = v_1^2$, $x_3 = v_2$, $x_4 = v_1 v_2$. One restriction on the x_i's is that they be linearly independent; e.g., if $x_1 = v_1$ and $x_2 = v_2$, then x_3 cannot be set to $(v_1 + v_2)$. The other restriction is that the x_i's be known or measured without error. (Both restrictions can be relaxed in more advanced work.) The model is termed "linear" because it is linear in the coefficients b_i. The b_i's are considered to be fixed but unknown and must be estimated from the data.

It should be noted that this is not the only statistical model possible and may not apply in some cases. However, it can be used successfully in a large number of situations and it does possess manipulative ease. The model should be constructed from physical and engineering considerations. As will be seen later, statistical tests can be used to determine which terms can be dropped from the model without seriously affecting the accuracy; however, they give no indication of which new terms should be added to the model.

The results on the missile discussed above, hereafter known as Example I, are of concern here. However, in order to illustrate the method with a small, uncomplicated case, a simple example (Example II) was concocted. The calculations of Example II can be done by hand in a "reasonable" (compared to Example I) length of time.

In Example II, the amount of catalyst added to each of two vats in a chemical plant was varied from 0 to 5 units. The resulting yields are listed in Table II.

TABLE II. CHEMICAL YIELD, EXAMPLE II

v_1 Amount of Catalyst	y	
	Yield for Vat 1	Yield for Vat 2
0	8.81	7.02
1	10.00	10.02
2	13.25	10.15
3	14.51	13.43
4	11.36	10.40
5	8.58	4.33

The yields are plotted in Figure 1 as functions of the amount of catalyst.

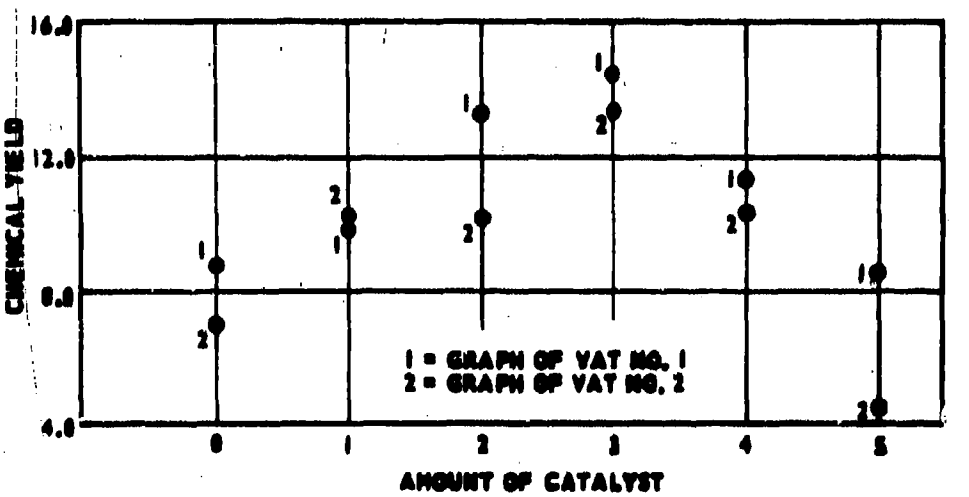


FIGURE 1. CHEMICAL YIELD VERSUS AMOUNT OF CATALYST, EXAMPLE II

The plot suggests that an appropriate model would be a second degree polynomial:

$$y = b_0 + b_1 v_1 + b_2 v_1^2 + e$$
$$= b_0 + b_1 x_1 + b_2 x_2 + e ,$$

where

$$x_1 = v_1, \; x_2 = v_1^2 .$$

Without loss of generality, it can be assumed that the noise e has zero mean [if not, the mean, E(e), could be included in the term b_0 so that the redefined noise $e^* = e - E(e)$ has zero mean].

If a total of n observations are taken, then the model can be written as:

$$y_j = b_0 x_0 + b_1 x_{1,j} + b_2 x_{2,j} + \ldots + b_k x_{k,j} + e_j, \quad j = 1, \ldots, n$$

where $x_0 = 1$ and $x_{i,j}$ is the value of x_i for the j^{th} data point. To shorten the above equations, the following vectors and matrix are defined:

$$\underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{pmatrix}, \quad \underline{b} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ \cdot \\ b_k \end{pmatrix}, \quad \underline{e} = \begin{pmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ e_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_{1,1} & x_{2,1} & \ldots & x_{k,1} \\ 1 & x_{1,2} & x_{2,2} & \ldots & x_{k,2} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ 1 & x_{1,n} & x_{2,n} & \ldots & x_{k,n} \end{pmatrix}.$$

The above equation becomes

$$\underline{y} = X\,\underline{b} + \underline{e}\ .$$

4. ESTIMATION OF THE COEFFICIENTS

A linear statistical model has been postulated in Section 2. In addition, the noise $\underline{e}$ is assumed to have zero mean and covariance matrix $= \sigma^2 I$, where I is the identity matrix and σ^2 is a constant that may be unknown. That is, for i, j = 1, ..., n, $E(e_j) = 0$, var $(e_j) = \sigma^2$ and cov $(e_i, e_j) = 0$ if $i \neq j$. If this assumption is not met, the proper transformation of variables, in most cases, will reduce the model to one in which the assumption does hold.

A method must be found for determining $\hat{\underline{b}}$, the estimate of the coefficients $\underline{b}$. There is usually a loss incurred when the estimate $\hat{\underline{b}}$ is not the true value $\underline{b}$. Usually, the further $\hat{\underline{b}}$ lies from the true value $\underline{b}$, the greater the loss becomes. Since the values of X and $\underline{y}$ are given, it is desirable to choose $\hat{\underline{b}}$ so that the predicted value of $\underline{y}$, $\hat{\underline{y}} = X\hat{\underline{b}}$ will be close, in some sense, to the actual observation vector $\underline{y}$. A convenient way of doing this is to choose $\hat{\underline{b}}$ so that the quadratic loss

$$\text{Loss} = (\underline{y} - X\,\hat{\underline{b}})^T\,(\underline{y} - X\,\hat{\underline{b}}) ,$$

where $(\underline{y} - X\,\hat{\underline{b}})^T$ is the transpose of $(\underline{y} - X\,\hat{\underline{b}})$, is minimized. This is equivalent to

$$\text{Loss} = \sum_{j=1}^{n} (y_j - \overline{x}_j\,\hat{\underline{b}})^2 ,$$

where $\overline{X}_j$ is the row vector

$$\overline{x}_j = \left(1,\ x_{1,j},\ x_{2,j},\ \ldots,\ x_{k,j}\right) .$$

Because of the above form, the estimate $\hat{\underline{b}}$ which minimizes the quadratic loss is called the <u>least squares estimator</u>.

The quadratic loss can be expanded

$$(\underline{y} - X\,\hat{\underline{b}})^T\,(\underline{y} - X\,\hat{\underline{b}}) = \underline{y}^T\,\underline{y} - 2\,\hat{\underline{b}}^T X^T \underline{y} + \hat{\underline{b}}^T X^T X\,\hat{\underline{b}} .$$

If this quantity is differentiated by $\hat{\underline{b}}$ and set equal to the zero vector, the result is

$$-2\,X^T\underline{y} + 2\,X^T X\,\hat{\underline{b}} = \underline{0}$$

$$X^T X\,\hat{\underline{b}} = X^T\underline{y}$$

$$\hat{\underline{b}} = (X^T X)^{-1} X^T \underline{y} .$$

In Example II, the quantities of interest are

$$\underline{y} = \begin{bmatrix} 8.81 \\ 10.00 \\ 13.25 \\ 14.51 \\ 11.36 \\ 8.58 \\ 7.02 \\ 10.02 \\ 10.15 \\ 13.43 \\ 10.40 \\ 4.33 \end{bmatrix}, \quad X = \begin{bmatrix} 1.0 & 0.0 & 0.0 \\ 1.0 & 1.0 & 1.0 \\ 1.0 & 2.0 & 4.0 \\ 1.0 & 3.0 & 9.0 \\ 1.0 & 4.0 & 16.0 \\ 1.0 & 5.0 & 25.0 \\ 1.0 & 0.0 & 0.0 \\ 1.0 & 1.0 & 1.0 \\ 1.0 & 2.0 & 4.0 \\ 1.0 & 3.0 & 9.0 \\ 1.0 & 4.0 & 16.0 \\ 1.0 & 5.0 & 25.0 \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} .$$

The quantities $X^T\underline{y}$ and X^TX are calculated to be

$$X^T\underline{y} = \begin{pmatrix} 121.86 \\ 302.23 \\ 1035.99 \end{pmatrix}, \quad X^TX = \begin{pmatrix} 12.0 & 30.0 & 110.0 \\ 30.0 & 110.0 & 450.0 \\ 110.0 & 450.0 & 1958.0 \end{pmatrix} .$$

The inverse of X^TX is

$$(X^TX)^{-1} = \begin{pmatrix} 0.4107 & -0.2946 & 0.0446 \\ -0.2946 & 0.3634 & -0.0670 \\ 0.0446 & -0.0670 & 0.0134 \end{pmatrix} .$$

The estimate of $\underline{b}$ is

$$\hat{\underline{b}} = \left(X^TX\right)^{-1} X^T\underline{y} = \begin{pmatrix} 7.249 \\ 4.549 \\ -0.924 \end{pmatrix} .$$

The prediction equation for y is thus

$$\hat{y} = 7.249 + 4.549\, x_1 - 0.924\, x_2$$
$$\hat{y} = 7.249 + 4.549\, v_1 - 0.924\, v_1^2 .$$

Listed below are $\underline{y}$, $\hat{\underline{y}}$, and the error in the prediction of $\underline{y}$, $\underline{y}-\hat{\underline{y}}$.

$\underline{y}$	$\hat{\underline{y}} = X\hat{\underline{b}}$	$\underline{y}-\hat{\underline{y}}$
8.81	7.249	1.561
10.00	10.874	-0.874
13.25	12.652	0.598
14.51	12.582	1.927
11.36	10.667	0.693
8.58	6.904	1.676
7.02	7.249	-0.229
10.02	10.874	-0.854
10.15	12.652	-2.502
13.43	12.582	0.848
10.40	10.667	-0.267
4.33	6.904	-2.574

Plotted in Figure 2 are the data points and the prediction equation.

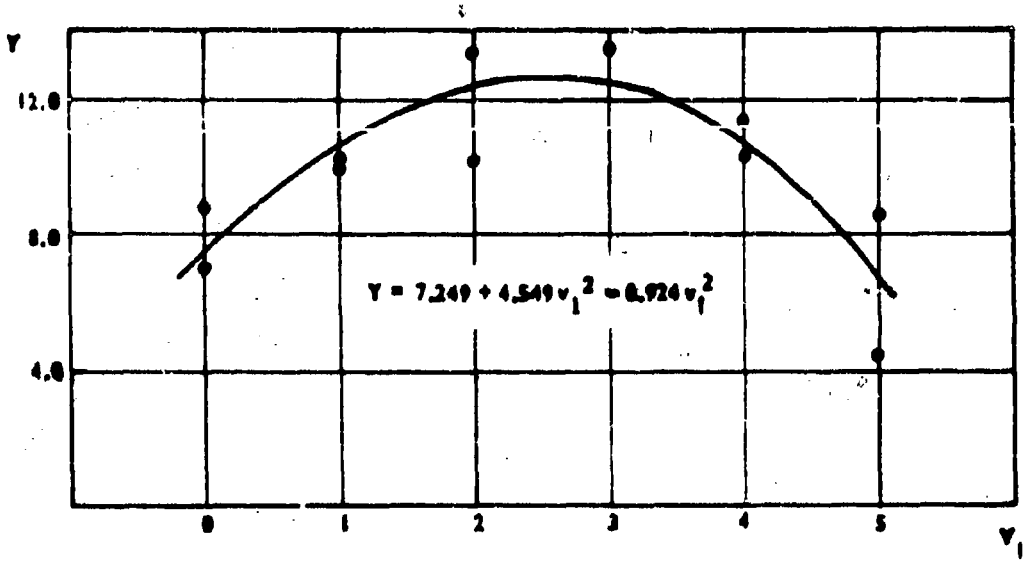


FIGURE 2. PREDICTED YIELD, EXAMPLE II

Notice that no mention has been made of the probability distribution of the measurement noise $\underline{e}$ except that the covariance matrix is $\sigma^2 I$ and the mean is the zero vector. Thus, the formula for the least squares estimator is free of the distribution of $\underline{e}$. Also, no matter what the distribution of $\underline{e}$, if $E(\underline{e}) = \underline{0}$ and cov $(\underline{e}) = \sigma^2 I$, then

$$E(\underline{y}) = E(X\underline{b} + \underline{e}) = X\underline{b}$$

$$\text{cov}(\underline{y}) = \text{cov}(X\underline{b} + \underline{e}) = \sigma^2 I$$

$$E(\hat{\underline{b}}) = E\left([X^T X]^{-1} X^T \underline{y}\right) = [X^T X]^{-1} X^T E(\underline{y}) = \underline{b}$$

$$\text{cov}(\hat{\underline{b}}) = [X^T X]^{-1} X^T \text{cov}\, \underline{y}\, X\, [X^T X] = \sigma^2 [X^T X]^{-1} .$$

Thus, no matter what the distribution, the least squares estimator is unbiased $[E(\hat{\underline{b}}) = \underline{b}]$ and has covariance matrix cov $(\hat{\underline{b}}) = \sigma^2 (X^T X)^{-1}$.

An appealing estimate of the variance σ^2 is the "average" loss. After the coefficient vector $\underline{b}$ has been estimated by $\hat{\underline{b}}$, the predicted value of the dependent variable at the j^{th} point is $\hat{y}_j = \overline{x}_j \hat{\underline{b}}$. The difference between the actual or measured value of y_j and the predicted $\hat{y}_j$ is called the j^{th} residual, $r_j = \left(y_j - \overline{x}_j \hat{\underline{b}}\right)$. The sum of the squares of the residuals is called the sum of squares for error (SSE) and can be shown to equal:

$$SSE = \sum_{j=1}^{n} r_j^2 = \sum_{j=1}^{n} \left(y_j - \tilde{x}_j \hat{b} \right)^2 = (\underline{y} - X \hat{\underline{b}})^T (\underline{y} - X \hat{\underline{b}}) .$$

One estimate of the variance σ^2 is then

$$s^2 = \frac{SSE}{n-k-1} .$$

How well s^2 estimates σ^2 depends upon the forms of the probability distribution of the noise e.

In Example II,

$$SSE = 25.02$$

$$s^2 = 2.78 .$$

The covariance matrix of $\hat{\underline{b}}$ is $(X^T X)^{-1} \sigma^2$ and is estimated by

$$(X^T X)^{-1} s^2 = \begin{pmatrix} 1.14 & -0.82 & 0.12 \\ -0.82 & 1.01 & -0.19 \\ 0.12 & -0.19 & 0.037 \end{pmatrix} .$$

5. TWO TYPES OF INDEPENDENT VARIABLES — QUANTITATIVE AND QUALITATIVE

For the missile model (Example I), y is the dependent variable, v_1 through v_7 are the basic independent variables; x_1 through x_{10} are the expanded variables. The expanded variables x_1 through x_6 are quantitative variables and x_7 through x_{10} are qualitative variables.

A quantitative variable is one to which such units as meters, degrees, and pounds can be attached. The quantitative variables include velocity, time, angle measurement, distance and amount.

The other kind of variable is the assigned or qualitative variable which represents such things as missile modification, type of stimuli, which of several measuring devices were used to obtain the data, etc. These variables must be assigned values and cannot logically be given units. Certain conventions for the assigning of values have been set up for this paper.

In Example II, the model could be expanded to include a term for the vat used. The expanded model is

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + e \ ,$$

where

$x_1 = v_1$

$x_2 = v_2$

$x_3 = \begin{cases} -0.5 \text{ if the first vat is used} \\ 0.5 \text{ if the second vat is used.} \end{cases}$

The matrix X becomes

$$X = \begin{bmatrix} 1.0 & 0.0 & 0.0 & -0.5 \\ 1.0 & 1.0 & 1.0 & -0.5 \\ 1.0 & 2.0 & 4.0 & -0.5 \\ 1.0 & 3.0 & 9.0 & -0.5 \\ 1.0 & 4.0 & 16.0 & -0.5 \\ 1.0 & 5.0 & 25.0 & -0.5 \\ 1.0 & 0.0 & 0.0 & 0.5 \\ 1.0 & 1.0 & 1.0 & 0.5 \\ 1.0 & 2.0 & 4.0 & 0.5 \\ 1.0 & 3.0 & 9.0 & 0.5 \\ 1.0 & 4.0 & 16.0 & 0.5 \\ 1.0 & 5.0 & 25.0 & 0.5 \end{bmatrix} .$$

The vector $X^T \underline{y}$ and the matrices $X^T X$ and $(X^T X)^{-1}$ are

$$X^T \underline{y} = \begin{pmatrix} 121.86 \\ 302.23 \\ 1035.99 \\ -5.585 \end{pmatrix}, \quad X^T X = \begin{pmatrix} 12.0 & 30.0 & 110.0 & 0.0 \\ 30.0 & 110.0 & 450.0 & 0.0 \\ 110.0 & 450.0 & 1958.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 3.0 \end{pmatrix},$$

$$(X^T X)^{-1} = \begin{pmatrix} 0.4107 & -0.2946 & 0.0446 & 0.0 \\ -0.2946 & 0.3634 & -0.0670 & 0.0 \\ 0.0446 & -0.0670 & 0.0134 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.3333 \end{pmatrix} .$$

The estimate $\hat{\underline{b}}$ is thus

$$\hat{\underline{b}} = \begin{pmatrix} 7.249 \\ 4.549 \\ -0.924 \\ -1.860 \end{pmatrix}.$$

The prediction equation is

$$\hat{y} = 7.249 + 4.549\, x_1 - 0.924\, x_2 - 1.860\, x_3 .$$

The estimate of the variance is

$$s^2 = 1.83 .$$

The covariance matrix and the correlation matrix of $\hat{\underline{b}}$ are given below:

$$\text{cov}\ (\hat{\underline{b}}) = \begin{pmatrix} 0.75 & -0.54 & 0.082 & 0.0 \\ -0.54 & 0.66 & -0.12 & 0.0 \\ 0.082 & -0.12 & 0.24 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.61 \end{pmatrix}$$

$$\text{cor}\ (\hat{\underline{b}}) = \begin{pmatrix} 1.00 & -0.77 & 0.65 & 0.00 \\ -0.77 & 1.00 & -0.96 & 0.00 \\ 0.65 & -0.96 & 1.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 1.00 \end{pmatrix}.$$

In this case, two curves are predicted

$$\hat{y} = 8.179 + 4.549\ v_1 - 0.924\ v_1^2 \text{ if vat 1 is used}$$

$$\hat{y} = 6.319 + 4.549\ v_1 - 0.924\ v_1^2 \text{ if vat 2 is used.}$$

Thus, the difference between the predicted yields from vat 1 and vat 2 with the same amount of catalyst is estimated as $\hat{y}_{\text{vat 2}} - \hat{y}_{\text{vat 1}} = \hat{b}_3 = -1.860$. The two curves are plotted in Figure 3.

This fit may be compared with the preceding fit without the term for vat differences.

In the previous example, there were two vats used and the values of -0.5 and 0.5 were rather arbitrarily assigned to represent the vat used. It is noticeable that the qualitative variable occupies one place in the model and one column in the X matrix; this corresponds to the one difference between two factors.

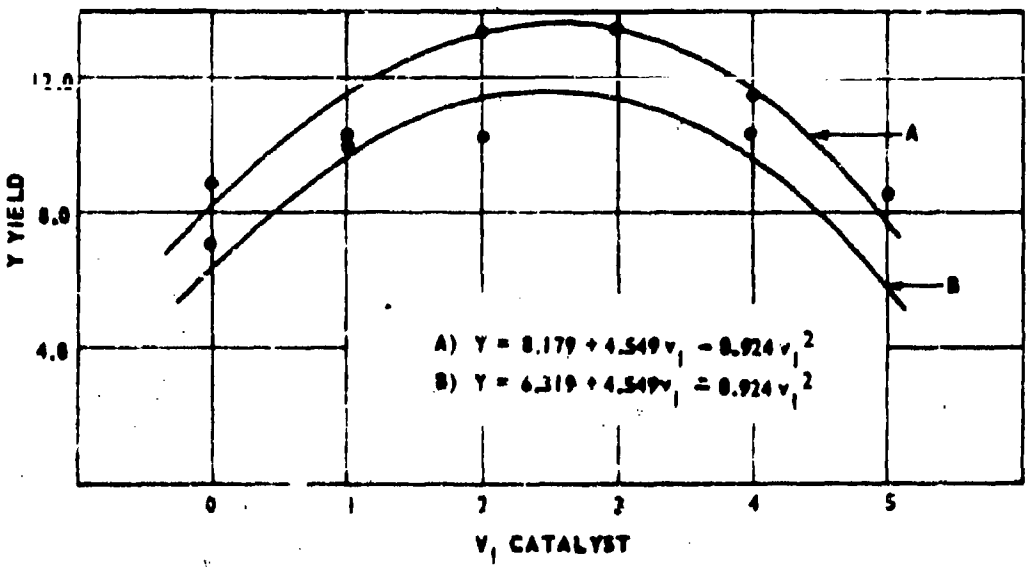


FIGURE 3. PREDICTED YIELD FOR THE EXPANDED MODEL, EXAMPLE II

In Example I, there were three modifications of the missile. There are two linearly independent differences among the three effects, A, B, and C of the modifications. Thus, one could choose B-A and C-B; in this case, C-A is a linear combination of the others, C-A = C-B + B-A. Another choice of linearly independent differences is B-A and 2C-B-A. In this case, two terms are added to the model and two vectors are added to the matrix.

In this work the following values were assigned to the expanded variables for the modifications:

Modification	x_7	x_8
1	-0.5	0.0
2	0.5	-0.5
3	0.0	0.5

If there are four types for a qualitative variable, then there are three independent vectors. They could be assigned the following values:

Type	x_1	x_2	x_3
1	-0.5	0.0	0.0
2	0.5	-0.5	0.0
3	0.0	0.5	-0.5
4	0.0	0.0	0.5

The same pattern is followed for other numbers of types.

6. THE ESTIMATED COEFFICIENTS FOR THE MISSILE MODEL

The data for the missile (Example I) were input into the Generalized Least Squares Fit (GELSF) digital computer program [2], which was written to do the above calculations. The results are shown in Tables III through VI. The predicted model is

$$\hat{y} = 24.103 + 0.086x_1 - 1.496x_2 + 0.020x_3 + 0.666x_4 - 0.016x_5 + 0.087x_6 + 6.719x_7 + 6.413x_8 - 2.061x_9 + 3.753x_{10}.$$

a. The Advantages of Normal Noise

In the special case of Gaussian or normal noise, the least squares estimator is also the maximum likelihood estimator. The likelihood function is the joint probability density of the observations. Assuming X is known perfectly, $\underline{b}$ is fixed but unknown, the noise $\underline{e}$ Gaussian with mean $\underline{0}$ and covariance matrix $\sigma^2 I$, the observations will be Gaussian with mean $X\underline{b}$ and covariance matrix $\sigma^2 I$. Thus, the likelihood function is given by

$$\text{Lh.} = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}(\underline{y} - X\underline{b})^T(\underline{y} - X\underline{b})\right] .$$

If the derivative of the likelihood function with respect to $\underline{b}$ is set to zero, the value $\hat{\underline{b}}$ which maximizes the likelihood is

$$\hat{\underline{b}} = (X^T X)^{-1} X^T \underline{y} .$$

This is identical to the least squares estimator. If, however, the distribution of the noise is other than Gaussian, the likelihood function and, thus, the maximum likelihood estimator, may be different from the least squares estimator.

Furthermore, if the covariance matrix is of the form $Q = \sigma^2 I$ and the noise is Gaussian, then it can be shown that the ratio $\dfrac{(\underline{y} - X\hat{\underline{b}})^T(\underline{y} - X\hat{\underline{b}})}{\sigma^2}$ has a Chi-square distribution with (n-k-1) degrees of freedom where n is the number of data points or observations and k is the number of x_i's in the model, and thus, $\dfrac{(\underline{y} - X\hat{\underline{b}})^T(\underline{y} - X\hat{\underline{b}})}{n-k-1}$ is an unbiased estimator of σ^2. This is the s^2 discussed in Section 4.

TABLE III. THE ESTIMATED COEFFICIENTS, EXAMPLE I

I	BETA(I)	BASIC TYPE	IDENTIFICATION	NUMBER OF VECTORS
0	0.24103171E 02		CONSTANT	
1	0.86253552E-01	1	TARGET ALT.	1
2	-0.14957100E 01	1	RANGE,LAUNCH	2
3	0.19629180E-01			
4	0.66613513E 00	1	RANGE,INTER.	2
5	-0.15850845E-01			
6	0.86580525E-01	1	TAR. CL.VEL.	1
7	0.67192430E 01	2	MISSILE MOD.	2
8	0.64127988E 01			
9	0.20606425E 01	2	TARGET TYPE	1
10	0.37527300E 01	2	RADAR POWER	1

THE SUM OF SQUARES FOR ERROR IS 0.13922442E 05

THE ESTIMATE OF THE VARIANCE IS 0.15643194E 03

THE ESTIMATE OF THE STANDARD DEVIATION IS 0.12507275E 02

THE NUMBER OF DEGREES OF FREEDOM IS 89.

TABLE IV. THE RESIDUALS, EXAMPLE I

THE RESIDUALS (ACTUAL OBSERVATIONS - PREDICTED)

1	0.12[illegible]613E C2	2	-0.121612[illegible]0E C2	3	-0.15827706E C2
4	0.13476972E 02	5	-C.15758877E 02	6	-0.13748894E C2
7	0.[illegible]E752C8E C1	8	-C.17314C40E C2	9	-0.17530940E C2
10	0.11537761E C2	11	-C.7C56C862E 01	12	-0.11932177E 02
13	C.35528126E C1	14	-C.88353692E 01	15	-0.53195298E C1
16	C.116C6226E 02	17	-C.661453[illegible]8E 01	18	-0.11071021E C2
19	C.106773895E C2	20	-C.4729C945E 01	21	-0.72510670E 01
22	C.13747760E 02	23	-C.5644265RE 01	24	-0.10487110E C2
25	C.61851535E 01	26	-C.12793162E C2	27	-0.76293404E C1
28	0.[illegible]C82127E 01	29	-C.93267230E 01	30	-0.10948503E C2
31	C.181446C54E C1	32	-C.31560862E 01	33	-0.98602482E 01
34	C.635062[illegible]8E 01	35	-C.75681361E C1	36	-0.21880586E C1
37	C.38325926E C1	38	-C.38519159E 01	39	-0.24691151E 01
40	0.254772[illegible]3E C1	41	-C.27746752E C1	42	-0.64946657E C1
43	C.2793C987E 00	44	-C.49562793E 01	45	-0.12232212E C0
46	C.35629303E 01	47	-C.76582400E 01	48	-0.28738209E 01
49	C.86656587E 01	50	-C.56360195E C1	51	0.68971837E C0
52	C.33733943E C1	53	-C.47334766E 01	54	-0.37444960E C0
55	C.29185003E C1	56	-C.41027260E 01	57	C.12000253E 00
58	C.103[illegible]1162E 01	59	C.27633111E 00	60	0.24198041E C1
61	C.2[illegible]E04529E 01	62	-C.492C5871E C1	63	0.19410922E 01
64	C.61723503E CC	65	-C.1C70C86[illegible]E C2	66	0.70708198E C1
67	C.5[illegible]8E32[illegible]E C0	68	C.61561C46E 00	69	0.20672664E C1
70	[illegible].15345565E 01	71	C.77161704E 01	72	0.[illegible]852933E C1
73	[illegible].735C6132E C1	74	C.62162[illegible]80E 01	75	C.51422264E 01
76	C.69429379E 01	77	C.36491225E C1	78	C.35065422E C1
79	[illegible].621233[illegible]4E C1	80	C.66303782E 01	81	0.62565392E C1
82	[illegible].7[illegible]402544E 01	83	C.69CC8548E C1	84	0.13307215E C2
85	C.1491180CE 02	86	C.79642540E C1	87	C.10505017E 02
88	[illegible].17[illegible]71762E C2	89	C.14072209E C2	90	0.[illegible]6639019E 01
91	[illegible].16[illegible]40611E C2	92	C.1339648CE C2	93	0.14830160E C2
94	C.24227803E C2	95	C.81107[illegible]40E 01	96	0.29744796E 02
97	C.3[illegible]3[illegible]6627E C2	98	C.12325[illegible]60E 02	99	0.40945884E C2
100	[illegible].51625[illegible]41E C2				

TABLE V. THE ESTIMATED COVARIANCE MATRIX, EXAMPLE I

COLUMNS 1 THROUGH 4 OF THE COVARIANCE MATRIX

0.93542284E 00	-0.36356652E-01	-0.56652870E-03	-0.56552546E-01
-0.36356652E-01	0.15345247E-01	-0.24059465E-01	-0.56836866E 00
-0.56652870E-03	-0.24059465E-01	0.35275723E-03	0.95503378E-02
-0.56552546E-01	-0.56836866E 00	0.95503378E-02	0.12636249E 01
0.27765373E 02	0.45275852E-02	-0.19510789E-03	-0.28402689E-01
0.86776596E-02	0.67653409E-02	-0.27832166E-03	-0.10729626E-01
-0.17877101E 01	-0.14537644E 00	-0.43267103E-02	-0.85554210E 00
-0.31085822E 00	-0.12754077E-01	-0.61321814E-02	-0.10186067E 01
0.16707153CE 01	-0.99729203E 00	0.93364145E-02	0.52244540E 00
0.40464304E 00	0.21969C26E 00	-0.48231689E-02	-0.62695605E 00

COLUMNS 5 THROUGH 8 OF THE COVARIANCE MATRIX

0.27765373E-02	0.86776596E-02	-0.17877101E 01	-0.31085822E 00
0.45275852E-02	0.67653409E-02	-0.14537644E 00	-0.12754077E-01
0.19510789E-03	-0.27832166E-03	-0.43267103E-02	-0.61321814E-02
0.28402689E-01	-0.10729626E-01	-0.85554210E 00	-0.10186067E 01
0.84151427E-03	0.53369615E-03	0.22198053E-01	0.32539614E-01
0.53369615E-03	0.21573451E-02	0.64845563E-01	0.49461851E-01
0.22198053E-01	0.64845563E-01	0.71528088E 02	0.30512465E 02
0.32539614E-01	0.49461851E-01	0.30512465E 02	0.29322505E 02
0.12007027E-01	-0.11804427E 00	0.10935714E 02	0.28468976E 01
0.14573285E-01	0.85635212E-02	0.22797647E 01	0.66327820E 01

TABLE V. THE ESTIMATED COVARIANCE MATRIX, EXAMPLE I (Concluded)

COLUMNS 9 THROUGH 10 OF THE COVARIANCE MATRIX

0.16707530E 01	0.84C4643C4E 0C
0.99729209E 00	0.21969C26E C0
0.93364145E-02	-0.48231689E-02
0.52244540E 00	-0.62695605E 00
0.12007027E-01	0.14573295E-01
0.11304427E 00	0.85635212E-02
0.10935714E 02	0.22797647E 01
0.28468976E 01	0.66327820E 01
0.26121725E 02	-0.14371421E 01
0.14371421E 01	0.8S3CC860E 01

TABLE VI. THE ESTIMATED CORRELATION MATRIX, EXAMPLE I

COLUMNS 1 THROUGH 4 OF THE CORRELATION MATRIX

0.10000000E 01	-0.28555596E-01	-0.33001279E-01	-0.55106910E-01
0.28595596E-01	0.10000000E 01	-0.92243847E 00	-0.36391936E 00
0.33001279E-01	-0.92243847E 00	0.10000000E 01	0.45288490E 00
0.55106910E-01	-0.36391936E 00	0.45288490E 00	0.10000000E 01
0.10472135E 00	0.11222024E 00	-0.35810149E 00	-0.87203876E 00
0.20440505E 00	0.10471233E 00	-0.31904300E 00	-0.20574622E 00
0.23126270E 00	-0.12357303E-01	-0.27238397E-01	-0.90096964E-01
0.62807087E-01	-0.16932320E-02	-0.60294333E-01	-0.16753790E 00
0.35764954E 00	-0.14027754E 00	0.97261560E-01	0.91043105E-01
0.14814639E 00	0.52850727E-01	-0.85934328E-01	-0.18686000E 00

COLUMNS 5 THROUGH 8 OF THE CORRELATION MATRIX

0.10472135E 00	0.20440505E 00	-0.23126270E 00	-0.62807087E-01
0.11222024E 00	0.10471233E 00	-0.12357303E-01	-0.16932320E-02
0.35810149E 00	-0.31904300E 00	-0.27238397E-01	-0.60294333E-01
0.87203876E 00	-0.20574622E 00	-0.90096964E-01	-0.16753790E 00
0.10000000E 01	0.35605876E 00	0.90478593E-01	0.20714181E 00
0.39609876E 00	0.10000000E 01	0.16598550E 00	0.19665727E 00
0.90478593E-01	0.16586550E 00	0.10000000E 01	0.66625189E 00
0.20714181E 00	0.19665727E 00	0.66625189E 00	0.10000000E 01
0.89898485E-01	-0.49726076E 00	0.25299265E 00	0.10286565E 00
0.16811207E 00	0.61697099E-01	0.90203603E-01	0.40989004E 00

TABLE VI. THE ESTIMATED CORRELATION MATRIX, EXAMPLE I (Concluded)

COLUMNS 6 THROUGH 10 OF THE CORRELATION MATRIX

0.357649S4E 00	0.14814639E 00
0.14027794E CC	0.52850727E-01
0.97261S6CE-01	-0.8SC3422BE-01
0.81S431C5E-01	-0.18686CC0E 00
0.9C9849S9E C1	0.168112C7E 00
0.497260C76E 0C	0.616970S9E-01
0.252997C5E CC	0.9C2C36C3E-01
0.102865C5E 0C	0.4CS89CC4E CC
0.190000CCE 01	-0.940955C56E-01
0.940955S6E-01	0.1CCCCCC0E 01

Further, the assumption of normally distributed noise is used in the statistical tests and confidence intervals to be discussed in Section 10.

The data for Example II was generated under the assumption of normal noise. However, there is a question to be asked about the missile data of Example I.

7 TESTING THE MISSILE DATA FOR NORMAL NOISE

Before conclusions based upon the assumption of normally distributed noise can be drawn for Example I, a test for normality must be made. The residuals (actual y - predicted y) estimate the error e and, thus, should be tested for normality.

It was decided to use the method of normal probability paper and control bands to test the residuals for normality. The Testing for Normality by Control Bands (TEN COB) digital computer program was used.

8. NORMAL PROBABILITY PAPER

The construction of normal probability paper is similar to that of logarithmic paper. Assume that the random variable r has a normal (Gaussian) distribution with mean μ and variance σ^2. Then the reduced variate $v = (r - \mu)/\sigma$ has a standard normal distribution, i.e., v has mean 0 and variance 1. If r is plotted on a horizontal linear scale and v is plotted on a vertical linear scale, the straight line $r = \sigma v + \mu$ will result (Figure 4).

Since v is a standard normal random variable, the cumulative distribution function of v is given by

$$F(v) = \int_{-\infty}^{v} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt .$$

On a second vertical scale the distribution function F (v) is plotted (Figure 5). The values of F (v) = 0 and F (v) = 1 never appear on the scale, since these correspond to values of $v = -\infty$ and $v = \infty$, respectively. If the lines are drawn for the function F (v) instead of v, a nonlinear vertical scale is shown (Figure 6).

a. Plotting of Points on Probability Paper

It is desired to test whether the underlying probability distribution of the n residuals is normal or can be approximated by a normal distribution.

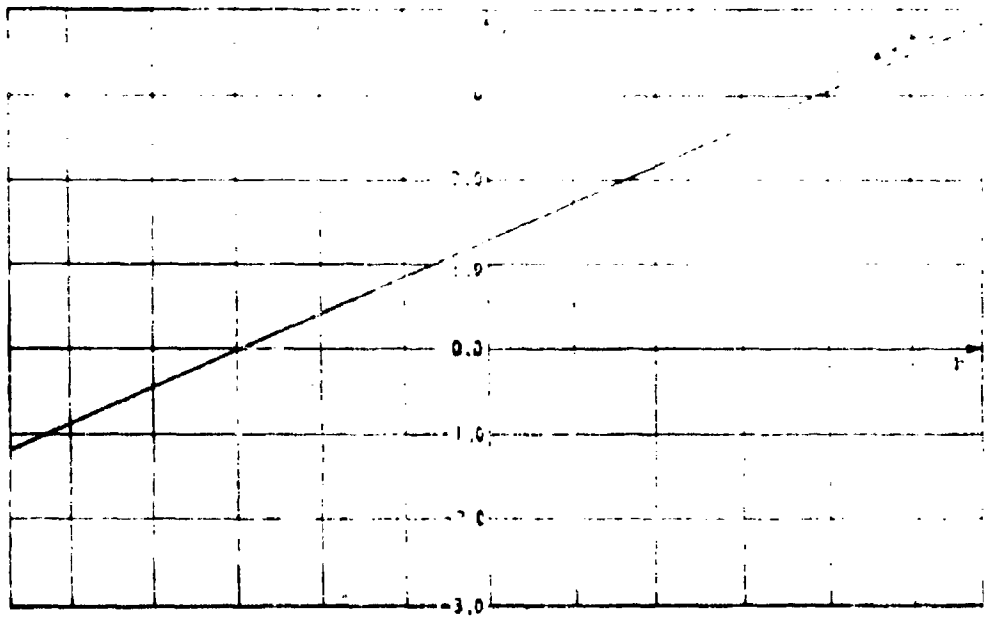


FIGURE 4. THE REDUCED VARIATE

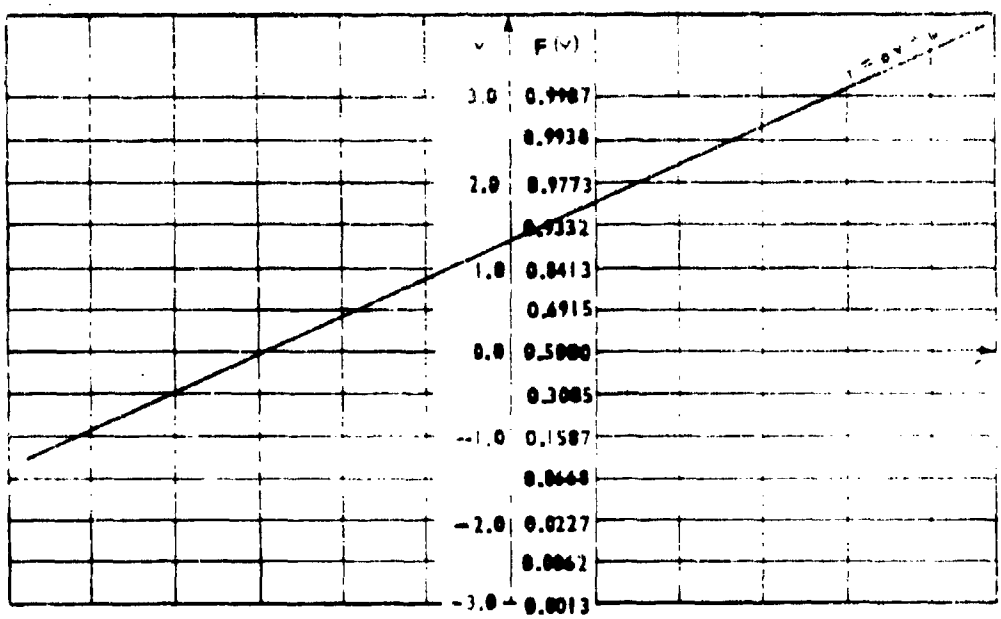


FIGURE 5. REDUCED VARIATE AND DISTRIBUTION FUNCTION

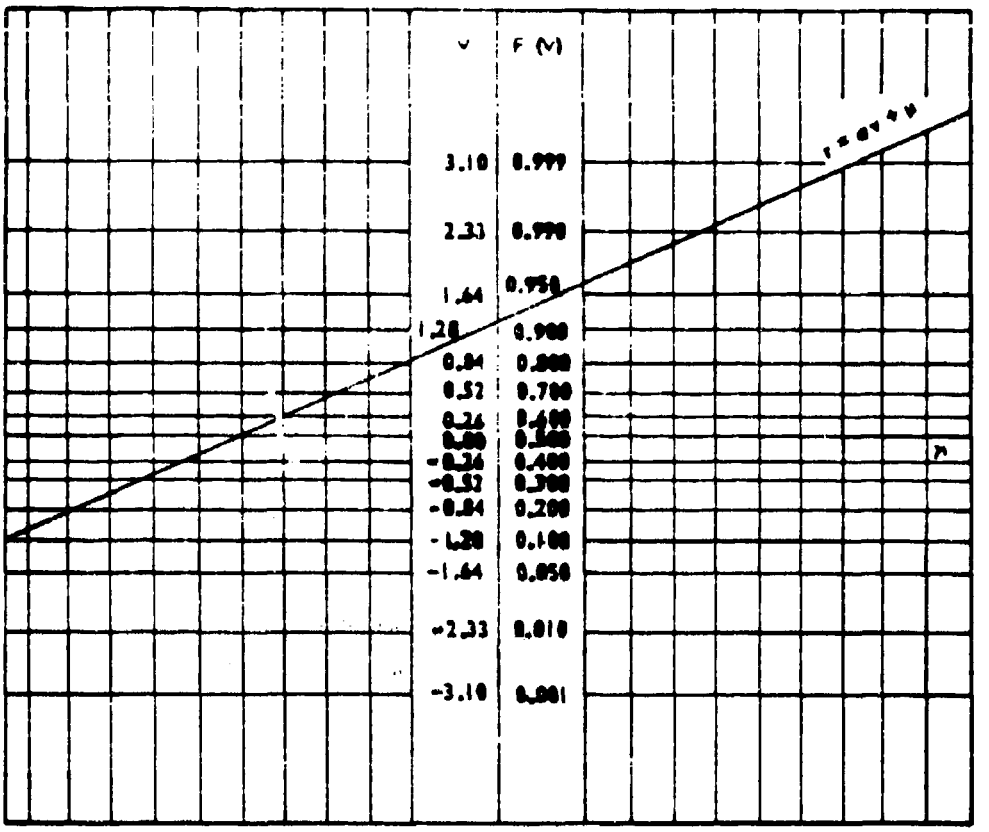


FIGURE 6. EXAMPLE OF NORMAL PROBABILITY PAPER

If the distribution is indeed normal, the residuals, when plotted against F (v) in the manner discussed below, should approximate a straight line, $r = \gamma + v/\delta$. The deviations from the line are caused by the finite random character of the sample of size n and such errors as round-off.

Let $r_1, r_2, \ldots, r_n$ be the n residuals arranged in ascending order. Several methods of plotting this sequence of numbers against F (v) are considered by Gumbel [3]. The best method is that of plotting r_j against $j/(n + 1)$. This method is distribution free and all observations can be plotted. Further, the plotting positions are simple to calculate.

Listed in Table VII are the 100 residuals arranged in ascending order. The corresponding values of $F(v) = j/(n + 1)$ and v are also listed.

The Residuals in Ascending Order	J The Rank	[illegible] Distribution Function Plotting Position	[illegible] Variate Plotting Position
-19.799	1	0.0099	-2.329
-17.531	2	0.0198	-2.05-
-17.314	3	0.0297	-1.885
-15.828	4	0.0396	-1.755
-13.749	5	0.0495	-1.649
-13.479	6	[illegible]	-1.559
-12.993	7	0.0693	-1.4[illegible]
-12.793	8	0.0792	-1.41[illegible]
-12.161	9	0.0891	-1.347
-11.932	10	0.0990	-1.285
-11.606	11	0.1089	-1.232
-11.548	12	0.1188	-1.180
-11.071	13	0.1287	-1.132
-10.949	14	0.1386	-1.086
-10.746	15	0.1485	-1.042
-10.701	16	0.1584	-1.001
-10.677	17	0.1683	-0.960
-10.487	18	0.1782	-0.922
-9.860	19	0.1881	-0.884
-9.328	20	0.1980	-0.848
-9.088	21	0.2079	-0.813
-9.008	22	0.2178	-0.779
-8.835	23	0.2277	-0.746
-8.666	24	0.2376	-0.714
-7.698	25	0.2475	-0.682
-7.629	26	0.2574	-0.651
-7.568	27	0.2673	-0.621
-7.251	28	0.2772	-0.591
-7.065	29	0.2871	-0.561
-6.615	30	0.2970	-0.532

TABLE VII. RESIDUALS OF EXAMPLE I WITH CORRESPONDING REDUCED VARIATE (Continued)

r_j The Residuals in Ascending Order	j The Rank	F (v) = j/101 Distribution Function Plotting Positions	v Reduced Variate Plotting Positions
-6.495	31	0.3069	-0.504
-6.351	32	0.3168	-0.476
-6.185	33	0.3267	-0.448
-5.644	34	0.3366	-0.421
-5.636	35	0.3465	-0.394
-5.319	36	0.3564	-0.367
-4.996	37	0.3663	-0.341
-4.921	38	0.3762	-0.315
-4.733	39	0.3861	-0.289
-4.729	40	0.3960	-0.263
-4.103	41	0.4059	-0.238
-3.852	42	0.4158	-0.212
-3.833	43	0.4257	-0.182
-3.593	44	0.4356	-0.162
-3.563	45	0.4455	-0.136
-3.373	46	0.4554	-0.112
-3.156	47	0.4653	-0.086
-2.919	48	0.4752	-0.062
-2.874	49	0.4851	-0.037
-2.775	50	0.4950	-0.012
-2.548	51	0.5050	0.012
-2.464	52	0.5149	0.036
-2.188	53	0.5248	0.062
-1.814	54	0.5347	0.086
-1.039	55	0.5446	0.112
-0.617	56	0.5545	0.136
-0.374	57	0.5644	0.162
-0.122	58	0.5743	0.187
0.120	59	0.5842	0.212
0.276	60	0.5941	0.238

TABLE VII. RESIDUALS OF EXAMPLE 1 WITH CORRESPONDING REDUCED VARIATE (Continued)

r_j The Residuals in Ascending Order	j The Rank	F (v) = j/101 Distribution Function Plotting Positions	v_j Reduced Variate Plotting Positions
0.279	61	0.6040	0.263
0.592	62	0.6139	0.289
0.616	63	0.6238	0.315
0.690	64	0.6337	0.341
1.535	65	0.6436	0.367
1.941	66	0.6535	0.394
2.067	67	0.6634	0.421
2.420	68	0.6733	0.448
2.880	69	0.6832	0.476
3.507	70	0.6931	0.504
3.649	71	0.7030	0.532
5.142	72	0.7129	0.561
6.212	73	0.7228	0.591
6.216	74	0.7327	0.621
6.254	75	0.7426	0.651
6.630	76	0.7525	0.682
6.901	77	0.7624	0.714
6.943	78	0.7723	0.746
7.071	79	0.7822	0.779
7.351	80	0.7921	0.813
7.716	81	0.8020	0.848
7.864	82	0.8119	0.884
7.940	83	0.8218	0.922
8.111	84	0.8317	0.960
9.085	85	0.8416	1.001
9.664	86	0.8515	1.042
10.505	87	0.8614	1.086
12.326	88	0.8713	1.132
13.307	89	0.8812	1.180
13.396	90	0.8911	1.232

TABLE VII. RESIDUALS OF EXAMPLE I WITH CORRESPONDING REDUCED VARIATE (Concluded)

r_j The Residuals in Ascending Order	j The Rank	F (v) = j/101 Distribution Function Plotting Positions	v Reduced Variate Plotting Positions
14.072	91	0.9010	1.285
14.830	92	0.9109	1.347
14.912	93	0.9208	1.410
16.841	94	0.9307	1.480
17.372	95	0.9406	1.559
24.228	96	0.9505	1.649
29.745	97	0.9604	1.755
30.698	98	0.9703	1.885
40.946	99	0.9802	2.058
51.625	100	0.9901	2.329

Figure 7 shows the points plotted on normal probability paper by a modified version of the TEN COB program. The horizontal scale is the r-scale; the horizontal line defining the grid extends from -37.655 to 69.981 for the case. The vertical scale is the reduced variate or v-scale; the vertical line defining the grid extends from v = -4.0 to v = 4.0.

b. Fitting the Straight Line

If the scatter of the plotted points is very small, the best fitting straight line can be found by lining up a ruler through the points. However, in many cases this is not satisfactory. It is then necessary to estimate δ and γ, the two parameters of the straight line;

$$r = \gamma + \frac{v}{\delta}; \quad v = \delta\ (r - \gamma)\ .$$

For the estimation of these parameters, the classical method of least squares will be employed.

In the method of least squares, either the sum of squares of the horizontal deviations of the points from the estimated straight line,

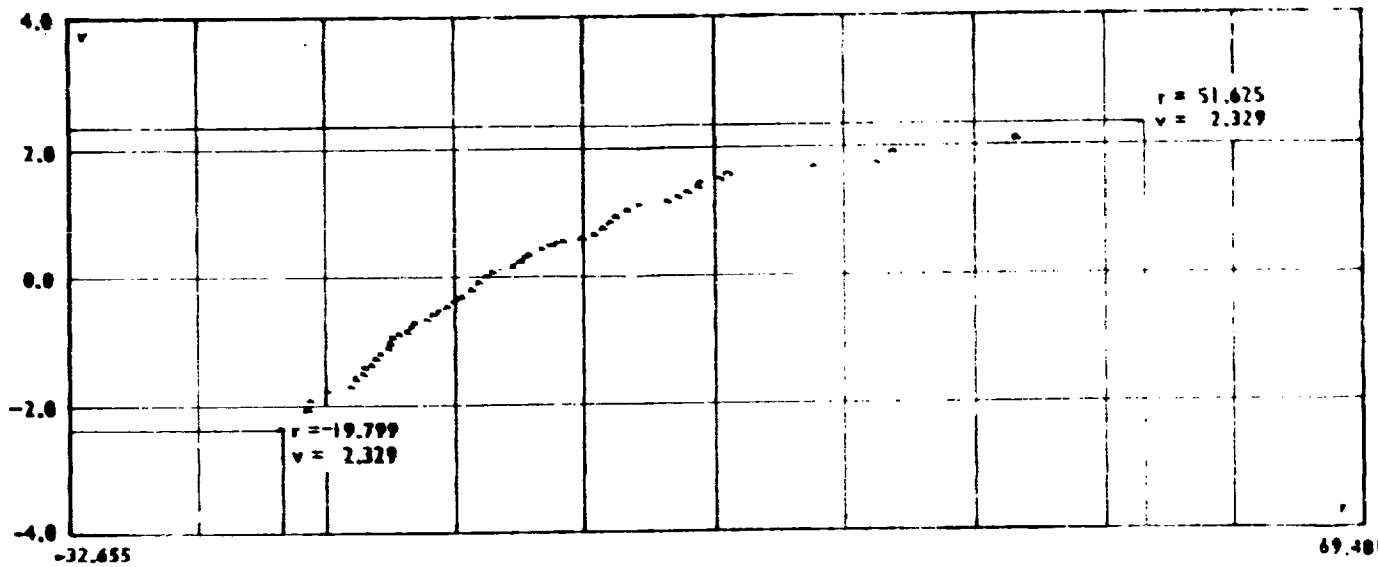


FIGURE 7. RESIDUALS OF EXAMPLE I PLOTTED ON NORMAL PROBABILITY PAPER

$$\sum_{i=1}^{n} \left(r_i - g_1 - \frac{v_i}{a_1} \right)^2 ,$$

or that of the vertical distances,

$$\sum_{i=1}^{n} \left[v_i - a_2 \left(r_i - g_2 \right) \right]^2 ,$$

is minimized. Here a_1 and a_2 represent the two estimates of δ, and g_1 and g_2 represent the two estimates of γ. If the partial derivatives of the first sum with respect to a_1 and g_1 are set to zero, the results are

$$-2 \sum_{i=1}^{n} \left(r_i - g_1 - \frac{1}{a_1} v_i \right) = 0$$

and

$$\frac{2}{a_1^2} \sum_{i=1}^{n} v_i \left(r_i - g_1 - \frac{1}{a_1} v_i \right) = 0$$

or

$$\overline{r} - g_1 - \frac{1}{a_1} \overline{v} = 0 ,$$

$$\overline{rv} - g_1 \overline{v} - \frac{1}{a_1} \overline{v^2} = 0 ,$$

where

$$\overline{r} = \frac{1}{n} \sum_{i=1}^{n} r_i , \quad \overline{v} = \frac{1}{n} \sum_{i=1}^{n} v_i ,$$

$$\overline{rv} = \frac{1}{n} \sum_{i=1}^{n} r_i v_i , \quad \overline{v^2} = \frac{1}{n} \sum_{i=1}^{n} v_i^2 .$$

For a normal distribution, $\overline{v} = 0$. Thus, the solution is

$$\frac{1}{a_2} = \frac{\overline{rv}}{s_v^2}$$

$$g_2 = \bar{r} ,$$

where

$$s_v^2 = \overline{v^2} - \bar{v}^2 = \overline{v^2} .$$

If the sum of squares of the vertical deviations is to be minimized, the partials with respect to g_2 and a_2 are set to 0,

$$-2 \sum_{i=1}^{n} \left[v_i - a_2 \left(r_i - g_2 \right) \right] = 0$$

and

$$-2 \sum_{i=1}^{n} \left(r_i - g_2 \right) \left[v_i - a_2 \left(r_i - g_2 \right) \right] = 0$$

or

$$\bar{v} - a_2 \bar{r} + a_2 g_2 = 0$$

$$\overline{rv} - a_2 \overline{r^2} + a_2 g_2 \bar{r} = 0 ,$$

where

$$\overline{r^2} = \frac{1}{n} \sum_{i=1}^{n} r_i^2 .$$

With $v = 0$, the solution is

$$\frac{1}{a_2} = \frac{\overline{r^2} - \bar{r}^2}{\overline{rv}}$$

$$g_2 = \bar{r} .$$

The estimate $1/a_2$ is altered slightly to

$$\frac{1}{a_2} = \frac{s_r^2}{\overline{rv}} ,$$

where s_r^2 is the sample variance,

$$s_r^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(r_i - \bar{r} \right)^2 = \frac{(\overline{r^2} - \bar{r}^2)\, n}{n-1} \quad .$$

In order to combine the two estimates of $1/\delta$ and eliminate the cross product $\bar{r}\bar{v}$, the geometric mean of $1/a_1$ and $1/a_2$ is found. Thus, the two combined estimates are

$$\frac{1}{a} = \sqrt{\frac{1}{a_1}\frac{1}{a_2}} = \frac{s_r}{s_v}$$

$$g = \bar{r} \quad .$$

These estimates require the calculation of only $\bar{r}$ and s_r from the sample. The value of s_v depends only upon the number of points n. As the sample size n increases, s_v approaches 1 and the estimate $\bar{r}$ and s_r approach the true values μ and σ; thus, the equation of the straight line approaches $r = \mu + \sigma v$, discussed in Section 8.

The estimated straight line,

$$r = \bar{r} + \frac{s_r}{s_v} v \, ,$$

is plotted on the same paper as the points. If the points lie close to the estimated line, the distribution is considered to be approximately normal. If the scatter is too great with respect to the line, the distribution is considered non-normal. To determine whether the scatter is too great, control bands are placed around the line.

For Example I, the values range from -19.799 to 51.625. The end-points of the graph are -37.655 and 69.481. The average residual, $\bar{r} = 0.00286$; the estimate of $1/\delta$ is $1/a = 12.353$. In Figure 8 are shown the points of Example I plotted along with the estimated straight line, $r = 0.00286 + 12.353\, v$.

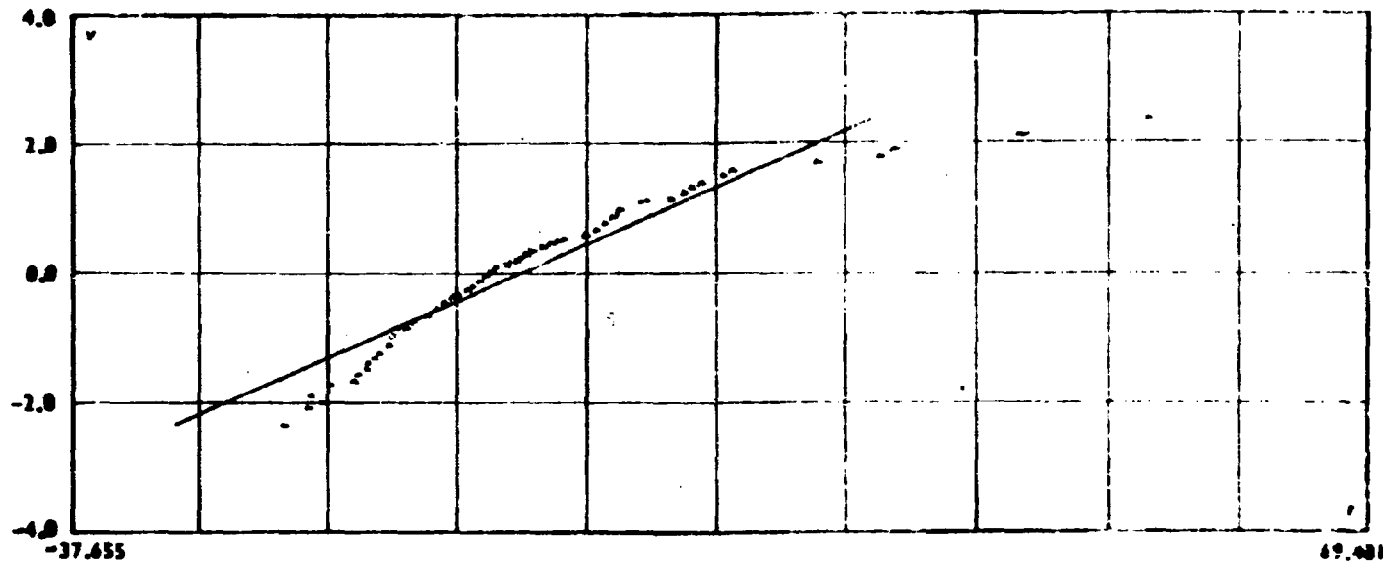


FIGURE 8. RESIDUALS AND BEST FITTING LINE OF EXAMPLE I

c. Control Bands

The j^{th} largest observation r_j in a sample of size n is called the j^{th} order statistic. Each observation was drawn from a population with density function f and distribution function F. In this case, the initial distribution is assumed to be normal. The j^{th} order statistic r_j has a derived density $f_n(r_j)$ that depends upon the initial distribution F and upon the values of j and n. It can be shown that this derived density function is

$$f_n(r_j) = \frac{n!}{(n-j)!\,(j-1)!} F^{j-1}(r_j)\left[1 - F(r_j)\right]^{n-j} f(r_j) .$$

Substituting the equations for a normal variate for $F(r_j)$ and $f(r_j)$, one obtains

$$f_n(r_j) = \frac{n!}{(n-j)!\,(j-1)!} \left\{\int_{-\infty}^{r_j} \frac{\exp\left[-\frac{1}{2\sigma^2}(t-\mu)^2\right]}{\sqrt{2\pi}\,\sigma} dt\right\}^{j-1}$$

$$\cdot \left\{1 - \int_{-\infty}^{r_j} \frac{\exp\left[-\frac{1}{2\sigma^2}(t-\mu)^2\right]}{\sqrt{2\pi}\,\sigma} dt\right\}^{n-j}$$

$$\cdot \left\{\frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2\sigma^2}\left(r_j - \mu\right)^2\right]\right\} .$$

This complicated form does not reduce to anything more reasonable and is difficult to manipulate and calculate. Thus, asymptotic distributions of the order statistics are used.

As n becomes larger, either j will increase with n so that j/n remains approximately constant or j will remain constant so that j/n decreases. In the former case, the j^{th} value r_j is called the j^{th} central value; in the second case, the values r_j and n-j + 1 are called extreme values.

It can be shown that as n increases, the distribution of the j^{th} central value, r_j, becomes asymptotically normal with mean and variance

$$E\, r_j = r_j^*$$

$$\text{var}\left(r_j\right) = \sigma_j^2 = \frac{\left[F\left(r_j^*\right)\right]\left[1 - F\left(r_j^*\right)\right]}{n\, f^2\left(r_j^*\right)} ,$$

where r_j^* is the solution of

$$F\left(r_j^*\right) = \frac{j}{n+1} .$$

The asymptotic distribution is used within the interval $0.15 \leq F \leq 0.85$, although its accuracy within this interval depends upon the sample size.

For the reduced variate $v_j = \alpha\left(r_j - \gamma\right)$, the variance is given by

$$\text{var}\left(v_j\right) = \alpha^2\, \text{var}\left(r_j\right) ,$$

and is independent of the parameters a and g. The product $\sqrt{n \cdot \text{var}\left(v_j\right)}$ is independent of the sample size and is a function of the initial distribution only. A chart of this product for several values of F, assuming that the initial standard deviation $\sigma = 1$, is shown in Table VIII.

TABLE VIII. REDUCED STANDARD ERRORS FOR UNIT STANDARD DEVIATION

Probability, F	$\sqrt{n \cdot \text{var}\left(v_j\right)}$
0.15	1.532
0.20	1.429
0.25	1.363
0.30	1.318
0.35	1.288
0.40	1.268
0.45	1.257
0.50	1.253
0.55	1.257
0.60	1.268
0.65	1.288
0.70	1.318
0.75	1.363
0.80	1.429
0.85	1.532

Because of the relationship $r = \frac{1}{\alpha} v + \gamma$, the standard deviations are related by $\text{s.d.}\left(r_j\right) = \text{s.d.}\left(v_j\right)/\alpha$. This can be written in the form

$$\text{s.d.}\left(r_j\right) = \frac{\sqrt{n \cdot \text{var}\left(v_j\right)}}{\alpha \sqrt{n}} \quad .$$

Thus, the standard error of the j^{th} value, r_j, can be estimated by multiplying the value $\sqrt{n \cdot \text{var}\left(v_j\right)}$ by $\left(\frac{1}{a} \times \frac{1}{\sqrt{n}}\right)$. For example, for the data given in Section 8.a, $n = 100$ and $\frac{1}{a} = 12.353$, so $\text{s.d.}\ (r_{20}) = 1.429\ (12.353)/10 = 1.756$. This standard error is used in the construction of control bands.

Assume that a and g have been estimated and that the observations and the estimated straight line

$$r = g + \frac{1}{a} v$$

have been plotted. For each probability value listed in Table VIII, the estimate of r_j is found by intercepting the estimated line with a horizontal line from the probability value and reading the corresponding r value, $r_{j,\,est}$. The standard error, $\text{s.d.}\left(r_j\right)$, is added and subtracted from the value $r_{j,\,est}$. The points $r_{j,\,est} + \text{s.d.}\left(r_j\right)$ are joined to form one curve; the points $r_{j,\,est} - \text{s.d.}\left(r_j\right)$ are joined to form another curve.

To complete the curves, it is necessary to find the standard error of some of the extreme values. The asymptotic distribution of the extremes is not normal and is, in fact, very complicated. In Table IX are given some values of the reduced standard error for the largest reduced order statistic. By the symmetry of the normal distribution, this is also the standard error for the smallest reduced order statistic.

The values in Table IX are approximated in the TEN COB program by the formula,

$$\text{s.d.}\left(v_n\right) = \text{s.d.}\ (v_1) = 0.71 - 0.061\, ln_e\ (n) \ ,$$

which is fairly accurate for samples of 20 to 500 points. The standard error of the reduced largest (or smallest) value is multiplied by the estimate 1/a to

TABLE IX. STANDARD ERRORS FOR REDUCED LARGEST AND SMALLEST VALUES

Sample Size	s.d. $\left(v_n\right)$ = s.d. (v_1)
20	0.52
25	0.51
30	0.50
40	0.48
50	0.46
75	0.45
100	0.43
200	0.40
500	0.36

obtain s.d. $\left(r_n\right)$ [or s.d. (r_1)]. This value is then added to and subtracted from $r_{n,\,est}$ or $\left(r_{1,\,est}\right)$ to give the points $r_{n,\,est}$ + s.d. $\left(r_n\right)$ and $r_{n,\,est}$ + s.d. $\left(r_n\right)$ or $\left[r_{1,est}\right.$ + s.d. (r_1) and $r_{1,\,est}$ - s.d. $\left.(r_1)\right]$. These points are added to the proper curves to extend the control curves. There is a probability of 0.68 for each j^{th} observation to lie within the band formed by the two control curves. Figure 9 shows the 68-percent control band for Example I.

Multiplication of s.d. $\left(r_j\right)$ by 0.6745, 1.960, 2.576, 2.807, and 3.290 leads to bands corresponding to the probabilities 0.50, 0.95, 0.99, 0.995, and 0.999. In Figure 10 the 95-percent control band has been added. This is the standard graph produced by the TEN COB program.

d. Testing for Normality by Control Bands

The method of control bands used by the TEN COB program gives a graphical criterion for the goodness of fit between the theoretic normal distribution and the observations.

If almost all of the observations fall within the 95-percent band, the underlying distribution can be assumed to be normal for most purposes and statistical tests and confidence intervals that depend upon an underlying normal distribution (such as the Student's t, the Chi-square, and the F tests) can be applied. If almost all of the observations fall within the 68-percent band, more confidence can be placed on the population's being normal. The term "almost all"

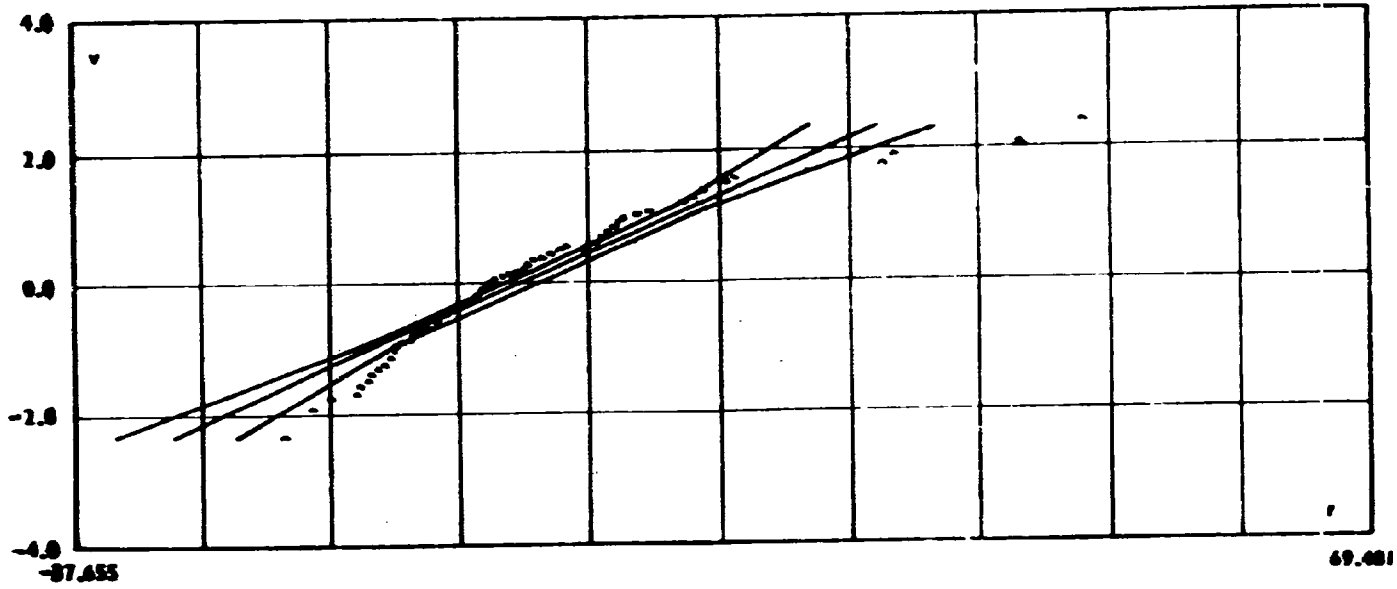


FIGURE 9. RESIDUALS, BEST FITTING LINE, AND 68-PERCENT CONTROL BAND FOR EXAMPLE 1

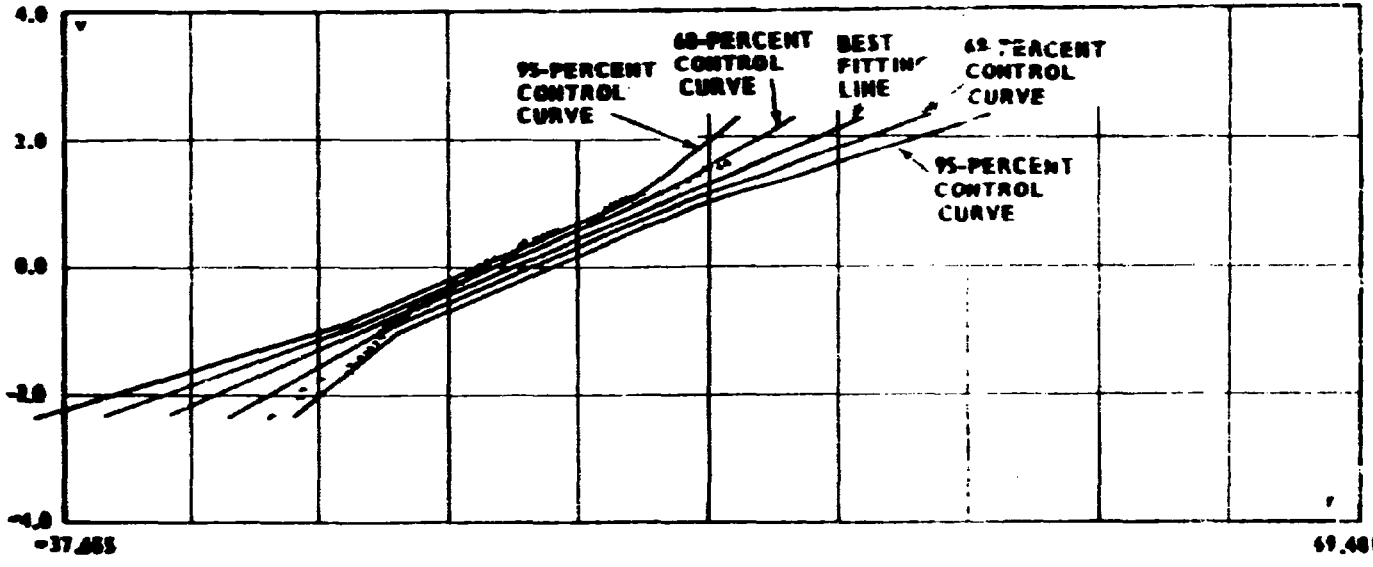


FIGURE 10. RESIDUALS, BEST FITTING LINE, 68-PERCENT CONTROL BAND, AND 95-PERCENT CONTROL BAND FOR EXAMPLE I

is necessarily vague since the degree to which the distribution must match a normal one varies with each set of observations.

In Example I, 67 percent of the residuals fell outside the 68-percent control curve and 18 percent fell outside the 95-percent control band. Thus, the data are probably not normally distributed.

When the observations are shown to be not normally distributed, one of two methods can be employed. The data can be tested against other types of distributions, such as the negative exponential, the log normal, etc. There are disadvantages to this approach. For one, the list of distributions to be tried is long. Further, even if a distribution is found that will approximate that of the observation, it would not have as many well-known associated tests and procedures as the normal distribution.

The other approach is to find a transformation of the observations that will result in normally distributed transformed observations. A list of suggested transformations can be found in Snedecor and Cochran [5].

9. REVISED MODEL FOR THE MISSILE DATA

Since the residuals for the radial miss distance of Example I were not normally distributed, it was decided to try transforming the miss distances to obtain normality of the residuals. The transformation $z = \ln_e y$ was made and the data were used in the GELSF program. The resulting calculation are shown in Tables X through XII. The prediction equation was:

$$\ln y = 2.758 - 0.00947\, x_1 - 0.0363\, x_2 + 0.000632\, x_3 + 0.0165\, x_4 - 0.000726\, x_5 + 0.00301\, x_6 + 0.0910\, x_7 + 0.250\, x_8 - 0.167\, x_9 + 0.134\, x_{10}.$$

The residuals were plotted by the TEN COB program, as shown in Figure 11. These residuals appear to be normally distributed. Thus, this was the model upon which the conclusions were drawn.

10. INTERPRETING THE RESULTS

The probability distribution of the residuals when the logarithms of the radial miss distances are used in Example I is approximately normal. Thus, several methods of testing hypothesis and setting confidence regions are applicable. A few of these are discussed below.

TABLE X. THE ESTIMATED COEFFICIENTS USING LOGARITHMS, EXAMPLE I

I	BETA(I)	BASIC TYPE	IDENTIFICATION	NUMBER OF VECTORS
0	C.27579525E 01		CONSTANT	
1	-3.94697267E-02	1	TARGET ALT.	1
2	C.363C0478E-01	1	RANGE,LAUNCH	2
3	C.63246976E-03			
4	C.16511960E-01	1	RANGE,INTER.	2
5	-C.7261375EE-03			
6	C.30117154E-02	1	TAR. CL.VEL.	1
7	C.91036746E-01	2	MISSILE MOD.	2
8	C.250C47C8E 00			
9	-0.16691358E 00	2	TARGET TYPE	1
10	C.13372551E 00	2	RADAR POWER	1

THE SUM OF SQUARES FOR ERROR IS 0.35121311E 02

THE ESTIMATE OF THE VARIANCE IS 0.3946214BE 00

THE ESTIMATE OF THE STANDARD DEVIATION IS 0.628189C4E 00

THE NUMBER OF DEGREES OF FREEDOM IS 89.

TABLE XI. THE RESIDUALS USING LOGARITHMS, EXAMPLE I

THE RESIDUALS (ACTUAL OBSERVATIONS - PREDICTED)

1	0.13385597E 01	2	-0.11799058E 01	3	-0.10448199E 01
4	0.13502299E 01	5	-0.13319544E 01	6	-0.11663169E 01
7	0.79629727E 00	8	-0.10740552E 01	9	-0.94477469E 00
10	0.60514525E 00	11	-0.81321748E 00	12	-0.77579290E 00
13	0.46630567E 00	14	-0.62421568E 00	15	-0.39156256E 00
16	0.75755597E 00	17	-0.52438433E 00	18	-0.53046221E 00
19	0.59447633E 00	20	-0.30190491E 00	21	-0.51173764E 00
22	0.60510261E 00	23	-0.37542816E 00	24	-0.45460771E 00
25	0.44047211E 00	26	-0.55882984E 00	27	-0.25522122E 00
28	0.40221058E 00	29	-0.43049059E 00	30	-0.47737660E 00
31	0.13784327E 00	32	-0.19021748E 00	33	-0.38079470E 00
34	0.22922132E 00	35	-0.34795099E 00	36	-0.31082278E-01
37	0.19269481E 00	38	-0.14447162E 00	39	-0.11162561E 00
40	0.61089530E-02	41	-0.87367991E-01	42	-0.16238207E 00
43	0.39692866E-01	44	-0.10614727E 00	45	0.17858613E 00
46	0.12051401E 00	47	-0.21784411E 00	48	-0.28523697E-01
49	0.19804837E 00	50	0.73008554E-02	51	0.13371450E 00
52	0.57462015E-01	53	0.24589617E-01	54	0.62577974E-01
55	0.32117526E-01	56	0.11823651E 00	57	0.17930253E 00
58	0.12092239E 00	59	0.17378875E 00	60	0.25149329E 00
61	0.25372824E 00	62	0.29134430E-01	63	0.18998992E 00
64	0.16246773E 00	65	-0.21211904E 00	66	0.28135192E 00
67	0.26003527E 00	68	0.31340526E 00	69	0.29511288E 00
70	0.28351834E 00	71	0.48460059E 00	72	0.65685284E 00
73	0.47215514E 00	74	0.40259142E 00	75	0.39993109E 00
76	0.41434345E 00	77	0.38470396E 00	78	0.45086261E 00
79	0.42076650E 00	80	0.46927812E 00	81	0.45927099E 00
82	0.58228585E 00	83	0.50373919E 00	84	0.71316219E 00
85	0.72576783E 00	86	0.56370444E 00	87	0.59241320E 00
88	0.94417223E 00	89	0.72913634E 00	90	0.57049624E 00
91	0.86563863E 00	92	0.69634185E 00	93	0.73910604E 00
94	0.10282168E 01	95	0.31617459E 00	96	0.12577172E 01
97	0.11652044E 01	98	0.25843141E 00	99	0.12695047E 01
100	0.16652924E 01				

TABLE XII. THE ESTIMATED COVARIANCE MATRIX USING LOGARITHMS, EXAMPLE I

COLUMNS 1 THROUGH 4 OF THE COVARIANCE MATRIX

0.21074712E-02	-0.51714E49E-04	-0.1429148CE-05	-0.14266172E-03
0.51714849E-04	0.48811196E-32	-0.6079427BE-04	-0.14337896E-02
0.1429148CE-05	-0.6079427BE-04	0.88987955E-06	0.24092065E-04
0.14266172E-03	-0.14337896E-02	0.24092065E-04	0.31801027E-02
0.7C0445E6E-05	0.11421468E-04	-0.4921869[illegible]	-0.71649763E-04
0.21890713E-04	0.17066520E-04	-0.7021053[illegible]	-0.27066986E-04
0.45097491E-02	-0.36673246E-03	-0.10914733E-04	-0.21582248E-02
0.78419342E-03	-0.32173C45E-04	-0.1546C287E-04	-0.25695781E-02
0.42147[illegible]6E-02	-0.2515EC89E-02	0.23552413E-04	0.13179417E-02
0.1020768[illegible]8E-02	0.5541C548E-03	-0.12167119E-04	-0.15815845E-02

COLUMNS 5 THROUGH 8 OF THE COVARIANCE MATRIX

0.7C0445E6E-05	0.2189C713E-04	-0.45097491E-02	-0.78418342E-03
0.11421468E-04	0.17066520E-04	-0.36673246E-03	-0.32173945E-04
0.4C2186[illegible]E-06	-0.7C21C537E-06	-0.10914733E-04	-0.15469287E-04
0.7164976[illegible]E-04	-0.27066986E-04	-0.2158224[illegible]	-0.25695781E-02
0.21228377E-05	0.13463233E-05	0.5599770CE-04	0.82093212E-04
0.13463233E-05	0.54422052E-05	0.16359211E-03	0.12477445E-03
0.5599770CE-04	0.16359211E-03	0.18043962E-02	0.76971968E-01
0.82093212E-04	0.12477445E-03	0.76971968E-01	0.73970126E-01
0.3C2894C5E-04	-0.2C778322E-03	0.27586871E-01	0.71816980E-02
0.36763153E-04	0.21602682E-04	0.57510257E-02	0.16732122E-01

TABLE XII. THE ESTIMATED COVARIANCE MATRIX USING LOGARITHMS, EXAMPLE I (Concluded)

COLUMNS 9 THROUGH 10 OF THE COVARIANCE MATRIX

0.42147C86E-02	0.10207688E-02
0.25158C89E-02	0.5541S548E-03
0.23552413E-04	-0.12167119E-04
0.13179417E-02	-0.15815E45E-02
0.3C2894C9E-04	0.36763153E-04
0.29778322E-03	0.21602682E-04
0.27586671E-01	0.57510257E-02
0.71816S80E-02	0.16732122E-01
0.65895710E-01	-0.36253522E-02
0.36253522E-02	0.22527353E-01

TABLE XIII. THE ESTIMATED CORRELATION MATRIX USING LOGARITHMS, EXAMPLE I

COLUMNS 1 THROUGH 4 OF THE CORRELATION MATRIX

0.100000CCE 01	-0.2F595596E-01	-0.33001279E-01	-0.55106910E-01
0.285955C6E-01	0.1CCCC0C0F 01	-0.92243847E 00	-0.36391936E 00
0.33001279E-01	-0.92243847E C0	0.100C00CCE 01	0.45288490E C0
0.55106910E-01	-0.363C1936E 00	0.4528849CE 00	0.10000000E C1
0.10472135E CC	0.1122C2C4E CC	-0.35810149E CC	-0.87203876E 00
0.20440505E 00	0.10471233E 00	-0.31904300E 00	-0.20574622E C0
0.2312627CE 00	-0.123573C3E-01	-0.27238397E-01	-0.90096964E-01
0.62807087E-01	-0.16932320E-02	-0.60294333E-01	-0.16753790E C0
0.35764994E 00	-0.14027794E 0C	0.97261560E-01	0.31043105E-01
0.14814439E C0	0.5285C727E-01	-0.8593432BE-01	-0.186960C0E 00

483

COLUMNS 5 THROUGH 8 OF THE CORRELATION MATRIX

0.10472135E CC	0.20440505E 00	-0.23126270E 00	-0.62807087E-01
0.11220284E C0	0.10471233E 00	-0.12357303E-01	-0.16932320E-02
0.35810149E CC	-0.31904300E-0C	-0.27238397E-01	-0.60294333E-01
0.87203876E 00	-0.20574622E C0	-0.90096964E-01	-0.16753790E C0
0.10000CCCE C1	0.3960C9E79E CC	0.90478593E-01	0.20714181E C0
0.39609879E 00	0.1CCCCCC0E 01	0.165C8550E 00	0.19665727E 00
0.90478593E-01	0.165C8550E 00	0.10000000E 01	0.66625189E 00
0.20714181E 00	0.19665727E 00	0.66625189E 00	0.10000000E 01
0.8C984859E-01	-0.45726C76E 00	0.252C9265E 00	0.10286565E 00
0.16811207E 0C	0.61697C90E-01	0.90203603E-01	0.40989004E 00

TABLE XIII. THE ESTIMATED CORRELATION MATRIX USING LOGARITHMS. EXAMPLE I (Concluded)

COLUMNS 9 THROUGH 10 OF THE CORRELATION MATRIX

0.35764994E 00	0.14814639E 00
-0.14027794E 00	0.52850727E-01
0.97261560E-01	-0.85934328E-01
0.91043105E-01	-0.18686CC0E 00
0.80984859E-01	0.16811207E 00
0.49726C76E 00	0.61697C99E-01
0.25299265E 00	0.9C2036C3E-01
0.10286565E 00	0.4C985CC4E 00
0.10000000E 01	-0.94055556E-01
0.94095956E-01	0.1CCCCC00E 01

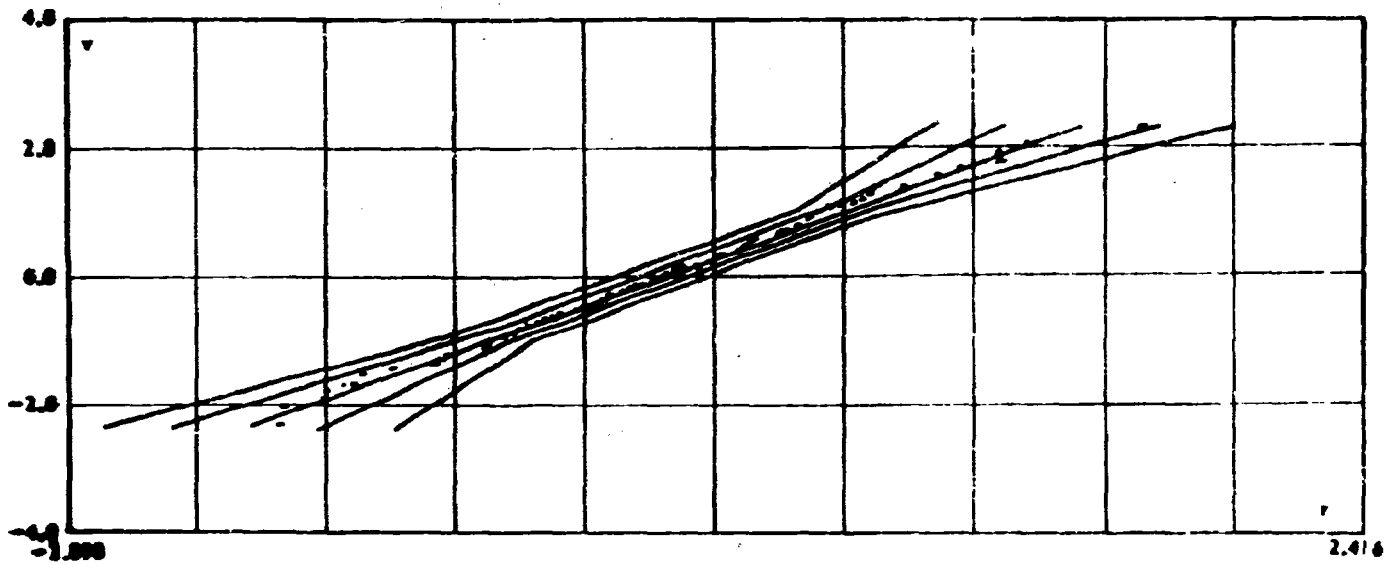


FIGURE 11. POINTS, BEST FITTING LINE, 68-PERCENT CONTROL BAND, AND 98-PERCENT CONTROL BAND FOR THE RESIDUALS USING LOGARITHMS IN EXAMPLE I

a. Testing $b_i = b_i^*$ Using a Student's t Random Table

The CELSP program prints the covariance matrix of the estimates $\hat{b}_1$ through $\hat{b}_k$ of the coefficients of the linear statistical model. The i^{th} diagonal element of the covariance matrix, c_{ii}, estimates the variance of $\hat{b}_i$. The ratio $\frac{\hat{b}_i - b_i}{\sqrt{c_{ii}}}$ has a Student's t distribution with n-k-1 degrees of freedom.

Suppose that a predetermined estimate b_i^*, of b_i, the true value, is available and the following hypotheses are postulated:

$$H_0: \; b_i = b_i^*; \; H_1: \; b_i \neq b_i^* \; .$$

This is a two-sided test since there is no advantage if $b_i < b_i^*$. As before, the level of the test is α. The critical t value is found for the level of the test, the number of degrees of freedom and the fact that it is a two-sided test, $t\left(\frac{\alpha}{2}, \text{n-k-1}\right)$. Since the t distribution is symmetric, the probability that a t random variable is greater than $t\left(\frac{\alpha}{2}, \text{n-k-1}\right)$ or less than $-t\left(\frac{\alpha}{2}, \text{n-k-1}\right)$ is α, i.e.,

$$\text{Prob} \left[|t| > t\left(\frac{\alpha}{2}, \text{n-k-1}\right) \right] = \alpha$$

where $|t|$ is the absolute value of t. This defines the critical region of the test. If H_0 is true, then $\frac{\hat{b}_i - b_i^*}{\sqrt{c_{ii}}}$ is a t random variable and

$$\text{Prob} \left[\left| \frac{\hat{b}_i - b_i^*}{\sqrt{c_{ii}}} \right| > t\left(\frac{\alpha}{2}, \text{n-k-1}\right) \right] = \alpha \; .$$

Thus, the decision is

$$\text{Reject } H_0 \text{ if } \left| \frac{\hat{b}_i - b_i^*}{\sqrt{c_{ii}}} \right| > t\left(\frac{\alpha}{2}, \text{n-k-1}\right) \; .$$

Do not reject H_0 otherwise.

If it is desired to test whether the variable x_i has a "significant effect" upon y, then set $b_i^* = 0$ and perform the above test.

b. Testing the Coefficients for Missile Modifications

In the revised model for the missile, using $z = \ln$ (radial miss distance), the coefficient b_7 represents the difference in z caused by a difference in Mod 1 and Mod 2 missile. That is, $b_7 = z_{\text{Mod 2}} - z_{\text{Mod 1}}$. From Table X, the estimate of this coefficient is

$$\hat{b}_7 = 0.0910 = \hat{z}_{\text{Mod 2}} - \hat{z}_{\text{Mod 1}} .$$

Since $z = \ln (y)$, a positive increase in z corresponds to a positive increase in y. Thus, it would appear that the radial miss distance would be larger for Mod 2 missile than for Mod 1 missiles. To see whether this difference is statistically significant, a t test is performed, with $b_7^* = 0$.

It was decided to set $\alpha = 0.10$. The number of degrees of freedom is, from Table X, 89. The hypotheses are:

$$H_0: b_7 = 0, \qquad H_1: b_7 \neq 0 .$$

The critical t value is

$$t\ (0.05,\ 89) = 1.64 .$$

The estimate of the variance of $\hat{b}_7$ is, from Table XII,

$$c_{77} = 0.1804 .$$

The test statistic is

$$\left| \frac{b_7 - 0}{\sqrt{c_{77}}} \right| = \frac{0.0910}{\sqrt{0.1804}} = 0.21 .$$

The test statistic is less than the critical value of 1.64. Therefore, it cannot be concluded that there is a significant difference in the radial miss distance for Mod 2 and Mod 1 missiles.

A similar test is made for the difference in z of Mod 3 and Mod 2 missiles. Again, $\alpha = 0.10$, $t\ (0.05,\ 89) = 1.64$. The hypotheses are $H_0: b_8 = 0$. $H_1: b_8 \neq 0$. From Tables X and XII, $\hat{b}_8 = 0.250$, $c_{88} = 0.0740$. The test statistic is thus

$$\left| \frac{\hat{b}_8 - 0}{\sqrt{c_{88}}} \right| = \frac{0.250}{\sqrt{0.0740}} = 0.92 .$$

This is less than the critical value. It cannot be concluded that there is a significant difference in the effect of Mod 3 and Mod 2 on radial miss distance.

There is one more difference to be examined; that between Mod 3 and Mod 1. To estimate this difference, the following sum is used:

$$\hat{b}_7 = \hat{z}_{\text{Mod 2}} - \hat{z}_{\text{Mod 1}}$$

$$\hat{b}_8 = \hat{z}_{\text{Mod 3}} - \hat{z}_{\text{Mod 2}}$$

$$\hat{b}_7 + \hat{b}_8 = \hat{z}_{\text{Mod 3}} - \hat{z}_{\text{Mod 1}}$$

In this case

$$\hat{b}_7 + \hat{b}_8 = 0.0910 + 0.250 = 0.341 .$$

Since $\hat{b}_7$ and $\hat{b}_8$ are normally distributed, then $\hat{b}_7 + \hat{b}_8$ is also normally distributed with mean $b_7 + b_8$ and variance

$$\text{var } (\hat{b}_7 + \hat{b}_8) = \text{var } (\hat{b}_7) + \text{var } (\hat{b}_8) + 2 \text{ cov } (\hat{b}_7, \hat{b}_8) .$$

This variance is estimated by

$$c_{77} + c_{88} + 2c_{78} .$$

From Table XII, $c_{78} = 0.0770$. Thus

$$c_{77} + c_{88} + 2c_{78} = 0.1804 + 0.0740 + 0.0770 = 0.3314 .$$

To test the hypotheses

$$H_0: \; b_7 + b_8 = 0; \qquad H_1: \; b_7 + b_8 \neq 0 ,$$

for $\alpha = 0.10$, the critical value is $t\ (0.05,\ 89) = 1.64$. The test statistic is

$$\left| \frac{\hat{b}_7 + \hat{b}_8 - 0}{\sqrt{c_{77} + c_{88} + 2c_{78}}} \right| = \frac{0.341}{\sqrt{0.3314}} = 0.59 .$$

Again, there is insufficient evidence of a difference in miss distance between Mod 3 and Mod 1.

c. Using the F Test in Analysis of Variance

Suppose that a model, called Model I or the complete model, has been fitted to a set of data and it is desired to reduce the model by dropping all terms which do not test as significant. Assume that the level of each test is 0.05 and that the tests are independent. If only one term is tested without positive results by a t test and dropped from the model, then the probability of falsely rejecting is 0.05. If two terms are tested separately with t tests and dropped, the probability of falsely rejecting at least one coefficient becomes $1 - (0.95)^2 = 0.0975$. If four terms are tested and dropped, the probability of falsely rejecting at least one coefficient is $1 - (0.95)^4 = 0.1855$. If this is extended to eight terms, it becomes $1 - (0.95)^8 = 0.3366$. Thus, the t test can be safely used only when one term alone is to be dropped. If more than one term is to be tested, the method of analysis of variance should be used.

Let Model I (the complete model) be the following:

$$\text{Model I: } y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k + e .$$

Model I has been fitted to the data and the sum of squares of error, SSE_1, and the estimate of the variance, s_1^2, have been calculated.

Now suppose that it is desired to test the significance of k-g of these terms. For convenience, assume that the last k-g terms are to be tested. The null hypothesis is then

$$H_0: b_{g+1} = b_{g+2} = \ldots = b_k = 0 .$$

If H_0 is true, then the system can be described by a shorter or deleted model, Model II:

$$\text{Model II: } y = b_0 + b_1 x_1 + \ldots + b_g x_g + e .$$

Model II is then fitted to the data and the sum of squares of error, SSE_2, and the estimate of the variance, s_2^2, are calculated.

Even if H_0 is true, the estimates of b_{g+1} through b_k will not be zero because of the finite random sample of observation noise. Thus, fitting Model I instead of Model II will reduce the sum of squares for error, $SSE_1 \leq SSE_2$. In fact, SSE_2 can be partitioned into two positive quantities

$$SSE_2 = SSE_1 + (SSE_2 - SSE_1) .$$

It can be shown that if H_0 is true, $\frac{SSE_1}{n-k-1}$ and $\frac{SSE_2 - SSE_1}{k-g}$ provide unbiased estimates of σ^2 and that $\frac{SSE_1}{\sigma^2}$ and $\frac{SSE_2 - SSE_1}{\sigma^2}$ are independent Chi-square random variables with n-k-1 and k-g degrees of freedom, respectively. Thus, the ratio of these two independent Chi-square random variables divided by the respective degrees of freedom is an F random variable, if H_0 is true.

$$F = \frac{\frac{SSE_2 - SSE_1}{\sigma^2 (k-g)}}{\frac{SSE_1}{\sigma^2 (n-k-1)}} = \frac{SSE_2 - SSE_1}{(k-g)\, s_1^2} .$$

To test H_0, a one-sided test on this F ratio is used. The level of the test α must again be specified. The critical value of F depends upon α and upon the degrees of freedom, k-g and n-k-1; $F_{crit} = F(\alpha, k-g, n-k-1)$. This is found in statistical tables.

The decision is then based upon a comparison of F with the critical value:

If $F > F(\alpha, k-g, n-k-1)$, reject H_0.

If $F \leq F(\alpha, k-g, n-k-1)$, do not reject H_0.

If H_0 is rejected, the model cannot be shortened by dropping the entire group of k-g terms. Possibly, a subgroup of these k-g terms can safely be dropped, but other models must be postulated and tested to decide which can be dropped.

d. Testing the Effect of Missile Modification Upon Radial Miss Distance Using the F Statistic

A reduced model, without the terms for missile modification, was set up as:

$$z = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_5 + b_6 x_6 + b_9 x_9 + b_{10} x_{10} + e .$$

The estimated coefficients for this model are shown in Table XIV.

The null hypothesis for this test is H_0: $b_7 = b_8 = 0$. It was decided to set $\alpha = 0.10$. The critical F value is

TABLE XIV. THE ESTIMATED COEFFICIENTS FOR THE REDUCED MODEL USING LOGARITHMS, EXAMPLE I

I	BETA(I)	BASIC TYPE	IDENTIFICATION	NUMBER OF VECTORS
0	0.2822?C44E 01		CONSTANT	
1	-C.13045744E-C1	1	TARGET ALT.	1
2	-?.36753511E-C1	1	RANGE,LAUNCH	2
3	C.69349830E-03			
4	?.26067385E-C1	1	RANGE,INTER.	2
5	-?.10532C11E-02			
6	C.2646?312E-02	1	TAR. CL.VEL.	1
9	-?.15728307E 00	2	TARGET TYPE	1
10	?.57511915E-01	2	RADAR POWER	1

THE SUM OF SQUARES FOR ERROR IS 0.35567401E 02

THE ESTIMATE OF THE VARIANCE IS C.39085057E 00

THE ESTIMATE OF THE STANDARD DEVIATION IS 0.62518042E 00

THE NUMBER OF DEGREES OF FREEDOM IS 91.

$$F\ (0.10,\ 2,\ 89) = 2.77\ .$$

From Table X for the complete model,

$$SSE_1 = 35.1213$$

$$s_1^2 = 0.3946\ .$$

From Table XIV for the reduced model,

$$SSE_2 = 35.5674\ .$$

The test statistic is

$$\frac{SSE_2 - SSE_1}{(k-g)\ s_1^2} = \frac{35.5674 - 35.1213}{(2)(0.3946)} = 0.57\ .$$

This value is less than the critical value. It cannot be concluded that missile modification affects the radial miss distance.

11. CONCLUSIONS

Often observations or results of experiments can logically be represented by a linear statistical model relating the observation y to various known quantitative and qualitative variables $x_1, \ldots, x_k$, random noise e, and fixed but unknown coefficients $b_0, b_1 \ldots, b_k$:

$$y = b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k + e\ .$$

The least squares estimates of the coefficients $b_0, b_1, \ldots, b_k$ have been developed theoretically and are calculated by the GELSF program.

If the noise e is Gaussian (normal) with zero mean and covariance matrix $\sigma^2 I$, where I is the identity matrix, then certain quantities have well known distributions. A graphical procedure for making tests for normality of the noise is described.

Crucial in the entire discussion is the model. The model must be of the correct form or the theory collapses. Statistical tests, in particular the analysis of variance, can suggest which terms should be retained in the model, but they cannot prescribe which new terms should be added. The model must be constructed from physical considerations and sound judgment.

Sound judgment should also be employed in the statistical evaluation of the data. In no case should statistical techniques described here or elsewhere be applied in a purely mechanical manner or divorced from the other aspects of the system under consideration.

REFERENCES

1. Mendenhall, W., An Introduction to Linear Models and the Design and Analysis of Experiments, Wadsworth Publishing Co., Belmont, California, 1968.

2. Rich, N. R., "The Theory Behind and the User's Manual for the Generalized Least Squares Fit (GELSF) Digital Computer Program," U. S. Army Missile Command, Redstone Arsenal, Alabama, October 1968, Report No. RD-TR-68-13.

3. Gumbel, E. J., Statistics of Extremes, Columbia University Press, New York, 1958.

4. Greene, L., and Rich, N. R., "Testing for Normality by Control Bands Digital Computer Program User's Manual and Theory," U. S. Army Missile Command, Redstone Arsenal, Alabama, May 1969, Report No. RD-TR-69-8.

5. Snedecor, G. W., and Cochran, W. G., Statistical Methods, Iowa State University Press, Ames, Iowa, 1967.

INFERENCE PROCEDURES BASED ON CENSORED TIME DEPENDENT OBSERVATIONS

Donald R. Barr and Toke Jayachandran
Litton Scientific Support Laboratory
Fort Ord, California

1. Introduction. Frequently, experiments are conducted in such a way that they may be considered to consist of n trials, where in each trial, the time required to achieve some objective ("success") is observed. If an upper limit, T_0, is placed upon the possible duration of each trial, then the outcome on each trial is either the time until success, or T_0, whichever is smaller. Such observations are said to be censored at T_0. For example, it might be desired to determine whether a certain type of combat aid, such as a target detection device, is effective, or whether one type of device is better than another. In order to make such inferences about a single device, often an experiment of the following design is conducted: n players are selected, and each player uses the device in an attempt to detect a target. If a given player has not succeeded within 4 minutes (say), his trial terminates and the next player begins. The observed data then consists of the times of detection for those trials terminating before $T_0 = 4$ (together with the number of trials terminating at $T_0 = 4$). If it is desired to compare two devices, then samples on each device may be taken as described above.

If the distributions of (uncensored) time until success are identical exponentials, this situation falls under the body of results generally known as "life testing." In what follows, we shall discuss statistical procedures for making inferences about the mean rate λ of success (the reciprocal of mean time to success), based upon such censored time dependent observations. These inference techniques include point estimators, confidence intervals and tests of hypotheses for both the single population case, and that of comparing two populations. Since there is a strong parallel between making a confidence interval for a parameter and tests of hypotheses concerning that parameter, we shall discuss only one or

This article has been reproduced photographically from the authors' manuscript.

the other of these in each of the approaches considered below. While it is assumed that the populations involved are exponentially distributed, attention is given to the robustness of the procedures under departures from exponential distributions.

In Section 2, we discuss procedures for making inferences about one population (one-sample inferences), including a review of several approaches in the literature and an easily applied approximate procedure that we have been investigating. A similar treatment for two-sample inferences is given in Section 3.

2. One-Sample Procedures. Suppose that $Z_1, \ldots, Z_n$ is a random sample of size n from an exponential population with mean rate λ. Let $X_i = \min \{ Z_i , T_o \}$; $i = 1, 2, \ldots, n$. Based upon observations $x_1, \ldots, x_n$ on the X's, we wish to make inferences concerning the parameter λ. The (point) estimation of λ has been discussed by several writers (see, for example, Bartholomew [1]). They have shown that the maximum likelihood estimate of λ is

$$\hat{\lambda} = k / \sum_{i=1}^{n} \{ a_i x_i + (1 - a_i) T_o \} , \qquad (1)$$

where

$$a_i = \begin{cases} 1 & \text{if } x_i < T_o \\ 0 & \text{otherwise} \end{cases} ; \quad i = 1, 2, \ldots, n,$$

and

$$k = \sum_{i=1}^{n} a_i .$$

While $1/\hat{\lambda}$ is asymptotically normal with mean $1/\lambda$ and variance

$\left[n\lambda^2(1 - e^{-\lambda T_o}) \right]^{-1}$, the convergence is slow [1] and, for small samples, $\hat{\lambda}$ is seriously biased [6] . The exact distribution of $\widehat{1/\lambda}$ is given by Bartholomew [1] in a form useful with small samples, along with some approximations that are useful with moderate sized samples.

Several procedures for obtaining confidence intervals for λ and $1/\lambda$, based upon approximate distributions for certain functions of $1/\lambda$, have been suggested. Bartholomew [1] discusses two procedures based upon normal distributions, and a technique given in NAVORD O. D. 29304 [6] uses an approach based upon a Poisson distribution. The statistical properties of confidence intervals obtained by these approximations are apparently not fully known at the present time.

We have investigated a method for obtaining confidence intervals for λ , based upon a general approach given by Halperin [4] , described as follows: Let

$$p = 1 - e^{-\lambda T_o} \tag{1}$$

denote the probability that the result in a given trial is not censored (i.e., $p = P[X_i < T_o]$) . Then each experimental trial may be viewed as a Bernoulli trial, where "success" is associated with non-censoring and occurs on each trial with probability p . Based upon the observed number k of success in the n experimental trials, a $100(1-\alpha)$ percent upper confidence bound, say P_U , for p can be constructed using well-known methods. Using equation (1), this bound can be "inverted" to obtain a $100(1-\alpha)$ percent upper confidence bound for λ as

$$\lambda_U = \frac{-\ln(1 - P_U)}{T_o} \tag{2}$$

Intervals and bounds of this type are very easy to compute, and appear to perform nearly as well as those based upon approximating distributions. The results of a Monte Carlo study on the performances of the two procedures based upon the Poisson distribution [6] and the binomial distribution [4] are presented in Tables 1, 2, and 3 below. 1000 samples of size n (n = 20, 30, 40, 50) were generated from exponential distributions with mean rate λ (λ = .1, .2, .5, 1, 2, 5, 10). For different truncation times T_0 the upper confidence limits λ_U were calculated using both procedures. Table 1 gives the average value of the upper confidence limit $\bar{\lambda}_U$. For each choice of λ and T_0 the first row contains $\bar{\lambda}_U$ for the O. D. 29304 [6] procedure (based on the poisson distribution); the numbers in the second row are those for the binomial procedure. Table 2 contains the sample variance of the upper bounds λ_U. In Table 3, the empirical coverage probability i.e., the proportion of the λ_U which actually exceed the true parameter value λ is given. The relative sensitivity of these procedures to departures from the exponential distribution are apparently not known at present.

Finally, we mention another approach to finding confidence intervals for λ, which seems to have received less attention in the literature than it deserves. Imagine that the time censored trials are conducted sequentially in time, and that we disregard the times between when each trial terminates and the next begins (see Figure 1). Then the times between successive successes is exponential with parameter λ, so that the "success arrival

λ	T_o	Sample Size n: 20		30		40		50	
		Confidence Coefficient 1 - α: .95	.99	.95	.99	.95	.99	.95	.99
.1	8.	.165	.199	.151	.177	.144	.164	.138	.156
		.178	.209	.158	.183	.149	.168	.142	.159
.2	2.	.379	.477	.340	.408	.316	.375	.299	.352
		.397	.478	.352	.408	.324	.376	.305	.352
.5	1.	.919	1.12	.804	.973	.752	.896	.723	.840
		.969	1.13	.836	.985	.772	.904	.738	.846
1.	.7	1.68	2.05	1.53	1.82	1.44	1.68	1.40	1.60
		1.78	2.13	1.60	1.86	1.49	1.71	1.43	1.62
2.	.3	3.50	4.28	3.10	3.74	2.97	3.47	2.84	3.26
		3.72	4.39	3.23	3.82	3.07	3.52	2.92	3.30
5.	.05	10.6	14.5	9.46	12.0	8.53	10.7	8.14	9.95
		11.2	14.3	9.80	11.9	8.76	10.6	8.33	9.88
10.	.06	17.0	21.6	15.6	18.8	14.6	17.2	14.1	16.3
		18.0	22.2	16.2	19.2	15.1	17.3	14.5	16.5

TABLE 1. Average upper 100 (1 - α) percent confidence bounds $\bar{\lambda}_U$ for the O. D. and the binomial procedures.

		Sample Size n							
		20		30		40		50	
		Confidence Coefficient 1 - α							
λ	T_o	.95	.99	.95	.99	.95	.99	.95	.99
.1	8.	.002	.002	.001	.001	.001	.001	.000	.001
		.003	.004	.001	.002	.001	.001	.001	.001
.2	2.	.012	.015	.008	.009	.005	.006	.004	.005
		.014	.017	.008	.010	.006	.007	.004	.005
.5	1.	.067	.076	.036	.045	.028	.033	.020	.022
		.080	.094	.042	.053	.030	.038	.021	.024
1.	.7	.194	.267	.124	.142	.086	.096	.063	.074
		.265	.366	.146	.175	.102	.119	.073	.088
2.	.3	.874	1.08	.448	.630	.375	.485	.264	.353
		1.11	1.46	.578	.759	.431	.567	.306	.386
5.	.05	12.7	16.8	7.30	9.34	5.02	6.68	3.88	4.74
		14.0	18.5	7.80	10.2	5.24	6.96	4.05	4.96
10.	.06	19.8	29.4	13 7	16.8	9.15	11.0	6.78	7.96
		25.0	40.1	16.2	20.0	10.6	12.4	7.60	9.29

TABLE 2. Sample Variances of λ_U

λ	T_o	Sample Size n: 20		30		40		50	
		Confidence Coefficient $1-\alpha$: .95	.99	.95	.99	.95	.99	.95	.99
.1	8.	.954	.995	.954	.994	.951	.993	.955	.996
		.977	.994	.963	.995	.959	.991	.955	.998
.2	2.	.950	.996	.967	.993	.960	.997	.950	.995
		.980	.996	.968	.993	.986	.993	.961	.995
.5	1.	.968	.997	.955	.993	.951	.995	.963	.992
		.985	.997	.984	.992	.952	.994	.969	.989
1.	.7	.951	.991	.958	.993	.956	.996	.950	.994
		.972	.991	.958	.993	.959	.993	.964	.996
2.	.3	.961	.995	.961	.993	.958	.996	.965	.988
		.985	.995	.972	.994	.961	.995	.966	.985
8.	.05	.949	1.00	.981	.997	.948	.995	.953	.995
		.949	.991	.981	.997	.948	.995	.975	.994
10.	.06	.951	.996	.954	.996	.948	.992	.954	.995
		.982	.995	.965	.997	.957	.991	.962	.991

TABLE 3. Proportion of the λ_U that exceed the true parameter value λ.

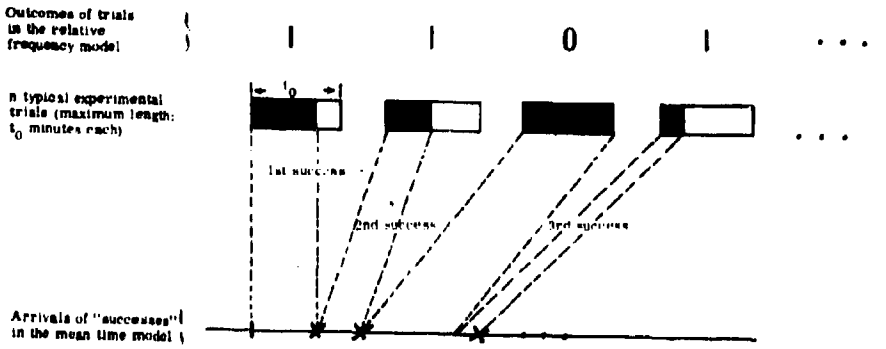


FIGURE 1.

process" is a poisson process with mean rate λ per unit time. Suppose that $k > 0$ successes are observed in n experimental trials. Then the conditional distribution of the waiting time to kth arrival, W_k, given k, is approximately gamma distributed. Thus $2\lambda W_k$ is chi-square distributed with $2k$ degrees of freedom. Using the tabulated χ^2 distribution, and given $k > 0$, one can easily find an interval (C_L, C_U) which will contain $2\lambda W_k$ with probability $1-\alpha$. Algebraic manipulation upon this interval yields

$$\left(\frac{C_L}{2W_k} \quad , \quad \frac{C_U}{2W_k} \right) \tag{3}$$

as a 100 $(1-\alpha)$ percent conditional confidence interval for λ.
The sensitivity of this approach to departures from exponential distributions depends upon T_0; with sufficiently small T_0, the interarrival times should be nearly exponential even with non-exponentially distributed times to success within the individual trials.

3. Two-Sample Procedures. In this section, techniques for comparing two exponential populations based on censored observations, are discussed. Suppose independent samples of sizes n_1 and n_2 are drawn from two exponential populations with means $1/\lambda_1$ and $1/\lambda_2$ respectively. It will be assumed that both sets of sample observations are censored at the same time point T_0 ; that is, an upper limit T_0, is placed upon the possible duratior of each trial. Let k_1 and k_2 denote the number of uncensored observations in the first and second sample respectively and let $p = k_1 + k_2$. Methods for testing the hypothesis $H_0 : \lambda_1 = \lambda_2$ and obtaining confidence intervals for $\rho = \lambda_1/\lambda_2$ and $\psi = \lambda_1 - \lambda_2$ are discussed below.

a. F test for $H_0 : \lambda_1 = \lambda_2$: Let W_{1,k_1} denote the total elapsed time time until k_1 uncensored observations are obtained from the first sample. Then, as mentioned earlier, $2\lambda_1 W_{1,k_1}$ has a chi-square distribution with $2k_1$ degrees of freedom. If W_{2,k_2} is similarly defined for the second sample, so that $2\lambda_2 W_{2,k_2}$ is distributed as a chi-square with $2k_2$ degrees of freedom, then, if H_0 is true, the ratio

$$W = \frac{2W_{1,k_1}/2k_1}{2W_{2,k_2}/2k_2} \tag{4}$$

has an F distribution with $(2k_1, 2k_2)$ degrees of freedom. The hypothesis H_0 will be rejected in favor of $H_1 : \lambda_1 < \lambda_2$, at significance level α , if the observed value W exceeds the 100αth percentile of the F distribution with $(2k_1, 2k_2)$ degrees of freedom. A confidence interval for $\rho = \lambda_1/\lambda_2$ can be obtained, based on the F distribution of $k_2\lambda_1 W_{1,k_1} \Big/ k_1\lambda_2 W_{2,k_2}$.

b. Cox's F' test : Let $n = n_1 + n_2$ and let the scores $t_{r,n}$ ($r = 1, 2, \ldots, n$) denote the expected values of the order statistics of a random sample of size n from an exponential distribution with mean equal to 1. It can be shown that

$$t_{r,n} = \sum_{s=0}^{r-1} \frac{1}{n-s} \qquad (r = 1, 2, \ldots, n)$$

Combine the $p = k_1 + k_2$ uncensored observations, defined in the beginning of the section and rank them. Replace the observation with rank r with the corresponding score t_{rn} ($r = 1, 2, \ldots, n$) . If two or more of the censored observations are equal, replace each one with the average of the corresponding scores t_{rn}. Let $\bar{t}_1$ denote the average of the scores assigned to the observations from the first sample and $\bar{t}_2$ the average of the scores assigned to the observations from the second sample. Cox [2] has shown that the ratio

$$W' = \frac{\left[k_1 \bar{t}_1 + (n_1 - k_1)\, t_{p+1,n} \right] / k_1}{\left[k_2 \bar{t}_2 + (n_2 - k_2)\, t_{p+1,n} \right] / k_2} \qquad (5)$$

is approximately distributed as an F with $(2k_1, 2k_2)$ degrees of freedom, when $H_0 : \lambda_1 = \lambda_2$ is true. The rejection region for the hypothesis $H_0 : \lambda_1 = \lambda_2$ can be determined from the F tables.
A confidence interval for $\rho = \lambda_1/\lambda_2$ can be obtained using this approximate distribution of W' , as follows: Multiply each of the uncensored observations from the second sample by a fixed number ρ_0 and apply Cox's procedure. This will lead to a test of the hypothesis $H_0' : \lambda_1 = \rho_0 \lambda_2$.

For a given significance level α, the set of all values of ρ_0 that will lead to rejection of H_0' will from a confidence interval for ρ with confidence coefficient $1-\alpha$.

Recently, Gehan and Thomas [3] have reported a Monte Carlo study comparing the powers of the F and F' tests for small sample sizes. It was found that, when the assumption that the two samples are from exponential distributions is valid, these tests have comparable operating characteristics, (see Figure 2 below). However, if the samples are from Weibull distributions, the F test is not robust and the F' test is superior. It should be noted that the F' test requires that both sets of sample observations are censored at the same time point T_0 ; the F test is not constrained with this requirement.

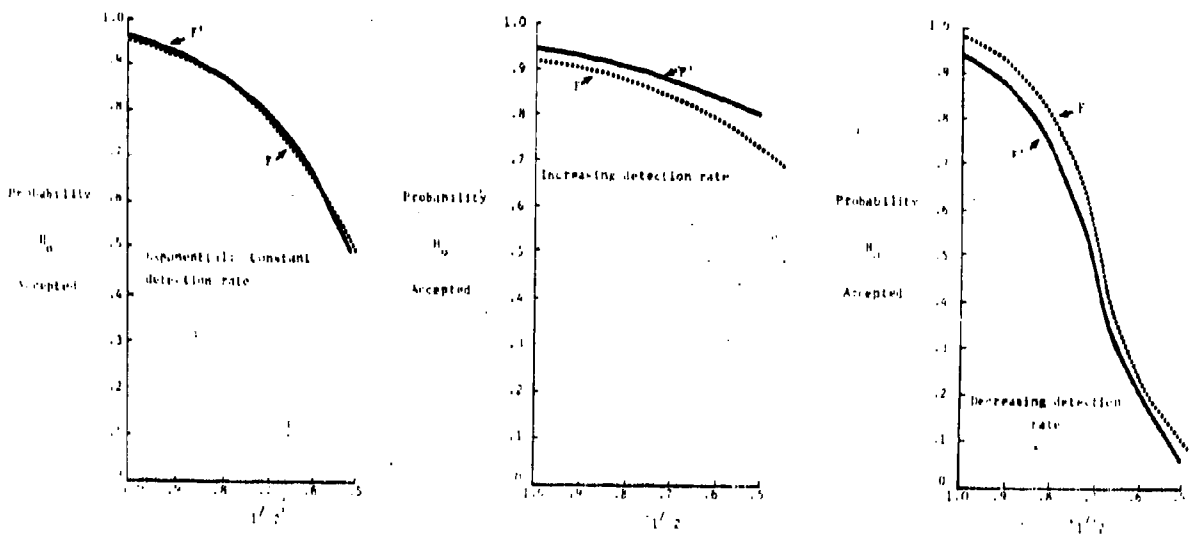


FIGURE 2. Operating Characteristic Curves for the F and F' Tests With Censored Observations for Exponential and Certain Weibull Distribution. ($n_1 = n_2 = 20$, $\alpha = .05$ and $\lambda_2 = 1.0$)

c. Confidence interval for $\psi = \lambda_1 - \lambda_2$: Suppose W_1 and W_2 are independent random variables with gamma distributions with parameters λ_1 and λ_2. Lenter and Buehler [5] obtained the conditional distribution of $H(u/v ; \psi)$ of $U = W_1$ given $V = W_1 + W_2$. This conditional distribution function involves only the parameter $\psi = \lambda_1 - \lambda_2$. A confidence interval for ψ can be obtained as follows: Set $H(u/v : \psi)$ equal to $\alpha/2$ and $1-\alpha/2$ respectively and solve for ψ. The two solutions for ψ will be the lower and upper confidence limits for ψ with confidence coefficient $1-\alpha$.

The Lenter-Buehler technique can be used to derive a confidence interval for the difference $\lambda_1 - \lambda_2$ for two exponential distributions when the observations are censored. As was pointed out earlier in this paper, if the censored observations from an exponential distribution are treated as having been obtained sequentially, then $2\lambda W_k$, where W_k is the waiting time till kth arrival, given k, has a Chi-square distribution with 2k degrees of freedom. Thus W_k has a gamma distribution. We can now form two gamma distributed variables from the two sets of observations from the exponential distributions and apply Lenter-Buehler [5] technique to obtain a confidence interval for $\lambda_1 - \lambda_2$.

REFERENCES

[1] Bartholomew, D. J. . The sampling distribution of an estimate arising in life testing. Technometrics 5. No. 3 pp 361-374 (1963)

[2] Cox, D. R. . Some applications of exponential order scores. Journal of the Royal Statistical Society, Series B. 26. pp 103-110 (1964)

[3] Gehan, E. A , and Thomas. D. G. . The performance of some two-sample tests in small samples with and without censoring. Biometrika 56, 1, pp 127-132 (1969)

[4] Halperin, M. , Confidence intervals from censored samples. Annals of Mathematical Statistics. 32, 3, pp 828-37 (1961)

[5] Lentner, M. M. , and Buehler, R. J. , Some inferences about gamma parameters with an application to a reliability problem. Journal of the American Statistical Association 58, pp 670-77 (1963)

[6] NAVORD OD 29304/ADDENDUM. Statistical exposition of guide manual for reliability measurement program. 15 November 1967.

OPTIMIZING A FOUR-PART ASSAY PROCEDURE

Walter D. Foster
Biomathematics Division, Department of the Army
Fort Detrick, Maryland

ABSTRACT. Optimizing the procedure of a four-part assay of a bulk material is considered on the basis of the variance of each part versus its cost. In d replications (defined as Days), a settling Chamber is used c times to deposit organisms on p Plates per chamber run; the plates are incubated and Read r times per plate. A nested analysis of variance provided estimates of the components for the variance function:

$$V(\text{Assay}) = \frac{\sigma_r^2}{rpcd} + \frac{\sigma_p^2}{pcd} + \frac{\sigma_c^2}{cd} + \frac{\sigma_d^2}{d}$$

The cost function is the sum of the cost rate x number of repeats for each part:

Cost = (unit Reading cost) rpcd + (unit Plating cost) pcd

+ (unit Chamber cost) cd + (unit Day cost) d.

A Lagrangian multiplier approach yielded minimum variance for a fixed total cost:

Min [V(Assay) + λ(Cost)].

This approach is contrasted with non-linear programming and integer programming.

INTRODUCTION. A problem common to both bulk sampling and sample surveys is to obtain as much information about a population, usually stratified or segmented in time or space, as possible for a given budget. For example, grain production in the mid-west or even from a specific farm would need to be sampled according to various strata defined by bags, bushels, wagon loads, etc. To maximize the information, i.e., minimize the variance of the sample mean for a given cost is the way the problem is usually stated.

Duncan[1], among others, has written extensively on procedures for bulk sampling and has a pertinent bibliography. Similarly, Cochran[2], to pick another outstanding writer, has worked extensively in the area of sample surveys. The problem addressed here is to optimize an assay

Preceding page blank

procedure for a sporulating organism where the steps may be considered as a three-phase procedure. This problem is answered partly by borrowing the technique from bulk and sample survey solutions; it is also approached from the point of view of integer programming.

A "day" is a natural block of work. A small amount of the spore material in dry state is encapsulated and dispersed with a CO_2 pistol into a small chamber. Several agar plates are set in the bottom of the chamber onto which the dispersed spores fall. After a suitable incubation period in a humidity chamber, the plates are "read" until 100 organisms are counted and the percent viability is estimated. Several such readings are made. It is the purpose here to minimize the variance of this assay procedure subject to a particular assay cost.

I. VARIANCE FUNCTION. An experiment was designed specifically to provide estimates of the various sources of variation. In each of several "days," repeated chamber runs provided an opportunity to observe variation from one chamber run to another within the same day. Six chamber runs per day were scheduled, three agar plates per chamber run were prepared, and each plate was read six times. The experiment continued over 18 days. For sake of simplicity the variation due to technicians is omitted in this report.

A nested analysis of variance gives both the structure of the analysis and the results.

ANALYSIS OF VARIANCE ON % VIABILITY

Source	df	MS	EMS
Days	17	315.013	$\sigma_R^2 + 2\sigma_p^2 + 6\sigma_c^2 + 36\sigma_D^2$
Chambers in Days	90	26.990	$\sigma_R^2 + 2\sigma_p^2 + 6\sigma_c^2$
Plates in chambers	216	20.981	$\sigma_R^2 + 2\sigma_p^2$
Readings in plates	324	15.012	σ_R^2

$$\sigma_R^2 \doteq 15; \quad \sigma_p^2 \doteq 3; \quad \sigma_c^2 \doteq 1; \quad \sigma_D^2 \doteq 8$$

Because of the balanced arrangement, there was no problem in estimating the components in this hierarchical analysis according to the expected mean square shown in the table above. The estimates of these components of variance are also shown there. In this completely randomized design with all factors considered random, the variance of the mean percent viability is given the following equation:

$$V(\overline{\overline{XV}}) = \frac{\sigma_R^2}{r} + \frac{\sigma_p^2}{rp} + \frac{\sigma_c^2}{rpc} + \frac{\sigma_D^2}{rpcd}$$

This variance function permits the experimenter to obtain any degree of precision he is willing to pay for merely by assigning appropriate values to the number of readings per plate, the number of plates per chamber, the number of chamber runs per day, and the number of days. One criterion for this degree of accuracy is the cost of the assay. A second criterion concerns the intended use of the material and to what degree of accuracy he must know its viability.

II. COST FUNCTION. Cost functions are notoriously difficult to describe realistically, especially when the assay procedure itself is conducted on a sporadic basis. The function chosen here was selected primarily because of its simplicity and was felt to be adequate for the problem at hand:

$S = rpcd\ R + pcd\ P + cd\ C + d\ D$, where $R = .20$; $P = .50$; $C = .50$; and $D = 2.50$ are the (fictitious) cost per unit for Reading, Plate, Chamber and Day, and r, p, c, d are the number of units of each.

As is clear, this function also follows a nested construction and assigns to each portion of the assay a fixed cost so that total cost is obtained merely by counting the number of component parts. Overhead costs such as equipment, utilities, laboratory space, etc., have been included in the "day" costs.

If the experimenter is willing to pay S dollars for one assay of the material, then the conjunctive use of the variance and the cost function will optimize the assay procedure for that cost. Conversely, the cost function will show him what he would need to pay to obtain an assay for a prescribed precision. The optimization processes are shown in the next section.

III. OPTIMIZATION PROCESSES.

A. Lagrangian Multipliers

Let us write the following function

$$L = V(\overline{X}) + \lambda S$$

which includes both the variance and the cost as a quantity to minimize. Following the standard procedure, partial derivatives of the function L are taken with respect to the parameters d, c, p, r giving four equations in five unknowns, the fifth unknown being the variable λ. The fifth equation is the cost function itself. These derivatives, which can be

seen in the Appendix, give rise to a set of five non-linear simultaneous equations whose solutions are also shown in the Appendix and were obtained as a unique set by virtue of the simplicity of the nested design.

It is clear, although not proved mathematically, that any hierarchical design with random effects and an associated nested cost function will have these two properties: non-linearity and analytic solution.

The procedure involving the Lagrangian multipliers clearly treats d, c, p, r as continuous variables. Obviously they must be positive as well as all of the other input values such as cost rates and the variance components. In keeping with this approach, we should then expect to find the solution for these variables to be non-integers. For a limit of S = \$9.00 per assay, the optimum solution gave

$$d = 1.98;\quad c = .79;\quad p = 1.73;\quad r = 3.54;\quad V(\bar{X}) = 7.4;\quad S = 9.0$$

Because the experimenter must work with integer values of these variables, the next highest and lowest integer value of each variable were examined in combinations with similar upper and lower values for all other variables giving rise in general to 2^4 combinations. The variance and cost functions were computed for 8 of the 16 combinations plus one other; and, the one that minimized the variance for the allotted cost was chosen: specifically where S = \$9.00, d = 2, c = 1, p = 1, r = 5. These 9 combinations are shown in the table below with the selected set shown with an asterisk.

d	c	p	r	$V(\bar{X})$	S
1	1	1	3	17.0	4.1
1	1	1	4	15.8	4.3
1	1	2	3	13.0	5.2
1	1	2	4	12.4	5.6
2	1	1	3	8.5	8.2
2	1	1	4	7.9	8.6
2	1	2	3	6.5	10.4
2	1	2	4	6.2	11.2
2	1	1	5	7.5	9.0*

By extending the perturbation range beyond adjacent values, such a table also has the advantage of displaying to the experimenter what a small increase in cost would yield in terms of increased precision. Similar solutions would be available at other levels of possible cost, S, per assay.

B. Integer Programming

Integer programming is that type of mathematical programming that seeks to optimize a particular function subject to certain constraints

where the constraints are expressed in integer form. Mathematical theory to date has not discovered a general procedure for problems in integer programming but several alternative procedures are available that are felt to be optimum for well-behaved surfaces. A fuller discussion of integer programming is beyond the scope of this paper and is not pursued here. An effective alternative procedure consists of asking the computer to map the surface for all combinations which are reasonable and to select the optimum value which satisfies the cost restraint.

APPENDIX

(1) $S = Dd + Cdc + Pdcp + Rdcpr$

(2) $$L = [V(\bar{X}) + \lambda(S)] = \frac{\sigma_d^3}{d} + \frac{\sigma_c^3}{dc} + \frac{\sigma_p^3}{dcp} + \frac{\sigma_r^2}{dcpr} - \lambda S + \lambda Dd + \lambda Cdc + \lambda Pdcp + \lambda Rdcpr$$

(3) $$\lambda = \frac{\sigma_r^2}{Rd^2c^2p^2r^2}$$

(4) $$\frac{\partial L}{\partial p} = \frac{-\sigma_p^2}{dcp^2} + \frac{-\sigma_r^2}{dcp^2r} + \lambda[Pdc + Rdcr] = 0$$

(5) $$\frac{\partial L}{\partial c} = \frac{-\sigma_c^2}{dc^2} + \frac{-\sigma_p^2}{dc^2p} + \frac{-\sigma_r^2}{dc^2pr} + \lambda[Cd + Pdp + Rdpr] = 0$$

(6) $$\frac{\partial L}{\partial d} = \frac{-\sigma_d^2}{d^2} + \frac{-\sigma_c^2}{d^2c} + \frac{-\sigma_p^2}{d^2cp} + \frac{-\sigma_r^2}{d^2cpr} + \lambda[D + Cc + Pcp + Rcpr] = 0$$

(7) $$r^2 = \frac{P}{R} \cdot \frac{\sigma_r^2}{\sigma_p^2}$$

(8) $$p^2 = \frac{C}{P} \cdot \frac{\sigma_p^2}{\sigma_c^2}$$

(9) $$c^2 = \frac{D}{C} \cdot \frac{\sigma_c^2}{\sigma_d^2}$$

(10) $$d = \frac{S}{D + C\sqrt{\frac{D}{C} \cdot \frac{\sigma_c^3}{d^2}} + P\sqrt{\frac{D}{P} \cdot \frac{\sigma_p^3}{d^2}} + R\sqrt{\frac{D}{R} \cdot \frac{\sigma_r^2}{d^2}}}$$

(11) $$d = \frac{\sqrt{D} \cdot \sigma_d \cdot S}{D[\,\sigma_d\sqrt{D} + \sigma_c\sqrt{C} + \sigma_p\sqrt{P} + \sigma_r\sqrt{R}\,]}$$

AN APPLICATION OF MATHEMATICAL PROGRAMMING TO EXPERIMENTAL DESIGN

J. Richard Moore
U. S. A. Aberdeen Research and Development Center
Aberdeen Proving Ground, Maryland 21005

It often happens that a number of observations (measurements) of interest can be made each time a complex system is operated. When there are several different modes in which such a system can be operated (missions) the set of observations may be different for each mission. Many observations may be common to several modes of operation while some may be peculiar to a particular mission.

When it is necessary to obtain a number of observations for each of a known set of subsystems (or functions) the question of how many times the system must perform each type mission arises. A reliability demonstration test for a complex system is an example of the above. This paper gives an example of a procedure by which the required set of observations can be obtained at a minimum cost. Hopefully this example will help those who are active in research and development to recognize situations in which appreciable saving can be effected by the use of mathematical programming. These procedures are relatively simple to use once the problem to which they apply has been identified.

As an example, consider a complex system which consists of six major subsystems. A reliability demonstration test must be conducted for the system and also for each of the subsystems operating as part of the system. This system is capable of performing four different types of mission. The amount of operating time (number of operating cycles) for each subsystem during a mission depends on the mission type but it is assumed that the probability that a specific subsystem will fail during a mission depends only on the amount of operating time (cycles) it accumulates during the mission and not on the type of mission which the system is performing. That is, it is assumed that the distribution of time (cycles) to failure does not depend on the type of mission the system is performing when the subsystem is accumulating operating time (cycles).

This article has been reproduced photographically from the author's manuscript.

It is necessary to use a reliability demonstration plan which allows no failures because of the limited availability of test resources. For the system to be tested this dictates that, in order to demonstrate the desired reliability with the required confidence, each of the subsystems must accumulate the operating experience given in Table I.

TABLE I

Required Subsystem Operating Experience

Subsystem	Operating Time (Cycles)
1	60 minutes
2	30 minutes
3	40 minutes
4	80 minutes
5	20 cycles
6	30 cycles

It is further required that the system must complete at least two of each type mission without failure.

The operating time (cycles) for each subsystem during each type mission is given in Table II

TABLE II

Operating Experience for Subsystems by Mission Type

Subsystem	Mission Type			
	1	2	3	4
1	6	12	0	3
2	9	0	4	6
3	8	2	0	5
4	0	6	10	3
5	5	1	2	0
6	3	0	2	3

The units of measurement for the entries in Table II are minutes for subsystems one through six and cycles for subsystems seven and eight. Thus, Table II says that a subsystem one operates for six minutes during a Type 1 mission, twelve minutes during a Type 2 mission, three minutes during a Type 4 mission and doesn't operate during a Type 3 mission, etc.

The problem is to find the best way to run this test program in the sense that the total cost of testing is minimized. The cost of running each type mission is given in Table III.

TABLE III

Cost of Running Each Type Mission

Mission Type	Cost of Mission
1	40
2	38
3	43
4	39

The cost data in Table II are relative costs which are in thousands of dollars in this example.

With the preceding information available, it is seen that this is a problem in linear programming in which one must determine the minimum number of times to require the system to complete each type mission, i.e., minimize the cost of testing, subject to the constraints that each subsystem must accumulate at least as much operating experience as is given in Table I and also the system must complete at least two of each type mission. Since the solution will be in terms of number of tests, it also follows that the solution vector for the linear program must be integer valued. A problem of this type is called an integer programming problem.

The first step in finding a solution to the problem will be to formulate it in standard terminology. Let Y be a column vector with jth element equal to the number of times the jth type mission must be run, that is

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}$$

where (y_j) are unknown which must be determined. Let C be a row vector with elements equal to the mission costs given in Table III, B be a column vector with elements equal to the required total operating requirement for each subsystem as given in Table I and A be a matrix with elements equal to the operating time (cycles) for each mission type as given in Table II. With this notation the problem is to find the vector Y such that:

$$CY = \text{minimum}$$

$$AY \geq B$$

$$Y \geq \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix}$$

and all of the elements of Y must be integers. Substituting the numbers given in Table I, Table II and Table III in the above relationships leads to the following statement of the problem:

Find a vector of integers, Y, such that

$$CY = 48y_1 + 38y_2 + 43y_3 + 39y_4 = \text{minimum}$$

and

$$\begin{aligned} 6y_1 + 12y_2 + 3y_4 &\geq 60 \\ 9y_1 + 4y_3 + 6y_4 &\geq 30 \\ 8y_1 + 2y_2 + 5y_4 &\geq 40 \end{aligned}$$

$$A \cdot Y = \begin{aligned} 6y_1 + 10y_2 + 3y_4 &\geq 80 \\ 5y_1 + y_2 + 2y_3 &\geq 20 \\ 3y_1 + 2y_2 + 3y_4 &\geq 30 \end{aligned} = B$$

$$Y = \begin{aligned} y_1 &\geq 2 \\ y_2 &\geq 2 \\ y_3 &\geq 2 \\ y_4 &\geq 2 \end{aligned}$$

The last four constraints can be eliminated from the problem by the translation of axis

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = X = Y - \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix} = Y - 2e$$

where $e = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

This transforms the original problem to the equivalent problem: Find a vector X such that

$$CX = \text{minimum}$$
$$AX \geq B_x$$
$$X \geq 0$$

where $B_x = B - AM$

and $M = 2e$ for this problem.

The next step is to inspect the vector B_x to determine whether the complexity of the problem can be reduced. If any element of B_x is zero or negative the operating requirements for the corresponding subsystem will have been satisfied when the system has completed the number of each type mission necessary to satisfy the constraints $Y \geq M$. Thus, the size of the problem can be reduced by dropping all constraints corresponding to zero or negative elements of B_x.

In the example $B_x = B - 2Ae$ is

$$B_x = \begin{pmatrix} 60 \\ 30 \\ 40 \\ 80 \\ 20 \\ 30 \end{pmatrix} - \begin{pmatrix} 6 & 12 & 0 & 3 \\ 9 & 0 & 4 & 6 \\ 8 & 2 & 0 & 5 \\ 0 & 6 & 10 & 3 \\ 5 & 1 & 2 & 0 \\ 3 & 0 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 18 \\ -8 \\ 10 \\ 42 \\ 4 \\ 14 \end{pmatrix}$$

The second element in B_x is negative, which means that if each type mission is run two times in order to satisfy the requirements $X \geq 0$ ($Y \geq 2e$) there will be (9)(2) + (0)(2) + (4)(2) + (6)(2) = 38 minutes operating time on subsystem number two, which is eight minutes more than the required thirty minute operating time for that subsystem; hence, the second element in B_x is minus eight. This allows a reduction in the size of the problem without changing the solution by dropping the second constraint equation.

Letting A_o and B°_x represent the reduced matrix and vector resulting from the deletions, the problem is rewritten as:
Find a vector X such that

$$CX = \text{minimum}$$
$$A_oX \geq B^\circ_x$$
$$X \geq 0$$

The next step is to inspect the columns of the matrix A_o to determine whether the dimension of the problem can be reduced by omitting one or more of the elements of the vector X. If any column of A_o contains all zeros then none of the subsystems for which data is needed will operate during the type mission corresponding to that column. Requiring the system to complete such a mission would only increase the cost of testing without generating any needed data. The dimension of the problem is reduced by eliminating the zero columns from A_o and the corresponding elements from the vectors X and C. Letting A°, X° and C° be the new matrix and vectors which result from this transformation, the problem is stated in final form as: Find a vector X° such that

$$C^\circ X^\circ = \text{minimum}$$
$$A^\circ X^\circ \geq B^\circ_x$$
$$X^\circ \geq 0$$

In the numerical example, omitting the second row in A and the second element in B_x reduces the problem to the form:

$$CX = (40 \quad 38 \quad 43 \quad 39) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \text{minimum}$$

$$A_o X = \begin{pmatrix} 6 & 12 & 0 & 3 \\ 8 & 2 & 0 & 5 \\ 0 & 6 & 10 & 3 \\ 5 & 1 & 2 & 0 \\ 3 & 0 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \geq \begin{pmatrix} 18 \\ 10 \\ 42 \\ 4 \\ 14 \end{pmatrix} = B_\lambda$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

The dimension of the problem can not be reduced because none of the columns of A_o contain only zero elements. Thus, it is in the final, reduced form.

With the exception of the requirement that the elements of X° must be integers, the above is simply a linear programming problem which can be solved using the simplex method. It can be shown that, if all of the elements of A° are either zero or one and all of the elements of B°_x are integers, then the solution vector obtained using the simplex method will have all integer elements. This would be the case if each of the subsystems were of cyclic nature, e.g., relays which operate for only one cycle when required during a mission. Also, if for each subsystem, the operating time (number of cycles) is the same during each mission in which it is required to operate and the operating requirement is a multiple of this, then the problem can be reduced to the zero-one-integer problem and the simplex method can be used to find the solution vector for the integer programming problem. It should be noted that the above remarks apply to the problem after it has been put in reduced form by making all possible decreases in size and dimension. That is, even though it is not possible to formulate the original problem as a zero-one-integer problem it may be possible to do so for the problem in reduced form.

It should be emphasized that the conditions of the preceding paragraph are sufficient but not necessary to assure that the simplex method will yield an integer solution to the problem. Since the simplex method is much easier to apply than the methods of integer programming, it is advisable to first find the simplex solution and see whether it is integer valued before proceeding further.

The simplex solution for the example given here is not integer valued so a search routine was used on the computer to find the solution, which is:

$$X = \begin{pmatrix} 0 \\ 1 \\ 3 \\ 3 \end{pmatrix} \text{ or } Y = X + 2e = \begin{pmatrix} 2 \\ 3 \\ 5 \\ 5 \end{pmatrix}$$

with minimum cost of

$$CY = (40 \quad 38 \quad 43 \quad 39) \cdot \begin{pmatrix} 2 \\ 3 \\ 5 \\ 5 \end{pmatrix} = 604$$

The reader may find it interesting to show that the solution satisfies all of the constraints.

The problem formulated here is of the same structure as the nutrition or diet model found in the literature. The reader is referred to Chapter 27 of "Linear Programming and Extensions" by George B. Dantizg for an excellent treatment of the formulation and treatment of that problem.

Transmission of Infrasonic Waves Generated by Large Missile Launches

by

R. E. Lacy and C. E. Sharp
Institute for Exploratory Research
U. S. Army Electronics Command
Fort Monmouth, New Jersey

Summary

Infrasonic pressure waves generated during launching of large missiles can be detected at distances greater than 2000 kilometers. On a number of occasions during recent years, two infrasonic wave trains spaced about 30 minutes apart have been recorded by sensor array areas located at Fort Monmouth, N. J. These areas are approximately 1400 kilometers from the launch pads at Cape Kennedy, Florida. To date, no explanation for these observed early arrivals of an infrasonic wave train which appears to travel at approximately twice the normal velocity of sound through the atmosphere, has proven to be acceptable. Most recently, we have observed a relationship between these sounds and the jet streams traveling the proximity of the path from Cape Kennedy to Fort Monmouth. On this basis, we have evolved a hypothesis that such correlation exists between the jet streams and this anomalous propagation of sound waves. Because of the large amount of jet stream data, but relatively few missile-firing events, a statistical design of further experiments adequate to test the hypothesis is planned. A theoretical analysis explains the pressure wave spectra received as consisting of three separate groups of spectra.

Introduction

Infrasonic wave propagation phenomena have been explored with renewed interest in recent years. In this report we summarize our current findings. We are concerned here only with infrasonic frequencies traversing the atmosphere. Infrasonic waves are inaudible sound waves whose frequency of oscillation are below 15 Hz; they have the same velocity as audible sounds, i.e., about 332 m/sec at 20°C and a pressure of 760 mmHg. The absorption of infrasound is considerably less than absorption of audible sound of the atmosphere due to heat conduction and viscosity.[1] Because of this, the detection at long distances of infrasonic energy generated by the large missiles, as represented by Titan III and Saturn V, launched from Cape Kennedy, Florida, is possible. Frequency analysis of a number of magnetic tape recordings made at Fort Monmouth, N. J. during the arrival of infrasonic waves generated by the missile launches show the maximal energy propagated to be centered within 0.5 to 5 Hz. The maxima of these traveling pressure oscillations recorded in the Fort Monmouth test areas seldom exceed 2.5 dynes/cm^2. The "normal" traveltime for infrasonic waves from Cape Kennedy to Fort Monmouth is between 70 to 85 minutes after launch. This

This article has been reproduced photographically from the author's manuscript.

variation in arrival time may be due to the directions and velocities of the surface winds along the travel path. The duration of the infrasonic wave trains also has been observed to vary between five and ten minutes. This variation may be the result of the relative signal-to-noise ratio at the sensor location for each launching. Furthermore a so-called precursor or wave train has been sporadically observed that arrived about 10 minutes earlier than the "normal" wave train. Figure 1 clearly shows this precursor recorded, when the wind turbulence was practically nil, at a site 1000 km north of Cape Kennedy, Florida.

The Wind Turbulence Problem

The detection at long distances of acoustical energy produced by the large missile launches (Titan, Atlas, Saturn, etc.) from Cape Kennedy is complicated by low signal-to-noise ratios. The acoustical energy to be detected is in the form of atmospheric pressure waves of usually less than 2 dynes/cm^2, contributed in the selected frequency spectra 0.01 to 10 Hz. At the distance of interest for this study, approximately 1400 km, the amplitude and frequency range of the pressure waves detected have been decreased considerably by atmospheric absorption, by the dispersive effects of wind and temperature variations in the lower and upper atmospheric regions, and by spherical spreading. At long ranges, the combined effect of these factors results in a reduction of the amplitude of the missile-produced atmospheric pressure waves to less than the random local atmospheric pressure variations, ambient at the sensor location. These ambient variations due to local atmospheric conditions can have pressure amplitudes in the order of 100 to 500 dynes/cm^2 and frequency spectra similar to missile-produced waves.

We found the most promising approach used to overcome these difficulties, when the position of the infrasonic source is known, to be in the employment of a large number of sensors (microphones, microbargraphs, or pressure transducers) in a linear array, parallel to the infrasonic pressure wave front and the use of real-time correlation techniques. The output of each sensor in the array is combined with the outputs of the others by summation or multiplication. If the sensors are sufficiently spaced in the linear array, the pressure variations caused by wind turbulence at each sensor tend to produce uncorrelated outputs. These uncorrelated outputs can be expected to sum power-wise to NA^2, where N is the number of sensors and A is the output amplitudes; whereas the missile-produced pressure waves arriving in phase at all sensors in the broadside linear array will sum up to $(NA)^2$. The improvement in the signal-to-noise ratio over that of a single sensor will be equal to 10 log $(NA)^2/NA^2$ or in dB equal to 10 log N. The greater the number of sensors in the array, of course, the greater the detection improvement[2] (Fig. 2). In addition, an effective wind screen (Fig. 3) has been devised to inclose each sensor. This screening can result in a further improvement in signal-to-noise ratio of 10 to 30 dB. The amount of improvement is a function of the degree of wind turbulence.

A 1000' broadside linear array of 20 screened and equally spaced microphone sensors have been in operation for several years at a test site in Wayside, N. J. (Fig. 4). In order to verify the direction of arrival and determine the velocity of propagation, a second linear array of 10 sensors is in operation at a site in Middletown, N. J., seven miles north of Wayside. Both arrays are arranged parallel to the acoustic-pressure wave-fronts arriving from Cape Kennedy. These arrays are remotely controlled and the outputs recorded via telephone circuits terminating in a laboratory room. The time of pressure wave arrival from launching time T_o, the wave durations, velocity, and other pertinent data are recorded for each array and each scheduled large missile launching.

Observations

The occasionally observed early arrival (precursor) of infrasonic waves in the time frame of 38 to 43 minutes (Fig. 1) preceding the normal arrival period is of considerable interest to those concerned with acoustic propagation phenomena. Observed amplitudes and durations of these early arrivals are about one-half that of the normal arrival. In the lower atmospheric regions, the absorption coefficient, which is a function of temperature, viscosity, and pressure, is extremely small at these low acoustic frequencies. In the upper atmospheric regions, the absorption coefficient is greater, principally because of the lower atmospheric pressure. This lends support to our hypothesis that upper atmospheric ducting is responsible for the early arriving pressure waves at the jet stream level (30 000 to 40 000'). Apparently, early arrival occurs only when the jet stream flows northward along the east coast from the direction of Florida. We have observed this phenomena during the winter months and it has been confirmed by studies of the tropopause wind analysis stream function charts published daily by the ESSA Environmental Data Service.[3] The velocity of the jet stream as given in these charts plus the velocity of sound at the relevant altitude could account for the early arrival of the infrasonic wave train. This theory is further confirmed by comparing the observed reduced wave amplitudes and durations of the early arrivals with the amplitudes and durations of the normal arrivals. A four-minute film showing the variation of these wind streams over a period of a year will accompany this presentation. These data show that the jet streams prevail in a west to east direction across the U.S.A. in the summer. As late fall occurs, the streams drop to the southern portion of the country with an attendant curvature to the north along the east coast. Figure 5 shows typical west to east summertime patterns, while Figure 6 shows the winter streams approximately parallel to the east coast. The latter is thereby in the correct orientation to account for the precursor, or earlier than normal wave arrival.

Design of Experiment-Statistical Problem Areas

During observations of the early infrasound arrivals from Cape Kennedy, the detection of the weak infrasonic wave trains from long distances is often masked by the effects of wind turbulence, particularly during the windy winter months in the northeastern part of the U. S. The broadside array of many sensors (100 or more) appears to be the most promising solution to the reduction of the effect of wind turbulence. The microphone sensors presently employed require considerable maintenance and therefore large arrays of a hundred or more sensors become rather impractical. At the present time, the problem of further reducing the effects of wind turbulence remains with us. Investigation of two types of infrasonic gradient[4] microphone sensors are currently in progress. The gradient types offer promise of greater reliability, lower costs per unit, and may be less affected by wind turbulence.

Then the problem of site selection for the sensors further broadens the experimental design problem. With but only two or three site instrumentations economically allowed, the question arises as to where the sampling would prove to be most profitable in terms of helpful data. Additionally, the correlation of the sparse experimental data with the voluminous wind data provide problem areas which may be amenable to statistical techniques. As shown in the accompanying film and Figures 5 and 6, the wind data shows distinct summer-winter variations. But, the launching of missiles are relatively few and far between.

To further complicate the real situation, there could be and are, other hypotheses attempting to explain the relative occurrences of the normal and precursor wave trains. For this experimental design consideration, though, we are confining our attention to the one hypothesis, i.e., the jet stream hypothesis. We plan to arrange experiments to confirm, or deny, this hypothesis. Initially, therefore, we must determine how best to proceed in this endeavor that is typical of many such quests in what might be termed macroscopic experimental research involving sporadic and uncontrollable conditions.

References

1. Richard K. Cook, "Strange sounds in the atmosphere, Part 1," Sound, Vol. 1, No. 2, pp. 12-33, May-June 1962.

2. F. B. Daniels, "Noise-reducing line microphone for frequencies below 1 cps," J. Acoust. Soc. Amer., Vol. 31, No. 4, 529-531, April 1959.

3. "Tropopause wind analysis stream function charts," U. S. Department of Commerce, ESSA, Environmental Data Service, Weather Bureau, Silver Springs, Maryland 20910.

4. B. B. Bauer, "Design and measurement of infrasonic gradient microphones and wind screens," J. Acoust. Soc. Amer., Vol. 44, No. 5, pp. 1428-1436, 1968.

Selected Bibliography

1. "Launch Acoustical Environmental Summaries," Published by the John F. Kennedy Space Center, NASA, TR 927, 30 January 1969, Office of the Director of Information Systems.

2. C. T. Johnson and F. E. Hale, "Abnormal sound propagation over the southwestern United States," J. Acoust. Soc. Amer. Vol. 25, No. 4, pp. 642-650, July 1953.

3. William L. Donn, Eric Posmentier, Uri Fehr, N. K. Balachandran, "Infrasound at long range from Saturn V, 1967," Science, Vol. 162, pp. 1116-1120, 6 December 1968.

4. Harold N. Ballard and Mike Izquierdo, "Reduction of Microphone Wind Noise by the Generation of a Proper Turbulent Flow," Environmental Sciences Div. Report ERDA-262, February 1965.

APPENDIX I

Wave Spectra

A theoretical analysis was performed in an attempt to explain mathematically the total process of generation, transmission, and reception of these pressure waves as typified by Figure 1. This theoretical approach is fundamental in nature as it provides understanding of the characteristics to be expected from signals so formulated. It is planned to present this original and extensive theory in a separate publication as yet undetermined.

Of interest, though, to those concerned with the physics is that a close examination of the experimentally obtained signals of Figure 1 show a high correlation with these theoretical findings. Pronounced characteristics that are readily apparent in Figure 1 coincide with the analytical results. This figure shows that between approximately 47 and 53 minutes from launch-time, the normal signal consists of relatively prominent groups of signal spectra having amplitudes clustered at positions located 48, 49, and 50.7 minutes from launch time. It may be further noted that the first group of spectra at 48 minutes is relatively small in amplitude, while the second two groups are contrastingly large.

This clustering and relative amplitudes are in close accord with the theoretical findings, which show an expectation that three major time periods of energies will be experienced in the reception of such pressure signals. The first period is mathematically predicted to consist of

relatively low-level high-frequency transients. Then follows two major groups of pressure spectra considerably larger in amplitude than the initial transient group.

This analysis therefore shows, as experienced experimentally, that one should expect a first-arrival low-level response to the initial missile initiated energy spike. Then the two larger groups of pressure waves follow, as Figure 1 clearly illustrates.

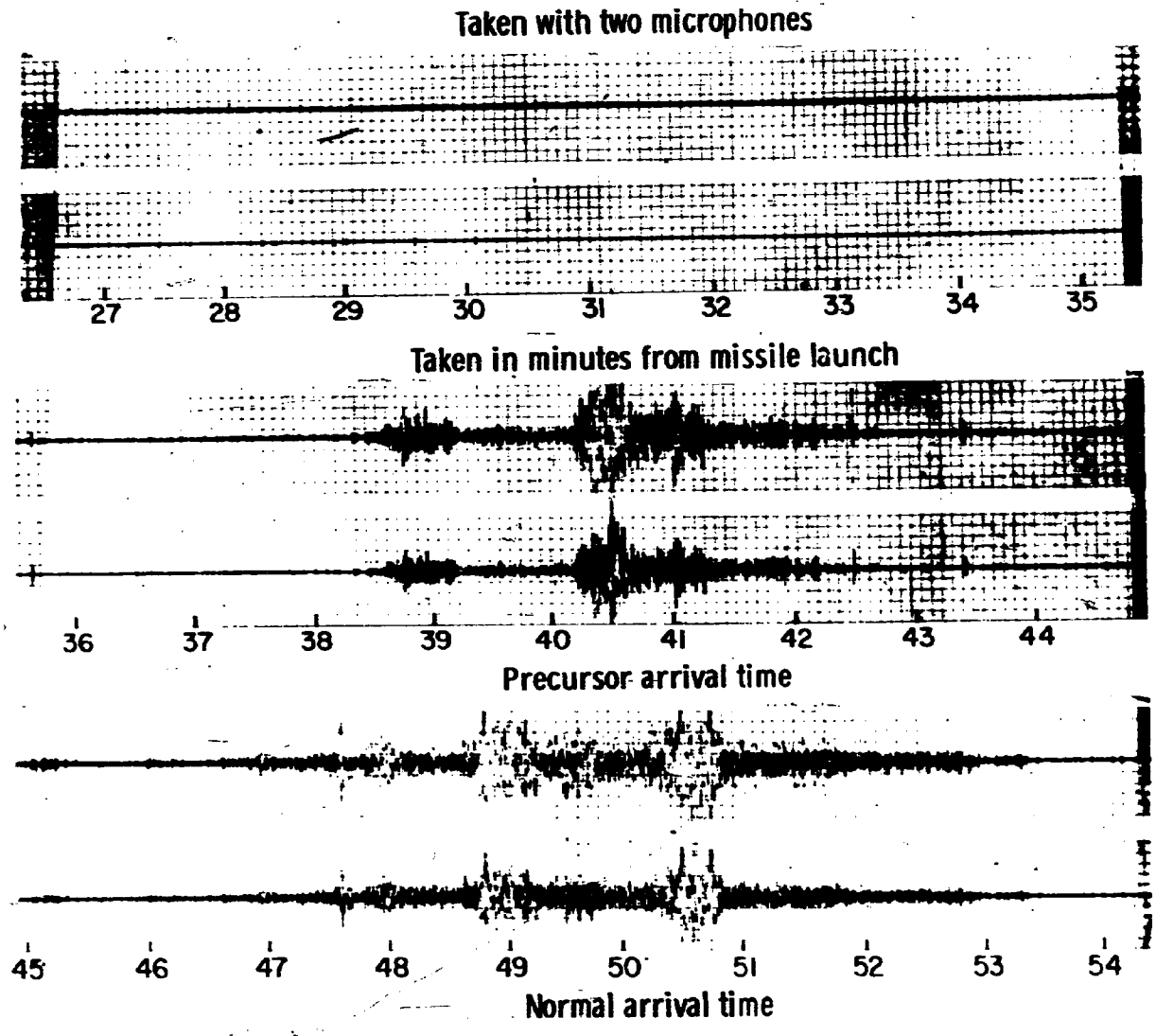


Fig. 1. Infrasound 0.7 to 7 Hz Recording at Cape Hatteras of a Saturn-Apollo Launch From Cape Kennedy

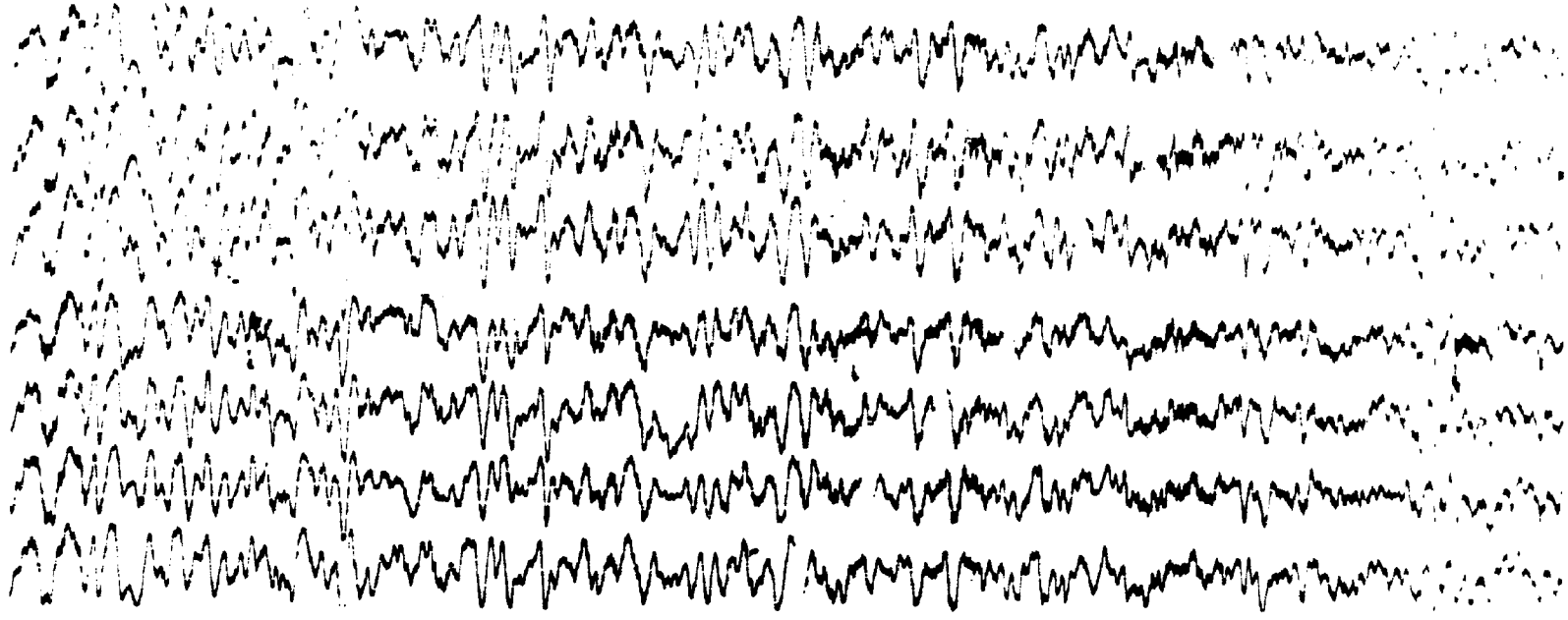


Fig. 2. An example of the correlated output from seven microphones in a broadside array.

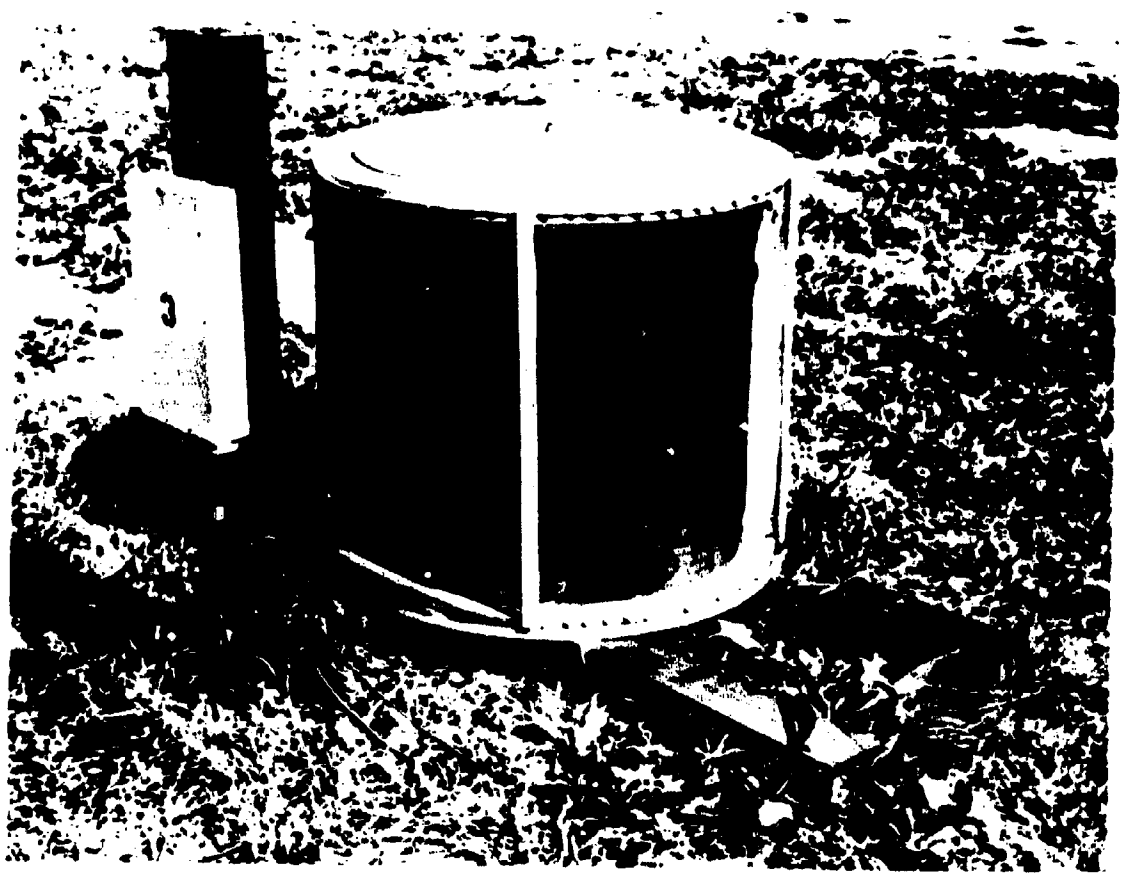


Fig. 3. An experimental design of an infrasonic wind screen inclosure.



Fig. 4. Experimental infrasonic broadside array of 20 microphones.

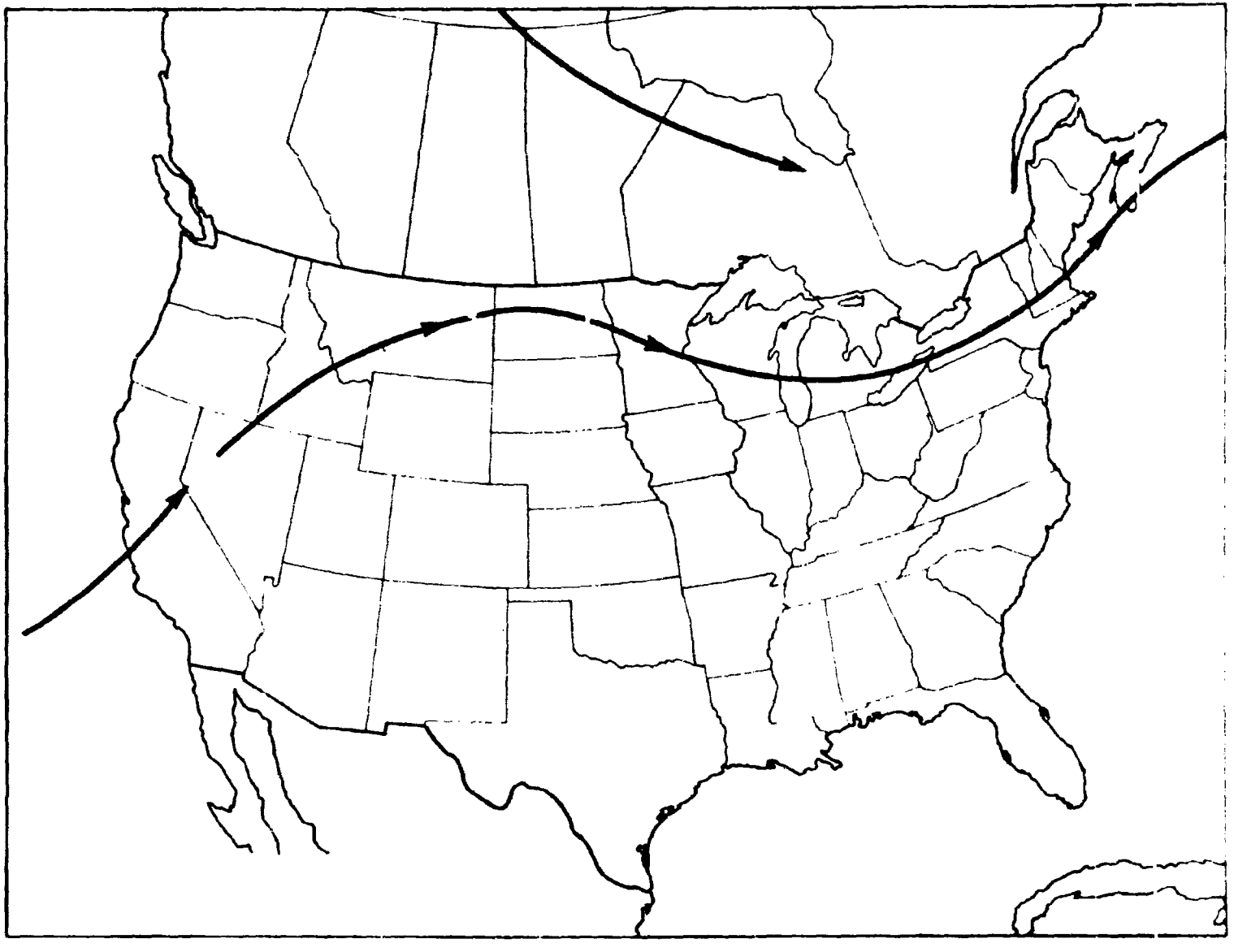


FIG. 5 TYPICAL SUMMERTIME JET STREAM PATTERNS

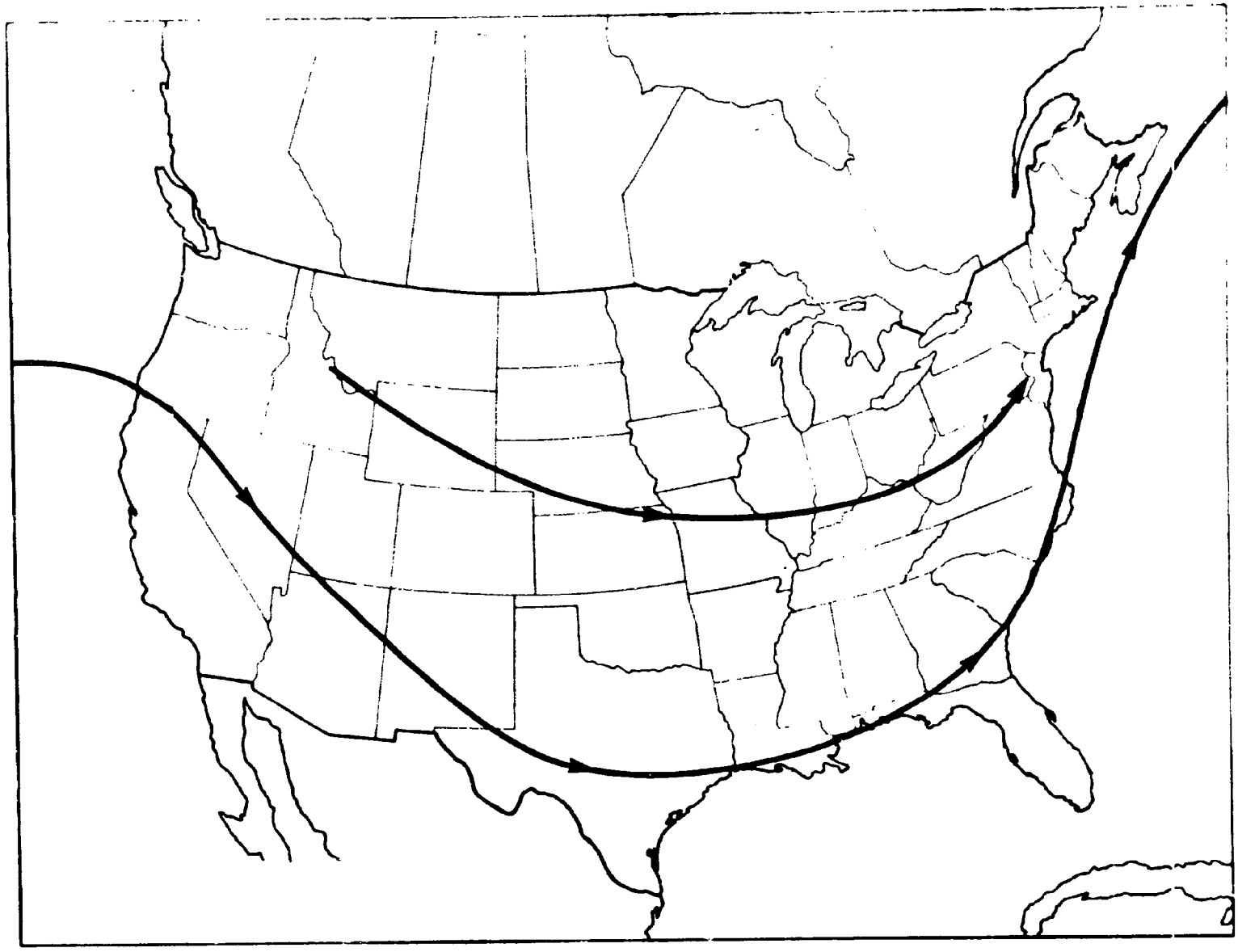


FIG. 6 TYPICAL WINTERTIME JET STREAM PATTERNS

DETERMINING THE FLIGHT RELIABILITY OF AN ANTITANK MISSILE WITH SIDE JETS

BY
R. G. CONARD AND N. R. RICH

U. S. Army Missile Command
Redstone Arsenal, Alabama

ABSTRACT

This paper deals with the probability of catastrophic failure and target miss of an antitank missile in flight resulting from failure of control jets to fire. When two failures occur in sequence, the missile deviates widely from the flight path but is recoverable. When three jets fail to fire in certain sequences, catastrophic missile failure will result.

A complete solution to this problem would involve the determination of the probability of failures of all possible combinations of successes and failures which would result in ground impact. It would consider the random distribution of the location of the target relative to launch position and the distribution of the missile about an average flight path.

A partial solution to this problem has been found in a system of "states," or intervals in the vertical plane. The recursive equations for the probability of lying in each state have been developed and tables of state probabilities for several values of the probability of success for one jet pair.

INTRODUCTION

The specific problem addressed in this paper deals with the probability of failure for a missile which employs jet pulses as a control mechanism and as a means for overcoming gravity.

In this system pairs of jet pulses fire sequentially as the missile, which maintains a fairly constant roll rate, flies toward the target. The pairs of jet pulses are so located about the center of gravity that little or no net torque is applied to the missile body (Figure 1). The firing position of the jets is in a near vertical plane. By varying the angle away from the vertical plane, a side component of thrust may be attained. This side force makes lateral corrections in the flight path upon command (Figure 2).

This article has been reproduced photographically from the authors' manuscript.

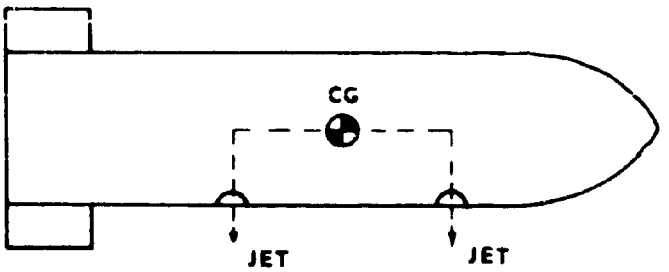


Figure 1. Location of Jets (Side View)

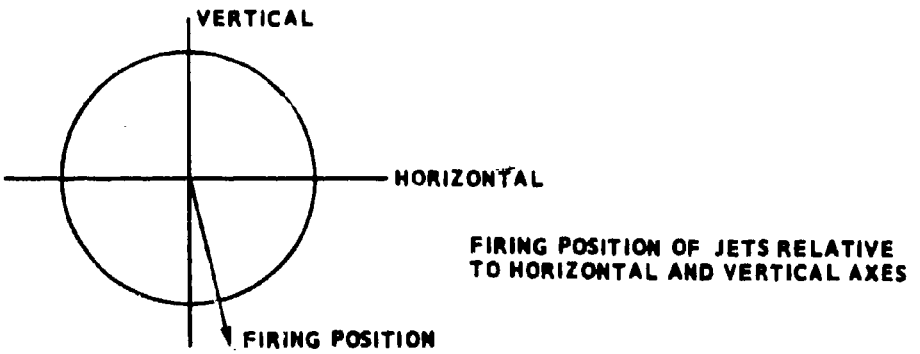


Figure 2. Location of Jet Pulse (Rear View)

The major concern is with the effect on the flight trajectory of a pair of jet pulses failing to fire at the correct time. Since the missile is acted on by gravity, it will drop immediately unless continually sustained by these jet pulses. The lateral dispersion will be affected also; this may or may not be serious depending on the location of the target with respect to location of missile when failure occurs and on the magnitude of such disturbing forces as cross wind. In this study, dispersion in the vertical plane alone is dealt with.

The missile is of a type which is continually commanded from launch to target. In this paper, it is assumed that failure problems do not exist in the guidance system, but occur only in the control mechanism.

So that the missile can be brought back to original flight path, the rate of firing increases as errors become larger. The time interval between jet firings is a discrete fraction of nominal firing intervals. The nominal firing rate will just balance gravitational forces, but a more rapid firing rate will force the missile to move upward.

OBSERVED FAILURE RATE

During the flight test program a known rate of failure has been observed. The failures have been in the circuitry so that either both jets of a pair fire simultaneously or both fail.

In the flight test program it has been observed that out of a total of M commands, L pairs have failed to fire. The estimated failure rate of the circuitry, therefore, has been L/M. There is no reason to believe that the rate will change unless the quality is improved at some additional cost or that the circuit is redesigned.

The effect on the trajectory has been investigated when combinations of failures have occurred in the flight program. Even though there has been a fairly high rate of component failure in the flight test program, the missile has not hit the ground. Although the missile exceeded a desirable control band for an interval of time, it would have missed the target only if the failure had occurred immediately before impact.

For this investigation two types of failures have been defined:

1) The missile deviates from the line-of-sight to the target by more than one unit. In this case it exceeds a desirable control band which will result in a miss if the deviation occurs just prior to impact.

2) The missile deviates downward three units and will impact the ground. This case is a catastrophic failure and can never recover or hit the target.

COMBINATIONS OF FAILURE

The first approach considered for this problem was that of determining the probability of certain combinations of failure.

If it is assumed that there are N jet pulse pairs, each with probability p of success and corresponding probability of failure q = 1 - p and that the pairs fail independently of one another, the probability of no failures will be p^N. The following recursive equations hold:

1) Probability of no repeated failure (FF) in a sequence of N - (Table 1)

$$A_N = pA_{N-1} + pqA_{N-2}, \quad A_0 = 1, \quad A_1 = 1 .$$

In a system of N consecutive independent subsystems, each with probability p of success, A_N is the probability of no repeated failures.

P	.5000	.6000	.7000	.8000	.9000	.9500	.9750	.9900	.9950
N									
1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	.7500	.8400	.9100	.9600	.9900	.9975	.9993	.9999	.9999
3	.6250	.7440	.8470	.9280	.9810	.9951	.9987	.9998	.9999
4	.5000	.6480	.7840	.8960	.9720	.9927	.9981	.9997	.9999
5	.4062	.5673	.7266	.8652	.9630	.9903	.9975	.9996	.9999
6	.3281	.4959	.6733	.8355	.9542	.9880	.9969	.9995	.9998
7	.2656	.4337	.6239	.8069	.9455	.9856	.9963	.9994	.9998
8	.2148	.3792	.5781	.7792	.9368	.9833	.9957	.9993	.9998
9	.1738	.3316	.5357	.7524	.9282	.9809	.9951	.9992	.9998
10	.1406	.2900	.4964	.7266	.9197	.9786	.9945	.9991	.9997
11	.1137	.2536	.4599	.7017	.9113	.9762	.9939	.9990	.9997
12	.0920	.2217	.4262	.6776	.9029	.9739	.9932	.9989	.9997
13	.0744	.1939	.3949	.6543	.8946	.9716	.9926	.9988	.9997
14	.0602	.1695	.3659	.6319	.8864	.9693	.9920	.9987	.9996
15	.0487	.1482	.3391	.6102	.8783	.9669	.9914	.9986	.9996
16	.0394	.1296	.3142	.5893	.8703	.9646	.9908	.9985	.9996
17	.0318	.1133	.2911	.5690	.8623	.9623	.9902	.9984	.9996
18	.0258	.0991	.2698	.5495	.8544	.9600	.9896	.9983	.9995
19	.0208	.0867	.2500	.5307	.8465	.9578	.9890	.9982	.9995
20	.0168	.0758	.2316	.5124	.8388	.9555	.9884	.9981	.9995
21	.0136	.0663	.2146	.4949	.8311	.9532	.9878	.9980	.9995
22	.0110	.0579	.1989	.4779	.8235	.9509	.9872	.9979	.9994
23	.0089	.0507	.1843	.4615	.8159	.9486	.9866	.9978	.9994
24	.0072	.0443	.1708	.4456	.8084	.9464	.9860	.9977	.9994
25	.0058	.0387	.1582	.4303	.8010	.9441	.9854	.9976	.9994

Table 1.

2) Probability of at least three failures in a row (FFF) in a sequence of N = (Table 2)

$$C_N = pC_{N-1} + qpC_{N-3} + q^2, \quad C_1 \quad C_2 \quad C_3 \quad \text{[illegible]}$$

3) Probability of at least one pair of failure (FF) but no subsequence with three or more failures in a row = (Table 3)

$$W_N = 1 - A_N - C_N$$

Trajectory simulation implies that when any combination of three out of four consecutive pairs or three out of five consecutive pairs of jets fail to fire in the right sequence, catastrophic failure results. Some combinations of three failures in six consecutive pairs result in marginal flight. A sequence of seven where the first, seventh, and any one other pair in between fail will not result in catastrophic failure. In fact, if every fifth jet pair failed to fire, the missile would recover but the dispersion about the desired flight path would be large.

This approach was abandoned since the important missile position could be found only by taking a particular combination of failures and running a trajectory simulation. The approach did not provide direct information on the probability of missile failure. For example, the missile would fail catastrophically if failures and successes alternated (SFSFSF...), but this case would be included in the calculation of A_N.

DISCRETE STATE STOCHASTIC MODEL OF VERTICAL DEVIATIONS

The next approach taken involved a simplified model of the vertical deviation from an average or nominal trajectory.

The average trajectory was computed by a least squares fit of the trajectory when all jets are firing in the proper sequence and at expected time intervals. The least squares fit established an average or expected trajectory about which the missile oscillated (Figure 3). In making the fit to the average trajectory the data immediately following the time when the failure occurs were discarded. The data were again used when the firing rate once more became normal.

At the time of firing of the N^{th} jet pair, the missile can be considered in one of seven "states." Each state is an interval in the vertical plane.

In a system of N independent consecutive subsystems, each with probability p of success, C_N is the probability of at least three failures in a row.

P / N	.5000	.6000	.7000	.8000	.9000	.9500	.9750	.9900	.9950
1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3	.1250	.0640	.0270	.0080	.0010	.0001	0.0000	0.0000	0.0000
4	.1875	.1024	.0459	.0144	.0019	.0002	0.0000	0.0000	0.0000
5	.2500	.1408	.0648	.0208	.0028	.0003	0.0000	0.0000	0.0000
6	.3125	.1792	.0837	.0272	.0037	.0004	0.0000	0.0000	0.0000
7	.3671	.2151	.1023	.0335	.0045	.0005	0.0000	0.0000	0.0000
8	.4179	.2496	.1201	.0398	.0054	.0007	0.0000	0.0000	0.0000
9	.4648	.2826	.1377	.0461	.0063	.0008	.0001	0.0000	0.0000
10	.5078	.3141	.1551	.0523	.0072	.0009	.0001	0.0000	0.0000
11	.5473	.3442	.1723	.0585	.0081	.0010	.0001	0.0000	0.0000
12	.5837	.3730	.1887	.0646	.0090	.0011	.0001	0.0000	0.0000
13	.6171	.4006	.2053	.0707	.0099	.0013	.0001	0.0000	0.0000
14	.6479	.4269	.2209	.0768	.0108	.0014	.0001	0.0000	0.0000
15	.6762	.4521	.2366	.0828	.0117	.0015	.0001	0.0000	0.0000
16	.7022	.4762	.2519	.0888	.0126	.0016	.0002	0.0000	0.0000
17	.7261	.4992	.2669	.0948	.0135	.0017	.0002	0.0000	0.0000
18	.7481	.5212	.2817	.1007	.0144	.0019	.0002	0.0000	0.0000
19	.7684	.5422	.2961	.1065	.0153	.0020	.0002	0.0000	0.0000
20	.7870	.5623	.3102	.1124	.0162	.0021	.0002	0.0000	0.0000
21	.8041	.5816	.3241	.1182	.0171	.0022	.0002	0.0000	0.0000
22	.8198	.6000	.3377	.1239	.0179	.0023	.0003	0.0000	0.0000
23	.8343	.6175	.3510	.1296	.0188	.0024	.0003	0.0000	0.0000
24	.8476	.6343	.3640	.1353	.0197	.0026	.0003	0.0000	0.0000
25	.8599	.6504	.3768	.1410	.0206	.0027	.0003	0.0000	0.0000

Table 2.

In a system of N independent subsystems that are employed consecutively, each with probability p of success, W_N is the probability of at least one pair of consecutive failures but no more than two failures in a row.

P	.5000	.6000	.7000	.8000	.9000	.9500	.9750	.9900	.9950
N									
1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	.3750	.2560	.1530	.0720	.0190	.0048	.0012	.0001	0.0000
3	.2500	.1920	.1260	.0640	.0180	.0047	.0012	.0001	0.0000
4	.3125	.2496	.1701	.0896	.0261	.0070	.0018	.0002	0.0000
5	.3437	.2918	.2085	.1139	.0341	.0092	.0024	.0003	0.0000
6	.3593	.3248	.2429	.1372	.0420	.0115	.0030	.0004	.0001
7	.3671	.3511	.2739	.1595	.0498	.0137	.0035	.0005	.0001
8	.3671	.3711	.3017	.1809	.0576	.0159	.0041	.0006	.0001
9	.3613	.3857	.3264	.2013	.0653	.0182	.0047	.0007	.0001
10	.3515	.3958	.3484	.2209	.0729	.0204	.0053	.0008	.0002
11	.3388	.4021	.3679	.2397	.0804	.0226	.0059	.0009	.0002
12	.3242	.4051	.3850	.2576	.0879	.0248	.0065	.0010	.0002
13	.3083	.4054	.4000	.2748	.0953	.0270	.0071	.0011	.0002
14	.2918	.4034	.4130	.2912	.1026	.0292	.0077	.0012	.0003
15	.2750	.3995	.4242	.3068	.1098	.0314	.0083	.0013	.0003
16	.2583	.3941	.4337	.3218	.1170	.0336	.0089	.0014	.0003
17	.2419	.3873	.4418	.3360	.1241	.0358	.0095	.0015	.0003
18	.2260	.3796	.4484	.3497	.1311	.0380	.0100	.0016	.0004
19	.2107	.3710	.4538	.3627	.1380	.0401	.0106	.0017	.0004
20	.1960	.3617	.4580	.3750	.1449	.0423	.0112	.0018	.0004
21	.1821	.3520	.4611	.3868	.1517	.0445	.0118	.0019	.0004
22	.1690	.3420	.4633	.3981	.1584	.0466	.0124	.0020	.0005
23	.1567	.3317	.4646	.4087	.1651	.0488	.0130	.0021	.0005
24	.1450	.3212	.4651	.4189	.1717	.0509	.0136	.0022	.0005
25	.1342	.3107	.4649	.4285	.1782	.0530	.0142	.0023	.0005

Table 3.

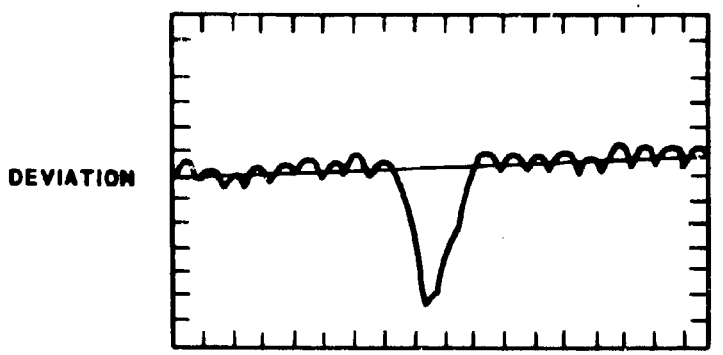


Figure 3. Vertical Deviation from Average Flight Path

If there are no jet failures, each pair will fire when the missile is in state 0 (Figure 4). The missile is considered to be in state 0 when the first pair is fired. Let X0(N) denote the probability of being in state 0 at the time of firing for N^{th} pair, then X0(1) = 1.

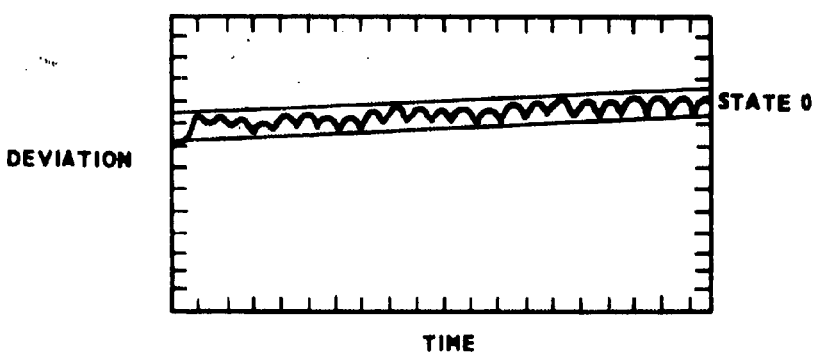


Figure 4. Nominal Trajectory (No Jet Failures)

When the first failure occurs, the missile will drop into state 2 (Figure 5). If the failure is followed by a success, the thrust of the jet pair does little more than arrest the missile in its fall; thus, the missile remains in state 2. The next success will bring the missile to state 1 and the next success brings the missile to state 0. For example, for the sequence S, S, F, S, S, S, the states will be 0, 0, 2, 3, 1, 0.

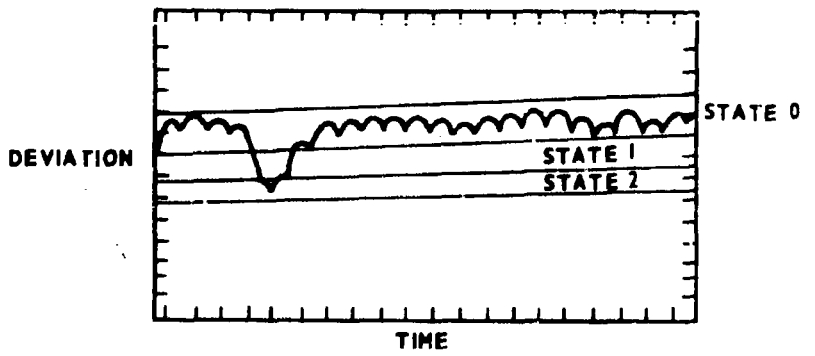


Figure 5. Vertical Deviation (One Jet Failure)

Two jet failures in a row after a series of successes will drop the missile into state 4 (Figure 6). If the next firings are successes, the missile will stay in state 4 one time and then climb up one state at a time. Thus, the sequence S,S,F,F,S,S,S,S,S, will result in the states 0,0,2,4,4,3,2,1,0. For the sequence S,S,F,S,F,S,S,S,S,S,S, as shown in Figure 7, the state sequence is 0,0,2,2,4,4,3,2,1,0.

Three failures in a row will place the missile in state 6 (Figure 8). This is a captive state representing the catastrophic failure of hitting the ground. Once in state 6, the missile remains in this state.

The recursive equations for this model are given in Table 4. Here XI(N) is the probability of being in state 1 at the time of the N^{th} pair firing. These were developed through the following type of reasoning: The missile can be in

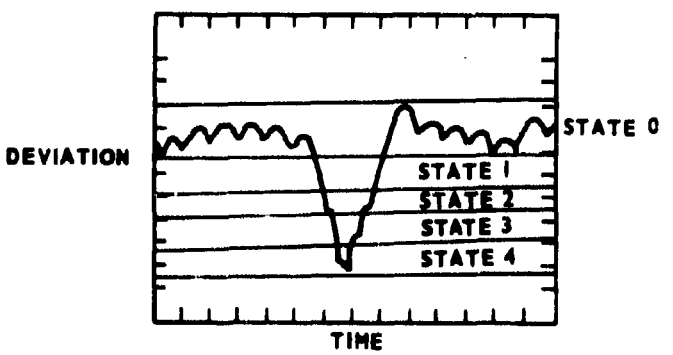


Figure 6. Vertical Deviation (Two Jet Failures)

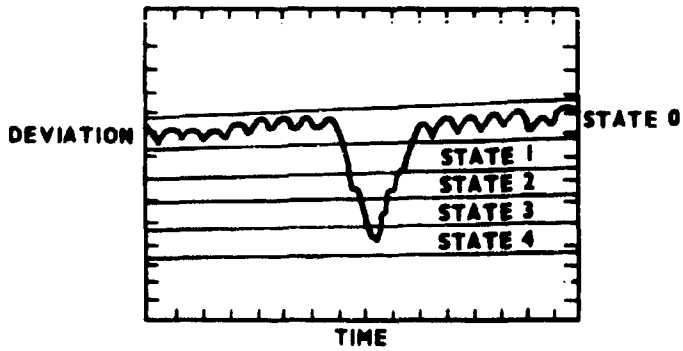


Figure 7. Vertical Deviation (Failure-Success-Failure Sequence)

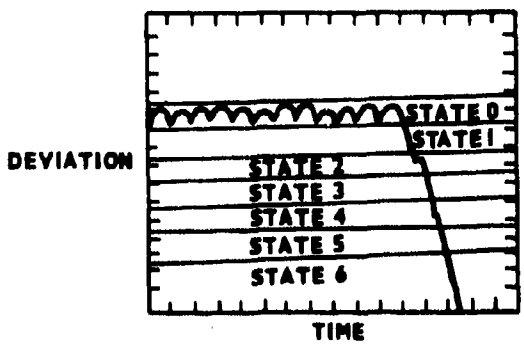


Figure 8. Vertical Deviation (Three Jet Failures)

(1) $X0(N) = pX0(N-1) + pX1(N-1)$

(2) $X1(N) = qp^2X0(N-3) + qp^1X1(N-1) + qp^1X2(N-5) + qp^5X3(N-6)$

(3) $X2(N) = qX0(N-1) + qpX0(N-2) + qp^2X1(N-3) + qp^3X2(N-4) + qp^4X3(N-5)$

(4) $X3(N = qX1(N-1) + qpX1(N-2) + qp^2X2(N-3) + qp^3X3N-4)$

(5) $X4(N) = qX2(N-1) + qpX2(N-2) + qp^2X3(N-3)$

(6) $X5(N) = qX3(N-1) + qpX3(N-2)$

(7) $X6(N) = X6(N-1) + qX5(N-1) + qX4(N-1)$

Table 4.

state 0 at the time of the N^{th} firing only if it were in state 0 or state 1 at the previous firing and that firing was successful. Thus, Prob (state 0 at time N = Prob (state 0 at time N - 1) · Prob (success) + Prob (state 1 at time N - 1) · Prob (success), or X0(N) = pX0(N - 1) + PX1(N - 1). It is assumed that the probability of success is independent of the state. The initial conditions for this set of recursive equations are given in Table 5. The values of the state probabilities for N = 1, ..., 25 and p = .50, .90, .95, .99 are shown in Tables 6 through 9.

States \ N	1	2	3	4	5	6
0	1	p	p^2	p^3	$p^4 + p^3q$	$p^5 + 2p^4q$
1				p^2q	p^3q	p^4q
2		q	2pq	$2p^2q$	$2p^3q$	$2p^4q + 2p^3q^2$
3					$2p^2q^2$	$4p^3q^3$
4			q^2	$3pq^2$	$4p^2q^2$	$4p^3q^2$
5						$2p^2q^3$
6				q^3	$4pq^3 + q^4$	$q^5 + 5pq^4 + 8p^2q^3$

Table 5.

The probability of catastrophic failure (hitting the ground) is the probability of being in state 6. If a target were of such a size as to cover states 0, 1, and 2 exactly, the probability of hitting the target would be the probability of being in any of these three states.

There are several drawbacks to this approach. The target might not cover an exact number of states so that the distribution within states is of importance. The range of the target should be considered to be variable: thus, the time between "times of firing" needs to be taken into account. As the pairs are fired, the number of available jet pairs decreases. Thus, if a failure occurs near the end of flight, the time until another unused pair is in position to fire is longer than it would be at the beginning of flight and the missile might drop further. For this reason, the model does not provide a uniformly good approximation to reality.

N	P	X0	X1	X2	X3	X4	X5	X6
1	.500	1.000000	.000000	.000000	.000000	.000000	.000000	.000000
2	.500	.500000	.000000	.500000	.000000	.000000	.000000	.000000
3	.500	.250000	.000000	.500000	.000000	.250000	.000000	.000000
4	.500	.125000	.125000	.250000	.000000	.375000	.000000	.125000
5	.500	.125000	.062500	.125000	.125000	.250000	.000000	.312500
6	.500	.093750	.031250	.125000	.125000	.125000	.062500	.437500
7	.500	.062500	.031250	.125000	.062500	.093750	.093750	.531250
8	.500	.046875	.046875	.078125	.093750	.109375	.062500	.625000
9	.500	.042968	.023437	.050781	.050781	.085937	.085156	.710937
10	.500	.033203	.013671	.048828	.044921	.052734	.085156	.771484
11	.500	.023437	.013671	.043945	.026367	.041992	.085156	.815429
12	.500	.018554	.013671	.029785	.019042	.040527	.024414	.854003
13	.500	.016113	.009033	.021240	.019531	.031494	.016113	.886474
14	.500	.012573	.005981	.019042	.016235	.021362	.014526	.910278
15	.500	.009277	.005493	.014174	.010620	.017211	.013000	.928222
16	.500	.007385	.004543	.011596	.008087	.015289	.009368	.94[illegible]328
17	.500	.006164	.003479	.008682	.007446	.011871	.006698	.955057
18	.500	.004821	.002494	.007415	.006011	.008567	.005744	.964942
19	.500	.003658	.002166	.006088	.004230	.006889	.004867	.972099
20	.500	.002912	.001538	.004526	.003297	.005828	.003618	.977977
21	.500	.002375	.001348	.003479	.002853	.004536	.002706	.982701
22	.500	.001862	.001010	.002882	.002270	.003399	.002251	.986322
23	.500	.001436	.000847	.002323	.001672	.002723	.001848	.989149
24	.500	.001142	.000696	.001767	.001317	.002238	.001494	.991434
25	.500	.000919	.000524	.001376	.001098	.001748	.001076	.993255

Table 6.

N	P	X0	X1	X2	X3	X4	X5	X6
1	.900	1.000000	.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	.900	.900000	.000000	.100000	0.000000	0.000000	0.000000	0.000000
3	.900	.810000	.000000	.180000	0.000000	.010000	0.000000	0.000000
4	.900	.729000	.081000	.162000	0.000000	.027000	0.000000	.001000
5	.900	.729000	.072900	.145800	.016200	.032400	0.000000	.003700
6	.900	.721710	.065610	.145800	.029160	.029160	.001620	.006940
7	.900	.708588	.065610	.157464	.026244	.027702	.004374	.010018
8	.900	.696778	.076763	.153527	.024275	.030180	.005248	.013225
9	.900	.696187	.074401	.149393	.026572	.031686	.004789	.016768
10	.900	.693530	.071744	.149334	.029229	.030882	.004842	.020436
11	.900	.688747	.071744	.151620	.028219	.030345	.005314	.024008
12	.900	.684442	.074040	.150233	.027502	.030754	.005452	.027574
13	.900	.682634	.073222	.148726	.027894	.031036	.005289	.031195
14	.900	.680271	.072253	.148304	.028397	.030679	.005264	.034827
15	.900	.677272	.072036	.148432	.028041	.030443	.005350	.038422
16	.900	.674378	.072364	.147686	.027758	.030450	.005359	.042001
17	.900	.672069	.071963	.146891	.027765	.030427	.005299	.045582
18	.900	.669629	.071508	.146377	.027802	.030252	.005274	.049155
19	.900	.667023	.071252	.145994	.027634	.030106	.005279	.052708
20	.900	.664447	.071124	.145404	.027492	.030022	.005265	.056246
21	.900	.662020	.070831	.144798	.027406	.029931	.005235	.059775
22	.900	.659566	.070518	.144266	.027357	.029804	.005214	.063292
23	.900	.657076	.070258	.143766	.027218	.029684	.005200	.066794
24	.900	.654601	.070029	.143219	.027104	.029580	.005182	.070282
25	.900	.652167	.069760	.142668	.027009	.029475	.005160	.073759

Table 7.

4								.000125
5								.000481
6								.000932
7								.001472
8								.001922
9								.002301
10								.002704
11								.003284
12								.003773
13								.004266
14								.004760
15								.005253
16								.005746
17								.006239
18								.006732
19								.007225
20								.007717
21								.008209
22								.008701
23								.009193
24								.009684
25	.950	.839604	.043751	.038031	.008517	.008978	.000840	.010176

Table 8.

N	P	X0	X1	X2	X3	X4	X5	X6
1	.990	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
2	.990	.990100	0.000000	.019800	0.000000	0.000000	0.000000	0.000000
3	.990	.980100	0.000000	.019800	0.000000	.000100	0.000000	0.000000
4	.990	.970299	.009801	.019602	0.000000	.000297	0.000000	.000001
5	.990	.970299	.009702	.019405	.000196	.000392	0.000000	.000003
6	.990	.970201	.009605	.019405	.000388	.000388	.000001	.000007
7	.990	.970099	.009605	.019596	.000384	.000386	.000005	.000011
8	.990	.969819	.009795	.019590	.000381	.000390	.000007	.000015
9	.990	.969818	.009791	.019583	.000385	.000393	.000007	.000019
10	.990	.969813	.009786	.019583	.000390	.000393	.000007	.000023
11	.990	.969804	.009786	.019569	.000390	.000393	.000007	.000027
12	.990	.969795	.009792	.019588	.000390	.000393	.000007	.000031
13	.990	.969791	.009791	.019588	.000390	.000393	.000007	.000035
14	.990	.969787	.009791	.019588	.000390	.000393	.000007	.000039
15	.990	.969783	.009791	.019588	.000390	.000393	.000007	.000043
16	.990	.969779	.009791	.019588	.000390	.000393	.000007	.000047
17	.990	.969775	.009791	.019588	.000390	.000393	.000007	.000051
18	.990	.969771	.009791	.019588	.000390	.000393	.000007	.000055
19	.990	.969768	.009791	.019588	.000390	.000393	.000007	.000059
20	.990	.969764	.009791	.019586	.000390	.000393	.000007	.000063
21	.990	.969760	.009791	.019588	.000390	.000393	.000007	.000067
22	.990	.969756	.009791	.019588	.000390	.000393	.000007	.000071
23	.990	.969752	.009791	.019587	.000390	.000393	.000007	.000075
24	.990	.969748	.009791	.019587	.000390	.000393	.000007	.000079
25	.990	.969744	.009791	.019587	.000390	.000393	.000007	.000083

Table 9.

One of the limitations of this approach is the difficulty of relating the position of the missile in one of the various states to its position along the trajectory since the rate of firing of the side thruster is not always a constant. In order to compensate exactly for gravity there must be a predictable amount of impulse imparted to the missile during a given time interval. If one side jet pair fails to fire at the proper time and the missile develops a vertical acceleration due to gravity, the acceleration must be corrected and the missile brought back to its appropriate flight path by more rapid firing of the side thrusters.

If the success-failure sequence is known, the number of successful side jet firings required to bring the missile back to its original flight path can be determined by flight simulation. For the simpler cases this has been done and a time relation between state position and the required number of successful firing results has been determined explicitly. The more complicated sequences would present some difficulty especially if the supply of available side jets is exhausted. On the average, however, a missile in a given state whose vertical acceleration has been arrested will have experienced a predictable number of successes within a given time.

QUESTIONS TO BE CONSIDERED

The following questions are submitted for consideration by the panel.

1) The distribution of the missile deviation about an average flight path will involve the dispersion of a properly functioning missile as well as the dispersion resulting from jet failure. A technique is required for combining the two sources of error.

2) Since the target is at a random position relative to the launcher, the probability of hit may be related to the range of the target. How should the target range be included as an error source?

3) Toward the end of flight, the number of jets to be fired becomes limited. Therefore, if a failure occurs late in the flight, there might be empty rows of jets. This would result in a delay in firing after the failure until a pair of jets would be in the proper position to be fired. How should this problem be handled?

A PROBABILITY APPROACH TO CATASTROPHIC THREAT

Clifford J. Maloney*
Bethesda, Maryland

I. INTRODUCTION. Prior to World War II statistical methods were viewed as a body of techniques appropriate to scientific research and to a limited range of other activities, primarily insurance. The war and the immediate post war years saw an enormous expansion of the application of statistics to quality control of incoming supplies, operations research as a replacement for trial and error in the choice of operational organization and technique, and sample surveys as a substitute for complete canvass for obtaining socio-economic data needed in decision-making. Developments more purely mathematical in nature, including linear programming, development of the electronic computer, finite mathematics and the like served to help foster an increasing mathematical, probabilistic, and statistical approach to decision-making in all aspects of human activity.

Perhaps the one aspect which has so far not attracted the attention of scholars has been the anticipation, prevention, and/or amelioration of unexpected but costly contingencies; an activity the businessman has long known as "putting out fires." But despite the fact that, superficially at least, "every case is different," the very fact that in the aggregate these occurrences are both frequent and costly, suggests that at the very least we can attempt to codify the techniques by which the successful practitioners cope with these pheonomena, however diversified, and that a possibility as well exists that some underlying structure will be discovered.

II. THE NATURE OF CATASTROPHE. The Statistical Department of the Metropolitan Life Insurance Company defines a catastrophe as any accident in which 25 or more lives are lost. Two incidents will show that an event may be a disaster despite the fact that no loss of life occurred or was even threatened. The first was the Pueblo incident where the chief cost was the severe, and continuing, but unmeasurable contribution to the threat to peace. The second was the public reaction to the characterization of a military briefing as a "brainwashing." One suspects that, two years later, the reaction would be very different. A first attempt to list the characteristics of a disaster is given in Figure 1. The ensuing six figures give a few ill-chosen examples under each of six disaster characteristics to which the reader can undoubtedly add an indefinite number of more suitable examples.

My claim that while, in fact, unexpected the disastrous event was "predictable" can be illustrated by the seizure of the Pueblo. All parties recognized that such an act was conceivable. The difference lay exclusively in their assessments of its likelihood. All I mean by

*Views expressed herein are those of the author and not to be construed as official.

predictable is that the catastrophic event can be visualized prior to occurrence. Contamination of the moon by earth microbes, or of the earth by organisms returning with the astronauts is another such case. Earth rupture, tsunami, or earthquakes are all conceivable consequences of underground nuclear tests. It is unnecessary to argue that the inconceivable never happens in order to make it clear that most catastrophes are of an ordinary garden variety; that their catastrophic nature lies in their (1) severity and (2) unexpectedness, not in their absolute novelty.

That catastrophe severity is as much psychological as real, may not be at once apparent. Such loss (other than one ship and the suffering of the men and their families) as was sustained in the Pueblo incident, lies in the threat to world peace - i.e., is psychological. The Dugway sheep kill created a great public outcry - yet the anticipation that several workmen will lose their lives when the Washington, D. C. subway is built, won't raise an eyebrow. The brainwashing episode was an extreme example.

III. DECISION-MAKING DOMAINS. To decide to ignore or to anticipate a threat is to make a decision, and decision-making is one of the more recent and more prestigious applications of probability. If, as claimed, catastrophic threat has not hitherto been recognized as an appropriate application of decision theory, there must be something about the material that conceals its appropriateness for the technique. We can at once expose this unconscious assumption, show how to remove it, and show the absolute necessity of doing so by reviewing the circumstances in which the application of probability is currently fully recognized. First, partly as a horrible example, partly because all problems not treated with a formal algorithm are necessarily so treated, and partly to show the absolutely fundamental role of probability in all of life's decisions, the universally applied method of common sense will be listed as a specific method and discussed. Its characteristics are given in Figure 8. Many, if not most, of life's decisions hinge on an estimate of probability. To portray this rule in a military context, I have extracted certain sentences from the chapter on the Anzio Campaign from the book COMMAND DECISIONS, published by the Office of the Chief of Military History in 1959, and present them in the appendix. Figure 9 is a map of the theater involved. By reading these extracts, which are limited to only those sentences in which a probabilistic concept is essential, in order one can gain a picture of the campaign almost as complete as by reading the full text.

Two further points should also be clear. The various generals assessed the probabilities very differently. In particular, General Mark Clark feared over-precipitate advance because of his one-time experience at Salerno. The English authorities seemed to feel that the German Cassino Line would crumble the moment the Allies had gained a foothold. No Allied commander seems to have contemplated the possibility of an unopposed landing - which in the event is what happened. What I

seek to infer from this example is that up to the time of World War II, no formalized procedure existed for dealing with the probability component of command decision-making and that that lack proved costly.

The application of probability considerations first occurred, and by now is well established, in the situations which grew out of, or which can be logically reduced to, games of chance. In practice, these cases arise in designed experiments or in sample surveys where the sample is drawn by "probability sampling." This situation is characterized in Figure 10.

In a closely related but distinguishable situation, characterized in Figure 11, the sample is not drawn from a preformed population but from a growing one. Both of these cases (Figures 10 and 11) involve probabilities and sample sizes in a middle range.

In the next situation (Figure 12) we are dealing with a case where, again, probabilities are, or at lease often are, in the middle range, but sample sizes are usually not even defined. During World War II, the Weather Bureau developed a so-called 30-year series of Northern Hemisphere maps to enable military forecasters to make forecasts by matching the current weather map with an earlier situation, on the assumption that the succeeding weather should also be a repetition. Unfortunately, matching is never exact. At this point a new factor is brought into play, which goes back to the earliest recognition of a role of chance assessment in decision-making, but which was relatively neglected in the century preceeding World War II, while an objective foundation for probability seemed to be proving adequate. This new factor was the concept of personal probability, not as a quantity having external reality, but as a measure of subjective mind state whether shared by several minds or peculiar to one. It will be argued below that this factor has been incorrectly apprehended. Here it is sufficient to note that on this foundation a super-structure for dealing with forecasting, technological and otherwise, has been elaborated and is being increasingly widely adopted. A good introduction is afforded by Bright (1968) and the references there included.

Finally, we come to the situation contemplated in this paper. Like the case above, sample size is essentially non-existent. But here the probabilities are so low that no hope exists of accumulating a fund of cases, sufficiently "similar" to provide more than the most tenuous basis for probability quantitation even as a belief state. This is the case described in Figure 13. The essential difference between this case and that of Figure 12 is that, whereas in the latter, probabilities are fairly high, here they are low. For example, stock market, weather or weapons system forecasting cannot successfully be based on simple sample statistics, as for instance is done when the number of telephones or children per household, or the weight gain due to a ration additive is

assayed in a controlled trial; still the market opens, the next day comes and brings with it an after the fact verifiable weather outcome, and whatever weapon system is chosen for development, five to ten years justifies the decision or reveals its invalidity. But catastrophic threat may or may not come to pass. The air defense of the United States against a bomber attack by the Soviets was incredibly expensive. Was it a success? No such attack occurred, but would one have had one if the defense been less effective or less costly? We will never know. The current debate over ABM deployment would not continue a day if all concerned were agreed that within the decade of the seventies one or more technologically effective nuclear missiles were highly likely to be launched against one or more of our cities. The ABM proponents, quite as much as the opponents, not only hope that the event will not occur, but assess its likelihood as low. The disagreement is entirely over how low is low enough.

IV. PROBABILITY FOUNDATIONS. Three times in the past the development of a new foundation for quantitative probability has resulted in a major success in the increase in understanding and control of the real world. The two "objective" foundations, the a priori or necessary of Cardan Galileo, Fermat and Pascal, and the frequency foundations of Graunt and Halley, developed about the same time, proved adequate as a foundation for all applications of probability prior to World War I . They were abstracted to a formal mathematical calculus at the end of that period by Kolmogoroff and others.

For the most part the necessary and the frequency foundation for probability are complementary. There is, however, one aspect in which they give conflicting counsel, the problem of outliers. On the frequency definition an outlier is meaningless - for that approach takes experience as given. On the contrary, the a priori or necessary approach rejects experience whenever it is in conflict with doctrine. Adherents of both approaches, being sensible men, adhere neither to the one extreme nor to the other. Nor, indeed, do they in practice achieve a comfortable compromise.

As suggested above, notions of "degree of belief" as a foundation of probability theory, while studies in the years following World War I, were not widely employed until after World War II. Most adherents of personal probability adopt a concensus form of the approach. It is not what one believes, but what one ought to believe that matters. DeFinetti and Savage, however, go all the way and allow everyone his own private degree of belief, divorcing the concept entirely from external world events.

A fifth foundation for the concept of probability underlies, if somewhat subconsciously, most work on curve fitting and indeed comes close to formal enunciation especially in the older least squares literature. Here every value of every quantitative property is assumed to be precisely determined, but as a function of a literally infinite number of independent or predictor variables, say

$$X = \alpha + \beta + \sigma + \delta + \zeta + \ldots .$$

One achieves a chance structure by neglecting most of these latter. Then the precise value of the predicted variable depends on the contribution of the neglected variables, which being unknown makes the predicted value a chance one. Each of the other theories can be obtained by making assumptions about this formula. The a priori definition assumes that the relative frequency of possible values of X can be inferred because the values of the independent variables are known, whenever individually important. The frequency definition conversely finds the values only of a fraction of the parameters, age, sex, occupation and so on in the case of mortality schedules ascertainable, but a number of essential parameters remain unknown, so that observations on X, as well as calculations on the known right hand terms, are required. Either subjective theory would differ from the a priori approach only in that the values of the parameters and the form of the function would be apprehended intuitively rather than be calculated from theory or inferred from sampling experiments. How does a batter hit a ball is the classic example to distinguish between the rational, algorithmic approach to nature and the intuitive subconscious unanalyzed approach. The batter seldom regards a graduate degree in mechanics as essential for success - though he does indeed bear in mind all the identifiable admonitions which can be derived either from theory or experience.

Personal probability is to be distinguished from objective probability not because it is an entirely different - an unrelated - species, but because it stresses a component always present when a mind contemplates the external world. The mind never knows the world, it has only beliefs about it. These beliefs are a substitute, or better, an alternative for objective measures. In some contexts, the objective criteria prove more successful, in others the subjective judgments. These beliefs can have little value if not verified in practice, and verified in just the way that objective deductions are - by the observed outcome. If the odds makers pick the wrong team to win the World's Series, all agree that the choice was an error. One is entitled to his opinion - but only before the game. The several bases for the concept of probability are set out in Figure 14.

V. A NEW VIEW OF PROBABILITY. It was remarked above that the requirements of supplying a foundation for decision theory in circumstances where evaluating parameters by sampling of many indistinguishable elements is not available, has since World War II been furnished by the concept of probability as "degree of belief." It was further asserted that the true nature of personal probability is not, as often thought, something apart from objective probability. There is only one probability. That is the success ratio in a series of trials. A logical complication arises from the fact that the series is necessarily infinite and therefore necessarily conceptual rather than actual. It may be estimated in each of three ways. The first way is the a priori or necessary approach.

If the possible results can be enumerated, if they possess an adequate degree of symmetry, and if they are independent, we apply the principle of insufficient reason. This approach has always been intuitively appealing, logically intractable, and limited in application. The device of just trying or observing a large number of trials, partially overcomes this latter difficulty, shares in a different way the intuitive appeal, but introduces its own logical difficulties. The degree of belief approach possesses the prime advantage of always being applicable, but only to the perceiving mind, never to the real world. For this latter to occur we must appeal to a method of verification. For example, consider a weather forecaster. Each newscast he presents a "fearless forecast" preceded by an explanation of why his last forecast failed. Weather forecasts are made on the basis of much data, much theory, historical records, or Farmer's Almanacs. The basis is irrelevant. The verification is the thing. If the forecasts are verified, the forecaster is a success. The sequence of trials is thus the successive forecasts periods. The successes are the verified forecasts, all others are failures. The process is just as objective as was Graunt's birth series. That the forecaster used his degree of belief in formulating his forecast is as incidental as the use by a batter of his in deciding to swing at a pitched ball. Each man functions as a measuring instrument. A person is often superior to a machine in such functions.

The catastrophic threat problem does introduce a complication. In most, if not all, realistic applications of personal probability the verification is not long in coming. That is why people forecast. But a catastrophic threat is necessarily a rare event, if it happens at all. Just as we cannot, before the event, accumulate an adequate fund of experience upon which to base an estimate of probability, so we cannot after the verification step accumulate a sufficient fund of verified or disproved forecasts to establish a success ratio in a trial sequence known to be homogeneous in the probability sense. Fortunately, the occurrence compensates for its severity by its rarity. The new element which the present analysis seeks to supply to the contribution of probability to decision-making is a new technique of verification.

For this purpose consider Figure 15. Here we have listed not one stochastic sequence, but three labelled as conditions A, B, C. A, B, and C could be threats of aggression by three different nations, or the threat of war, of famine and of a major epidemic. Each represents what would usually be chosen to characterize one series of trials as ordinarily described in discussions of probability. But the types of catastrophic threat are not limited; they are legion. Suppose we employ the methods of personal probability to assign "degrees of belief" to occurrence of each. Further, let this be done not by one individual, or by one forecasting team, but by many, each applying whatever "betting system" is most attractive. We would get as a result a two-fold matrix of assigned probabilities such as is diagrammed in Figure 16.

While it is true that the probabilities subjectively assigned to the contingency that a catastrophe of specified nature would occur in a specific interval of time is low, and in general varies as between the types of contingencies involved, the possible contingencies are legion, so that it is to be expected that on any given estimation procedure a fairly high number of possible catastrophes would be assigned the same or nearly the same value of the probability of occurrence in a specific time frame. In Figure 17, this is indicated by assigning probabilities not to individual catastrophic contingencies, but to classes characterized by the property that within a class, each individual catastrophe is regarded as equally likely. The grouping may well be different as between rows. Hence, the columns do not refer to the same contingencies in every row. The letter k in each cell represents the number of contingencies all assigned the same probability of occurrence by the technique of that given row.

As each period of observation passes, the occurrence of the various types of catastrophe would be noted. The subjective probabilities assigned by any one procedure (in any one row) would be verified or refuted according as the empirical success ratio was sufficiently close to the ex ante assigned probability or not. This empirical success ratio would be calculated in a slightly different manner from that utilized at present, where it is assumed that every sequence of trials is studied in isolation and without regard to what is happening in any other. Let the success ratio at a certain observation period be n/m; where n is the number of catastrophes, all having been assigned the same probability, which have so far occurred and m is the product of the number of such catastrophes by the number of observation periods. Here k equals the number of catastrophes grouped as having the same subjective probability of occurring, and p is the probability itself. Let h be the number of catastrophes of this class which occur in the next observation period. Of course, h will almost always be zero, rarely be unity and almost never be greater. Then the success ratio at the end of this period becomes

$$\frac{n + h}{m + k}$$

The essence of this procedure is that we are judging, not the success in assigning subjective probabilities to specific types of catastrophes, but the betting system itself (at this level of probability). If the probability assignment for one class of catastrophe is verified, our confidence in the assignment of all classes is strengthened.

So far we have pooled experience in assessing the efficacy of a particular "betting system," i.e., method of assigning subjective probabilities, but only at a specific level (or limited range of levels) of the assigned probability. As first shown by Karl Pearson, we can use chi-square or some alternative approach to get a combined test of the procedure irrespective of level.

Particularly in a medical environment the need to ensure a very high level of safety for products produced for human consumption has received increasing attention. The Polio and Thalidomide incidents are instances. The cigarette and cyclamate episodes provide an illustration of variation in response to evidence. These cases have been studied (so far as the author is aware) by extrapolation of conventional statistical techniques from regions of relatively high probability (so that evidence is attainable) to the region of interest where probability of occurrence is extremely low - one in a million exposures or less. The present paper seeks to provide a procedure for catastrophes where even this extrapolatory technique is unavailing, but there is no reason why, when it is, that the evidence from both approaches should not be pooled.

VI. THE COST FACTOR. Costs, like threats, are some real and some imaginary. A perfect system of probability assessments would still not be adequate for decision-making, if costs are ignored, except in those instances where the probability concerned is negligible.

There is a vague appreciation that in a military context "negligible" probabilities are a trap. Defenders, particularly budget officers who are to fund protection from such "impossible" threats, or competitors for those funds inevitably dismiss such probabilities as negligible, yet often in fact, just that stratagem will be selected by attackers in the very knowledge that it will be unanticipated by the defenders. Instances are the Pearl Harbor attack; the choice of land over sea approach at Singapore; the Black Forest end run of the Maginot Line; the "post season" sea assault at the Battle of Hastings. Two famous examples from the history of mathematics describe the failure of Saccheri, and the success of Hamilton from a confrontation with the unthinkable.

As this paper treats catastrophic threat the costs are by assumption high. But a closer estimate is needed. Perhaps the outstanding characteristic of disaster costs is that the ex post[2] estimates of costs "always" outweigh the ex ante[1] cost estimates, and by a huge factor. This is recognized by every parent who tries to get his teenage offspring to cross a college campus or to drive the family car with due caution. The military meets this problem in stark terms when it tries to indoctrinate habits of safety in new recruits. Live ammunition in field exercises represents one attempt to secure credible reality. However, no adequate technique has been found.

---

[1] before the observation period.
[2] after the observation period.

But while a closer approximation of ex ante to ex post costs assessments by new troops would reduce the incidence of personal tragedy, and shorten the battle seasoning of troops, the wide discrepancy inevitably existing in assessing the costs of catastrophes, is the essence of the catastrophic threat problem. Indeed, in practice all or nearly all the discussion will be found to turn on an estimation of the probabilities, and little of it on assessment of the costs, though a great deal of attention is devoted to deprecating the costs of prevention or of protection predicated on this presumed negligible risk. With a one-sided consideration of both costs and probabilities, there is little chance of efficient decision-making.

The instinctive impulse of the mathematically, or philosophically minded is to generalize the problem and then return to the specific example - in our case the unreality of ex ante estimates - on the basis of supposedly clearer insight. To estimate is to make a decision, and decision-making is intrinsically emotional. This point was graphically demonstrated by Koehler in experiments during World War I, Figure 18.

If a chicken is placed close to a short stretch of fence and a handful of grain is placed on the other side close to the fence, but beyond reach through it, the chicken will excitedly press against the fence and never step back to seek a way around it. If the grain is moved farther away, the emotional attraction will be reduced, and the hen will have a better chance of solving the problem.

So true is it that decision-making is an emotional process that advertising and political discussion are directed almost exclusively at emotional, and only incidentally at logical components of issues. In consequence persons who will make decisions are selected on the basis of emotional, not rational, characteristics. So widespread is this process that it almost becomes a "first law of management" that, when it comes to making decisions "those with the power lack the knowledge, those with the knowledge lack the power." In consequence, the "on tap" become the "untapped."

This lack of knowledge has not merely been recognized, but has been deplored through all of history. Even so, courts of inquiry, inquests, grand juries, staffs, study groups and a host of other techniques, have evolved to supply knowledge deficiencies of those in power. However, in a power rather than information dominated environment, all such devices tend to be ineffective because they tend to be ignored. Where power makes decisions, only power can influence decisions. Where decisions are out of touch with reality because information is ignored, and countervailing power is laking feedback from consequences is avoided. Figure 19 lists certain principles which seem applicable.

Ascendency of emotional over informational factors in decision-making leads to a widely recognized, widely deplored, oscillation in efforts at averting or dealing with catastrophic threat. The soldier

knows this oscillation in the contrast between his hero's role in times of threat. and scapegoat's role in seeming peace. The tin-roof parable is known to all, but heeded by few. The technique outlined in this paper represents only what could be done, but not what will be done, until the forces producing this violent oscillation are rendered impotent. A single incident at this Design Conference will illustrate the situation. One talk at the Conference was given by Dr. Condon, in charge of the safety program for the moon landings. After explaining that the success of the safety program was due, in large part at least, to his refusal to adopt a "good enough" attitude, the speaker conceded that the program was thereby made costly and commented that, in the days ahead there will be "pressures on us in the way of costs."

This is the universal story of the fight against catastrophe. The expense ensures success; the success breeds lack of support. It will always be thus where decision-making is a budgetary process.

REFERENCES

Bright, James R. ed. [1968]. Technological Forecasting for Industry and Government. Prentice-Hall, Englewood Cliffs, New Jersey.

DeFinetti, Bruno [1937]. Foresight: Its Logical Laws, Its Subjective Sources. Annales de l'Institut Henri Poincare, Vol. 7. Reprinted in Studies in Subjective Probability, Kyburg & Smokler, eds. Wiley.

Quade, E. S. ed. [1964]. Analysis for Military Decisions. The Rand Corporation, Santa Monica, California Report No. R-387-PR.

Savage, Leonard J. [1954]. The Foundation of Statistics. John Wiley & Sons, Inc. Available from University Microfilms, Ann Arbor, Michigan.

SUDDEN

EXTREME

UNEXPECTED

PREDICTABLE

PSYCHOLOGICAL

RARE

Figure 1. Characteristics of Catastrophes

SUDDEN

Donora vs. Los Angeles Smog

Iran Earthquake vs. Erosion

Epidemics vs. Normal Death Rate

Air Crash vs. Automobile Deaths

Poliomyelitis vs. Cardiac Arrest

Figure 2. Illustrative Examples

EXTREME

Date	Place	Type	Deaths
1947	Texas City	Fire-exp.	561
1942	Boston	Club Fire	492
1944	Port Chicago	Ammo Ships	322
1944	Hartford	Circus Fire	168
1947	Southwest	Tornado	167

Figure 3. Illustrative examples

UNEXPECTED

Thalidomide Teratogency

Pearl Harbor Attack

Prince of Wales and Repulse

Virginia Flood (1969)

Figure 4. Illustrative examples

PREDICTABLE

Radium illuminated dials

Giant Solar Flare

Unknown Moon Agent

Smallpox among natives

Nuclear test triggered tsunami

Figure 5. Illustrative examples

PSYCHOLOGICAL

Battle of Big Bethel

Sheep deaths near Dugway

D. C. Subway deaths

"Brainwash" political death

Figure 6. Illustrative examples

RARE

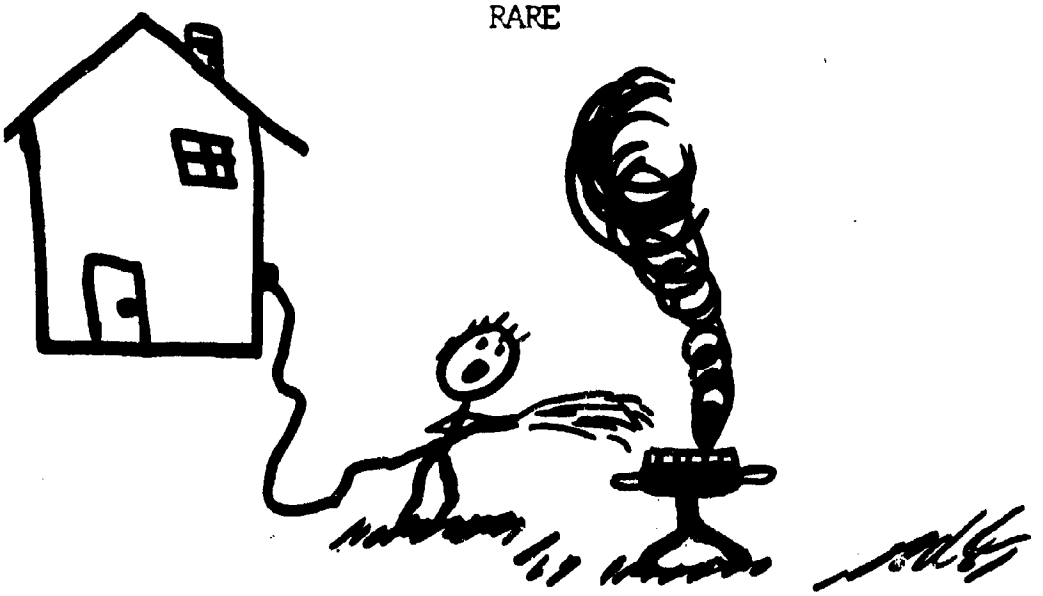


Figure 7. Putting out fires

RANGE - - - - - - - - - - - - - all

TECHNIQUE - - - - - - - - - - - "educated" guess

SAMPLE SIZE - - - - - - - - - - unknown

VALIDATION- - - - - - - - - - - experience

Figure 8. Common sense

Preceding page blank

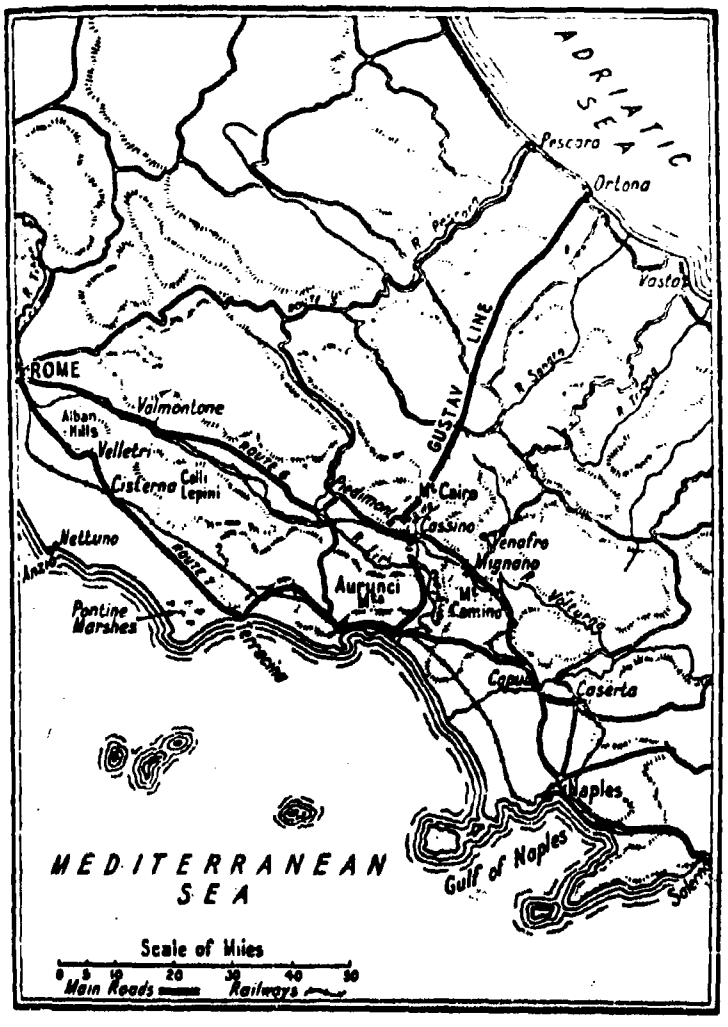


Figure 9. Theater of Anzio campaign

RANGE - - - - - - - middle

TECHNIQUE - - - - - formal statistics

SAMPLE SIZE - - - - large

VALIDATION- - - - - experiment

Figure 10. Formal statistics

RANGE - - - - - - - middle

TECHNIQUE - - - - - quality control

SAMPLE SIZE - - - - large

VALIDATION- - - - - feed back

Figure 11. Process monitoring

RANGE - - - - - - - middle

TECHNIQUE - - - - - forecast

SAMPLE SIZE - - - - none

VALIDATION- - - - - feed back

Figure 12. Technological forecasting

RANGE - - - - - - - extreme

TECHNIQUE - - - - - - - - -

SAMPLE SIZE - - - - none

VALIDATION- - - - - feedback

Figure 13. Catastrophic threat

PROBABILITY FOUNDATIONS

Necessary	**Fermat, Pascal**
Frequency	**Graunt, von Mises**
Abstract	**Kolmogoroff**
Personal	**Bernoulli, Bayes**
Neglected Causes	

Figure 14.

Cond	Rslt	Cond	Rslt	Cond	Rslt
A	F	B	S	C	S
A	S	B	S	C	F
A	F	B	S	C	F
A	F	B	F	C	S
A	S	B	S	C	F
A	F	B	S	C	F

Figure 15. Examples of chance sequences

$$\begin{matrix} P_{11} & \cdots & P_{12} & \cdots & P_{1j} & \cdots & P_{1c} \\ P_{21} & \cdots & P_{22} & \cdots & P_{2j} & \cdots & P_{2c} \\ \cdot & \cdots & \cdot & \cdots & \cdot & \cdots & \cdot \\ P_{r1} & \cdots & P_{r2} & \cdots & P_{rj} & \cdots & P_{rc} \end{matrix}$$

Figure 16. Matrix of probabilities assigned to catastrophes (columns) by different procedures (rows)

K_{11}, P_{11}	K_{12}, P_{12}	....	K_{1c}, P_{1c}
K_{21}, P_{21}	K_{22}, P_{22}	....	K_{2c}, P_{2c}
. .	. .	....	. .
K_{r1}, P_{r1}	K_{r2}, P_{r2}		K_{rc}, P_{rc}

Figure 17. Matrix of catastrophes assigned a common probability

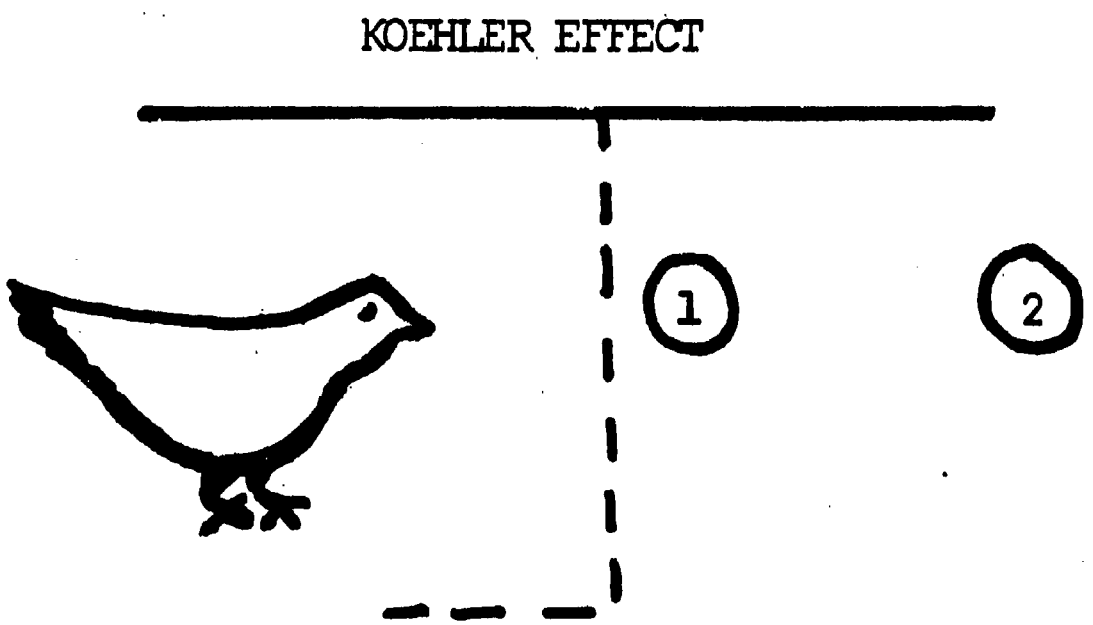


Figure 18. Effect of emotional intensity on decision-making

Feedback necessarily occurs

Information feedback is cheapest

Information feedback is ignorable

Feedback through channels is no feedback

Figure 19. Principles of decision-making verification

APPENDIX

1. A commander can make a decision simply by ruling out what appears to him to be impractical or unfeasible. (244)

2. General Alexander felt that ... allied troops on the enemy flank below Rome might so threaten German communications as to compel the enemy to retreat. (248)

3. Link-up between the main and the Anzio fronts, it was assumed, would take place no later than seven days after the landing. (248)

4. A strengthened Anzio force, if assured continuous resupply by water, could, he believed, consolidate a beachhead.... (249)

5. Whether the 60 miles between Anzio and the Garigliano was too great a distance for action on one front to influence the other was discussed, but it was accepted as an unavoidable risk. (249)

6. It was, impossible to predict the exact German reaction to a landing, but the most probable reactions seemed desirable from the Allied standpoint. (250)

7. The Anzio force might provoke the Germans ... to withdraw. (250)

8. ... intelligence officers of the 15th Army Group were rather optimistic. (251)

9. ... they 'counted on' the effect of weather and on harassment by the Allied air forces to interfere.... (251)

10. The [ambiguity of orders to Sixth Corps] arose from the difficulty of judging .... (252)

11. Fifth Army intelligence estimates were less optimistic.... (252)

12. The enemy was judged to have. (252)

13. By the third day the Germans could perhaps. (252)

14. Two additional divisions could probably. (252)

15. The Fifth Army assumed that the VI Corps would meet strong resistance on the beaches (252)

16. It expected the Corps to receive heavy counterattacks.... (252)

17. ... having underestimated German strength at Salerno.... (252)

18. The Fifth Army - and with it the VI Corps - expected the same pattern.... (252)

19. The Fifth Army expected the VI Corps to be ready to do one of two things upon landing. (253)

20. The operation becomes such a desperate undertaking. (254)

21. Otherwise "a crack on the chin is certain" (254)

22. A failure now would ruin Clark, probably kill me, and certainly prolong the war... (254)

23. A week of fine weather at the proper time and I (Lucas) will make it.... (255)

24. Alexander told Lucas "we have every Confidence in you" (255)

25. What troubled General Lucas...was the contrast between his own concern...and nonchalance in the higher echelons.... (255)

26. Lucas was not so sure. (255)

27. The chances are seventy to thirty that (256)

28. He [Lucas] believed his forces lacked the strength (256)

29. The general idea seems to be... (256)

30. I wish the higher headquarters were not so optimistic.... (256)

31. Securing a beachhead was all Fifth Army expected.... (257)

32. Lucas [was not] to push on to the Alban Hill mass at the risk of sacrificing his corps (257)

33. Such a possibility [moving on the Alban Hills] appeared slim to the Fifth Army Staff (257)

34. The staff questioned Lucas' ability.... (257)

35. It was obvious what the loss of the supply base would mean (257)

36. If the enemy came to Anzio in strength (257)

37. The British feared they might mistake Americans for Germans (257)

38. What everyone had overlooked ... was the possibility of achieving complete surprise (258)

39. The Germans always regarded the long sea flanks in Italy as exposed.... (258)

40. To reinforce [local troops Kesselring] expected to call on Tenth Army for a division.... (259)

41. He hoped to have the Fourteenth Army in north Italy move ... the equivalent of about one or two divisions (259)

42. Fearing that the Fifth Army was about to make a breakthrough.... (259)

43. Feeling that the fate of the Tenth Army .... (259)

44. According to the German estimate the landing had a good chance... (259)

45. Field Marshall Kesselring assumed that the troops would probably try to seize the Alban Hills (260)

46. The Germans were considerably reassured by Allied behavior at the landing (261)

47. Kesselring's order to stand fast on the Garigliano-Rapido line was... in the nature of a gamble.... (261)

48. If the Allies attacked on January 23 or 24, German forces would not be strong enough to hold (261)

49. The evening of 23 January, Kesselring "believed" that the danger of a beachhead expansion was no longer imminent (261)

50. By 24 January the German command considered the danger of an Allied breakthrough removed. (261)

51. Alexander was very optimistic, Clark somewhat subdued (261)

52. Lucas' concern with logistical aspects came not only from prudence (262)

53. He believed the Germans could increase their build-up (262)

54. He believed the Germans would stop his VI Corps before it could cut their line of communication (262)

55. His intelligence officers informed him that the Germans were taking troops from the Fifth Army main front to oppose him (262)

56. This might permit the Fifth Army to advance (262)

57. The Fifth Army, Lucas was certain, would still have to fight powerful rear guards (262)

58. He expected no spectacular rapidity of movement (262)

59. ... he sought to build up his strength and his supplies to remain intact even though isolated (263)

60. I feel now [January 25] the beachhead is safe (263)

61. Lucas expected the 1st Armored Division to arrive soon (263)

62. That is about all I can supply but I think it will be enough (263)

63. I must do nothing foolish (263)

64. I must hold it "..." I think I can (263)

65. Kesselring came to the conclusion that the Allies were preparing a full scale attack (263)

66. The best defense, he felt, was an attack on his own (263)

67. Lucas thought he could attack in a few days (264)

68. He expected 30 LST's to be unloaded at Anzio 27 January (264)

69. Clark "received the impression" that the outcome of the struggle depended on who could increase his forces more quickly (264)

70. Though the situation was not clear to Clark (264)

72. Apparently, some of the higher levels think I have not advanced with maximum speed (264)

73. I think more has been accomplished than anyone had a right to expect. (264)

71. He urged Lucas to take bold offensive action (264)

74. This venture was always a desperate one (264)

75. I could never see much chance for it to succeed (264)

76. Without Anzio our situation would have been desperate (264)

77. Had I rushed troops to Albano and Velletri they would have been destroyed (264)

78. The only thing to do was what I did (264)

79. Keep the enemy off balance until the Corp was ashore and everything was set (264)

Lucas could envisage holding some troops in corps reserve (265)

The situation is crowded with doubt and uncertainty (265)

I expect to be counterattacked in the morning. (265)

I think he realized the seriousness (265)

He [Clark] thinks I should have been more agressive on D-Day (265)

There has been no chance to build "Shingle" up to decisive strength (265)

Anyone could have seen that from the start (265)

I can win if I am left alone (265)

I don't know whether I can stand the strain (265)

Clark and those above him thought Anzio would shake the Cassino loose at once (266)

They had no right to think that (266)

It was clear that the attack had not accomplished much (266)

The Ranger force met unanticipated opposition (266)

The enemy had an unexpectedly strong and well organized defensive [illegible] (266)

It seemed clear the Germans has built up their forces around Anzio (266)

What the Allies did not know was how close they came to breaking out [illegible] broad (266)

The Allied intelligence officers seemed like overwhelming German [illegible]

[illegible] intelligence officers had to assume (266)

[illegible] thought he could support two more divisions at Anzio on [illegible] (267)

[illegible] and Clark decided that the enemy build-up dictated a [illegible] defensive tactics (267)

[illegible] February "keeping the enemy off balance" was a forlorn hope (267)

[illegible] are disappointed but there is no military reason why they [illegible]

102. General Devers thought Lucas should have gone on - on landing (268)

103. "Had I done so, I would have lost my corps" (268)

104. Clark thought Lucas had done all he could at Anzio (268)

105. I thought I was winning something of a victory (268)

106. General Clark thought Lucas could have taken the Alban Hills but could not have held them (268)

107. Clark thought British G-2 intelligence was always over optimistic (269)

108. The Germans built up their differences at Anzio much faster than the British believed possible (269)

109. Clark had always felt that Anzio had little chance of success (269)

110. In retrospect Clark felt that the total losses at Garigliano and at Anzio might have been safer and as productive at Garigliano alone (269)

111. A powerful counter attack at Anzio could well have wrecked the entire Italian Campaign (269)

112. By the end of January, Clark was disappointed by Lucas' lack of aggressiveness (269)

113. Clark believed Lucas should have made a reconnaissance in force to capture Cisterna and Campoleone.... (269)

114. Clark thought such an effort to be not incommensurate with Lucas' forces (269)

115. Others felt much the same (269)

116. General Marshall thought Lucas could have taken the Alban Hills (269)

117. However, he thought Lucas had acted wisely (269)

118. Marshall felt Lucas could not have held the Alban Hills and the port at Anzio (269)

119. The theater G-2 had held the same opinion at the Christmas Day Conference at Tunis (269)

120. G-2 thought Lucas would have been in a bad way without a main front breakthrough. (269)

121. The Allies would be unable to keep the Germans from shifting forces to Anzio from south Italy as well as elsewhere (269)

122. General Lemnitzer also felt the Allies did not have the strength to hold the Alban Hills (269)

123. Lemnitzer thought that Alexander hoped that the Anzio operation plus a main force attack "might" force a German withdrawal (270)

124. The advance "on" the hills was exactly what Alexander thought possible (270)

125. When Alexander visited the beachhead on D-Day he approved the decision not to push out far from Anzio. (270)

126. Lemnitzer thought that Alexander thought that Lucas had done no wrong but was under too much strain (270)

127. By that time it was clear the Anzio operation would involve a long, hard struggle (270)

128. It would seem that Lucas' action during the first few days was justified (270)

129. The main German Army showed no signs of withdrawing (270)

130. The Allies saw no immediate prospect of forcing a general retreat (270)

131. It became far more likely the Germans would move in strength against Anzio (270)

132. If the VI Corps went too far inland it would risk annihilation (270)

133. Allied intelligence judged the German strength as sufficient but not overwhelming (270)

134. It would seem that the Allied hesitation on the Anzio shore stemmed from a belief in German invincibility (270)

135. This belief was a product of doubt and uncertainty both before and during the operation (270)

136. This belief was used later to explain the inevitability of the actual course of events (270)

137. The only thing that disturbed Lucas was the necessity to safeguard the port (271)

138. Without it the swift destruction of the corps was inevitable (271)

139. Lucas thought he could not have done differently (271)

140. Nevertheless the alternative remained a disturbing possibility to him (271)

141. He admitted a mass of armor and motorized infantry might have reached the Alban Hills (271)

142. He was sure he could not have remained there (271)

143. Any force that far from Anzio would have been in the greatest jeopardy (271)

144. Lucas did not see how it would have escaped annihilation (271)

145. As it turned out he believed he had reached positions from which the enemy was unable to dislodge him (271)

146. Lucas believed the whole operation a mistake (271)

147. Anyone who expected him to push to the Alban Hills was bound to be disappointed (271)

148. Lucas had never considered doing so (271)

149. He considered his mission to be taking the port and its surroundings (271)

150. Perhaps this was an influence of the Navy (271)

151. Admiral Cunningham asserted no reliance could be placed on over the beaches maintenance (271)

152. Unfavorable weather was probable (271)

153. General Clark said: "You can forget this goddam Rome Business" (271)

154. The capture of Anzio was an obvious objective (271)

155. But early occupation of the Alban Hills was vital (271)

156. The Anzio forces later realized the importance of the hills (272)

157. Was General Lucas justified in delaying seven days before starting his offensive? (272)

158. Could he have gotten away with the gamble of an immediate drive to the Alban Hills (272)

159. Certainly the complete surprise achieved at the landing could have been exploited (272)

160. According to Tenth Army estimates only a quick cutting of lines of communication would have led to major Allied success (272)

161. Such a success would be more likely to capture Rome (272)

162. According to Kesselring's Chief of Staff, an audacious flying column could have penetrated to the city (272)

163. He was astonished at the Allied passivity (272)

164. Could the Germans have withstood a dynamic front as they did the static front? (272)

165. Would they have dared to hold both at Anzio and at Garigliano (272)

166. An Allied force ensconced on the Alban Hills would have been a much greater threat than those or Anzio (272)

167. The answer can only be speculation (272)

168. Alexander thought an aggressive commander would have acted differently than Lucas (272)

169. He would and could have pushed regimental strength patrols to the hills (272)

170. The shock of Allied troops directly threatening Rome might have by itself permitted Allied retention of both the hills and a supply corridor (272)

171. A bluff might have worked (272)

172. General Patton might have been successful (272)

FROM: COMMAND DECISIONS - Office of the Chief of Military History
Department of the Army 1959

EMPIRICAL BAYES AND THE DESIGN AND ANALYSIS OF EXPERIMENTS

Richard G. Krutchkoff
Virginia Polytechnic Institute
Blacksburg, Virginia

Let me start with an introduction to Empirical Bayes. Consider the simple estimation situation in which we observe a value x of the random variable X which has distribution function $F(x|\theta)$, and must estimate θ with small squared error. In the parametric situation, the form of the distribution function is known except for the value of the parameter θ. Both x and θ may be vector valued.

For Empirical Bayes to be applicable here we consider the case in which the estimation problem is routine. That is, we observe x_1 from $F(x_1|\theta_1)$ and must estimate θ_1; then some time later, in a similar but independent situation, we observe x_2 from $F(x_2|\theta_2)$ and must estimate θ_2. This routine situation continues until at present we have the observation x_n from $F(x_n|\theta_n)$ and we must estimate θ_n. These estimating situations we call experiences. As an example consider the situation encountered at the Radford Arsenal. Every six weeks base grain was mixed and subsequently cured with Nitroglycerine in order to form propellent for the Nike missile. It was desired to estimate the parameters for each base grain separately since it was believed that the parameters would vary from base grain to base grain in some unpredictable manner. Since the θ values θ_1, θ_2, θ_3,..., θ_n, vary in an arbitrary and unpredictable manner, we assume that θ is a random variable but with a completely unknown distribution. It is important to note that we do not use our ignorance of the distribution as a justification for choosing a diffuse or uniform distribution.

If one were to take a completely classical approach to the problem he would note that X_n is a sufficient statistic for θ_n. This is easily seen by noting that

$$F(x_1,x_2,x_3,\ldots,x_n|\theta_1,\theta_2,\theta_3,\ldots,\theta_n) = \prod_{i=1}^{n} F(x_i|\theta_i) \quad .$$

The classical solution to the problem therefore must completely ignore the observations $x_1,x_2,x_3,\ldots,x_{n-1}$ and use only the observation x_n in estimating θ_n. Even intuitively this is an unfortunate result.

The pure Bayesian approach to the problem assumes a form for the distribution of θ, say $G(\theta)$, and then obtains the estimator

$$E(\theta|x_n) = \frac{\int \theta dF(x_n|\theta)dG(\theta)}{\int dF(x_n|\theta)dG(\theta)}$$

(the posterior mean) as the minimizing estimator. If the choice of $G(\theta)$ is correct then this is indeed the minimizing estimator. If the choice of $G(\theta)$ is not correct then the estimator may have a very large mean squared error. Note that the Bayes estimator ignores the past experience $x_1, x_2, x_3, \ldots, x_{n-1}$, as did the classical estimator. Surely, we should be able to use this experience in some way.

The Empirical Bayes approach to this problem is now very simply stated. We find the Bayes estimator $E(\theta|x_n)$, which is usually given in terms of the unknown distribution function, and express it in a form which can be estimated from the data, $x_1, x_2, x_3, \ldots, x_n$ without knowledge of, or assumptions about, the unknown prior distribution. The proper forms for $E(\theta|x_n)$ are given in Rutherford and Krutchkoff [11] for four general families of distributions. Examples of members of these general families are the Poisson, Negative Binomial, Logarithmic, Gamma, Normal (unknown mean), Normal (unknown variance), Exponential, and the Uniform Distributions. In Lemon and Krutchkoff [5] an Empirical Bayes estimating procedure is proposed for any discrete conditional distribution. This procedure has now been extended to include any conditional distribution.

Let me now briefly mention some recent applications of this approach. First, consider the simple linear orthogonal model

$$Y_i = \alpha + \beta(X_i - \bar{X}) + \varepsilon_i$$

where the errors are assumed to be normal and where we must routinely estimate α and β. This problem is considered in Clemmer and Krutchkoff [1] and the example analyzed there is worth rementioning here.

Every six weeks Radford Army Arsenal mixed Base Grain for their Nike missiles. The Base Grain was then cured with Nitroglycerine to form rocket propellent. Estimates of the parameters in a linear model were required for each Base Grain. Since the chemicals were purchased at different times and mixed at different times under different atmospheric conditions, the parameters were expected to vary in an unpredictable manner. Using the estimator for the normal distribution given in Rutherford and Krutchkoff [11], Clemmer and Krutchkoff [1] found the desired estimators as

$$E(\alpha|\hat{\alpha}) = \hat{\alpha} + \frac{\sigma^2}{N} \frac{f'(\hat{\alpha})}{f(\hat{\alpha})}$$

and

$$E(\beta|\hat{\beta}) = \hat{\beta} + \frac{\sigma^2}{S_{XX}} \frac{f'(\hat{\beta})}{f(\hat{\beta})}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the usual Least Squares or Maximum Likelihood estimators for α and β, σ^2 is the error variance, N is the number of observations taken in the n^{th} experience, S_{XX} is the usual sum of squares of the independent variable $\left(\sum_{i=1}^{n} (X_i - \bar{X})^2 \right)$; $f(\hat{\alpha})$ is the marginal density of the least squares estimator; and, $f'(\hat{\alpha})$ is the derivative of this density estimated at the same point. In general, the form of the estimator is simply the least squares estimator plus a correction factor. The correction factor is the variance of the least squares estimator times the ratio of the derivative of the marginal density to the marginal density itself evaluated at the present value of the least squares estimator. It is worth noting here that if the parameter has a diffuse prior distribution then the ratio of the derivative of the density to the density will in effect be zero and the Empirical Bayes estimator will be the Classical estimator. Thus, when the prior information is of little value this correction term disappears rather than biasing the result undully.

The estimate of the ratio recommended in the paper [1], can be simply written as

$$\frac{f'(\hat{\alpha})}{f(\hat{\alpha})} = \frac{\sum_{i=1}^{n} \left\{ \left[\frac{S_{in} A_i^*}{A_i^*} \right]^2 - \left[\frac{S_{in} A_i}{A_i} \right]^2 \right\}}{h \sum_{i=1}^{n} \left[\frac{S_{in} A_i}{A_i} \right]^2}$$

where

$$A_i = \frac{\hat{\alpha}_n - \hat{\alpha}_i}{2h}$$

$$A_i^* = \frac{\hat{\alpha}_n - \hat{\alpha}_i + h}{2h}$$

and where

$$h = n^{-1/5} \left\{ \max \left[\frac{1}{n} \sum_{i=1}^{n} (\hat{\alpha}_i - \bar{\hat{\alpha}})^2, \frac{1}{N} \sum_{i=1}^{N} (y_i - \bar{y})^2 \right] \right\}$$

with $\bar{\hat{\alpha}} = \frac{1}{n} \sum_{i=1}^{n} \hat{\alpha}_i$, and $\bar{y} = \frac{1}{N} \sum_{i=1}^{N} y_i$.

Note that we are using $\hat{\alpha}_i$ to represent the least squares estimate for α in the i^{th} experience and $\frac{\sin 0}{0}$ is to be interpreted as unity. An estimate of $\frac{f'(\hat{\beta})}{f(\hat{\beta})}$ is obtained by simply replacing the α's with β's.

The improvement of this Empirical Bayes estimating procedure is then compared with the Classical procedure by taking a ratio of the mean squared errors. This was done by choosing a distribution for α and β and generating α_1, β_1 from this distribution and then generating several observed values y_i from the regression equation $y_i = \alpha_1 + \beta_1 (x_i - \bar{x}) + \varepsilon_i$ with ε_i being random normal errors. This was done fifty times obtaining new values for the observations at each experience. The Empirical Bayes estimator was obtained using the i-1 previous sets of data as past experience. The entire run of 50 experiences was then repeated 500 times. The average ratio of the Empirical Bayes squared errors to the Classical variance was then plotted as a function of the number of experiences. This was then repeated for many different prior distributions, error variances and experimental designs. It was found that the mean squared error for the Empirical Bayes procedure was never greater than that of the classical procedure with the ratio of the Empirical Bayes mean squared error to the Classical mean squared error often dropping well below unity. It was also determined that the ratio of the mean squared errors depends not on the prior distribution or the error variance or the design but solely on one relation involving them; namely,

$$Z = \frac{\text{Var}\ (\hat{\alpha} | \alpha)}{\text{Var}\ \alpha}$$

for α and a similar expression for β. This index is the ratio of the least squares variance to the variation in the parameter. Intuitively, if the least squares variance is small and the parameter variation large not much information can be extracted from past experience. This is, in fact, the case. When Z is extremely small, below 0.1, we find that the derivative of the marginal on α is small compared with the density itself and the correction factor disappears. On the other hand, when the least squares variance is large and the parameter variation small, much is to be gained from past experience. However, when Z is very large (say 10) then one might as well assume that the parameter is not varying at all and pool all the data. The interesting and realistic range is when Z is about unity. Figures 1, 2, and 3 given here are for Z values of 0.5, 1 and 2 with past experience ranging from one to fifty. The solid line is for σ^2 known, obtained by pooled data or estimated from the present data with $N \geq 20$. The broken line is for σ^2 estimated from the present set of data with $N = 8$. The reduction in mean squares error obtained from nineteen batches of Base Grain is given here in Table 1.

Then in Martz and Krutchkoff [6] the regression model was extended to the multilinear model

$$y_i = \alpha + \beta X_i + \gamma X_i^2 + \ldots + \varepsilon_i$$

where orthogonality was not required. This required obtaining the multivariate extension to the estimators presented in Clemmer and Krutchkoff [1] and finding estimators for joint marginal densities and their vector derivatives. The mean squared errors once again were never greater than those of the least squares estimators with their ratio often dropping well below unity for the usual Z values.

The model was then extended to allow for the possibility that σ^2 varies in an unpredictable way from experience to experience. Consider the model

$$\underset{\sim}{y} = \underset{\sim\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

where $\underset{\sim\sim}{X}$ is a $k \times p$ matrix of known fixed quantities which remains the same from experiment to experiment and $\underset{\sim}{\varepsilon}$ is distributed $N(\underset{\sim}{0}, \sigma^2\underset{\sim\sim}{I})$. We assume $\underset{\sim}{\beta}$ and σ^2 vary randomly from experiment to experiment according to the unknown prior distribution $G(\underset{\sim}{\beta}, \sigma^2)$.

If $\underset{\sim\sim}{X}$ is of rank p, the usual least squares estimators for $\underset{\sim}{\beta}$ and σ^2 are

$$\underset{\sim}{\hat{\beta}} = (\underset{\sim\sim}{X}'\underset{\sim\sim}{X})^{-1}\underset{\sim\sim}{X}'\underset{\sim}{y}$$

and

$$\hat{\sigma}^2 = \frac{(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta})' (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{\beta})}{k-p}$$

Denote $(k-p)\hat{\sigma}^2$ by S. For this situation, the Empirical Bayes estimators are given by

$$\tilde{\underset{\sim}{\beta}} = \hat{\underset{\sim}{\beta}} + \frac{S(X'X)^{-1}}{k-p-2} \frac{f_{N,\hat{\underset{\sim}{\beta}},\, k-p-2}(\hat{\underset{\sim}{\beta}},S)}{f_{N,k-p}(\hat{\underset{\sim}{\beta}},S)}$$

and

$$\tilde{\sigma}^2 = \frac{S}{k-p-2} \frac{f_{N,k-p-2}(\hat{\underset{\sim}{\beta}},S)}{f_{N,k-p}(\hat{\underset{\sim}{\beta}},S)}$$

The past experience for this is in the form of the vectors

$$\begin{pmatrix} \hat{\underset{\sim}{\beta}}_1 \\ S_1 \end{pmatrix}, \begin{pmatrix} \hat{\underset{\sim}{\beta}}_2 \\ S_2 \end{pmatrix} \ldots \begin{pmatrix} \hat{\underset{\sim}{\beta}}_n \\ S_n \end{pmatrix} .$$

The ratio of densities given here are estimated in a way similar to the expressions already given but a bit more complicated. The actual formulas are not yet published, but can be found in the Virginia Polytechnic Institute Ph.D. dissertation of one of my students (see Rencher [8]). Needless to say, many simulations were run and never was the Classical squared error for any component of $\underset{\sim}{\beta}$ smaller than that of the Empirical Bayes procedure for as few as one past experience. The amount of improvement was similar to the figures already shown. On the other hand, there were cases in which we needed as many as five experiences before the Empirical Bayes procedure had a smaller squared error than the Maximum Likelihood procedure when estimating σ^2. See for example figure 4.

Another example of an Empirical Bayes application is in Sequential Estimation. Here we considered the case in which one must sequentially estimate the mean of a Normal distribution, the cost being the sum of the mean squared error and a constant times the number of observations

taken. Although the big problem was the stopping rule, we had some difficulty in handling the past experience, since the number of observations taken differed from experience to experience. This problem was solved, however, and the solution is generally applicable to this type of past experience. The results, i.e., the ratio of Empirical Bayes cost to Classical cost plotted as a function of the number of experiences is typically as shown in the solid line of figure 5. The dotted line is the improvement obtained by using the Classical stopping rule and then the Empirical Bayes estimator. Since determining the stopping time by the Empirical Bayes approach is so very tedious we recommend using this hybrid approach. Unfortunately, the details of these procedures have not as yet been submitted for publication. They are available, however, in the Ph.D. dissertation of another one of my students (see Lemon [4]).

Another project presently underway is the estimation of the power spectral density function in a time series. In a time series situation one often has past experience from similar situations or one can break the present time series into parts which can be considered experiences. For example, in testing stress on airplane wings in a wind tunnel one has the results of tests on other wings. When the Navy obtains a time series signal from the path of a submarine, it is merely one experience in many such experiences. In each of these the object is to obtain the power amplitude of the various frequency components. We have used the Empirical Bayes approach to obtain efficiencies of the order of 150% that of the standard approach. This work is still in progress.

A project which is just about complete now involves estimating the arrival and service parameters in a Que. We have estimators for Ques involving the exponential and the Ehrlang distributions. As usual, the Empirical Bayes estimators have a significantly smaller mean squared error than the usual estimators. A typical example is depicted in Figure 6. λ is the mean arrival time, μ the mean service time, and ρ is the traffic intensity for an M/M/1 Que.

Now let us discuss the Analysis of Variance. First, we considered the random effects model:

$$y_{ij} = \mu + a_i + e_{ij}$$

with I effects and J repetitions per effect. We were able to estimate the variance σ_A^2 of the effects as well as the error variance by using past experience. By making a ratio of these two statistics, we came up with an analog to the F statistic. The percentage points of this statistic were found by Monte Carlo simulation for several values of I, J and numbers of experiences. We found that these tables depend only on J and N and not on I. Many typical situations were then

simulated and the power for the Empirical Bayes test of the Hypothesis $\sigma_A^2 = 0$ was always significantly greater than that of the usual F-test for the same size. After as few as ten or fifteen past experiences the power was as much as 50-80% higher for the Empirical Bayes test.

For the fixed effects model

$$\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$

we had to reparameterize to full rank before proceeding. Once this was done, the estimates were the same as for the linear regression situation. In order to test the Hypothesis $\underset{\sim}{\beta} = 0$, we made an analogy to the F statistic by using the sum of squares of the Empirical Bayes estimators for the parameters in the numerator, and the estimated error variance in the denominator. Here the percentage points were found to depend on the number of repetitions, the number of experiences and also the number of effects. Only tables for up to six effects were simulated. Once again the Empirical Bayes test was always more powerful than the usual F-test (when there were at least four past experiences). Unfortunately, the details of this topic are not yet in print, but can be found in Rencher [8].

Before leaving parametric Empirical Bayes, I've been asked to briefly mention the results we obtained in long range prediction of rainfall. The Weather Bureau puts out a map, twice a month, predicting rainfall in the categories of light, moderate and heavy for a period of 30 days. The predictions are for large areas and not for particular locations. We were asked to use this map and predict for each city the amount of rainfall within the next 30 days. We were able to do just this. We found a procedure for predicting the probability distribution of rainfall in inches for any location that had been collecting such data for at least fifteen years. The results were remarkably successful. One could only compare with the Weather Bureau, however, for the categories light, moderate, and heavy. The Weather Bureau, for example, was correct in Roanoke, Virginia, but 30% of the time while we were correct more like 70% of the time. The details of this project can be found in Philpot and Krutchkoff [7].

Let us now turn to another type of Empirical Bayes Estimation. Consider the situation where the distribution of the observation is itself unknown; a non-parametric situation. Here, we observe the value x of the random variable X whose distribution depends in some unknown way on θ and we are asked to estimate θ with small squared error. Such estimation is of course impossible. For this situation, we assume that there is a supplementary observation obtained after the estimate is given, perhaps in the form of customer feedback. To be more specific, let us say we have observed an x from some unknown distribution and must estimate the θ value related to it in some unknown way. Later,

we are given an observation y from the random variable Y whose distribution may likewise be unknown but for which $EY = \theta$. That is, our supplementary sample is an unbiased estimate of θ. Unfortunately, this estimate is too late. In Krutchkoff [3] the problem is assumed to be routine with the θ values varying in an unpredictable way. The Empirical Bayes estimating procedure for this situation is very simple. If the present observed values is x and there are several past experiences with this same value of x, then use as your estimate the average of the supplementary values y which occurred after the occurrence of the value x. If there is not a sufficient number of past experience at $x_1 = x$, then make a linear regression using the past y values as the dependent variable and the past x values as the independent variable and find the regression value of y at x. The results of such a situation are given in Krutchkoff [3]. Generally, after a few past experiences the Empirical Bayes mean squared error drops below the mean squared error of the Classical estimator which would be used if the distribution of X were actually known. Here we have not only an estimator which we can use when nothing else exists, but one which is better than the usual estimator when one does exist.

An extension of this non-parametric approach was given in Gabbert and Krutchkoff [2]. Here we assumed that a machine producing items was to be checked to determine when it was Out of Control. The linear regression form of the estimate was employed but using only the past fifteen experiences.

A sample of defectives was taken and x, the sample proportion of defectives found. The value y was later supplied by some other procedure such that $EY = p$, the true proportion of defectives. Clearly, each box is an experience with the true proportion of defectives varying randomly. The estimate of the present proportion was obtained from x by using the past fifteen values of x,y, in a linear regression and obtaining the regression value of y for the present value of x. The variance needed in the control chart was also obtained from the linear regression as the variance of the regression line at the present value of x. A typical power function for the Empirical Bayes procedure is given in figure 7. Here the machine was caused to go out of control, producing a proportion of defectives varying randomly about some undesirable proportion P_1 (P_0 is the in control value). The power for the procedure is seen to start out below the usual control chart procedure but after a few experiences with the out of control machine the power increases rapidly. In this example, the Empirical Bayes procedure detects the out of control situation at about the twelfth consecutive sample whereas it takes the usual procedure about 18 samples. Of course, since this Empirical Bayes procedure does not make use of the fact that the sample is Binomial, it can be applied in situations where the distribution of the first sample is itself unknown. We are presently working on a single sample non-parametric approach but not enough results are as yet available to present anything here.

Although the title of this paper contains the "design of experiments," I have virtually nothing as yet to report. Several results, however, can easily be predicted. For the non-parametric supplementary sample situation we did not require the present supplementary sample to determine the estimator or its squared error. If the squared error is within tolerable limits, we need not take the supplementary sample at all. We can, in effect, calibrate the preliminary sample eliminating the need for talking the more costly supplementary samples.

In the parametric situation we recall that our squared error was smaller than that of the Classical procedure. By estimating the prior variance we can estimate the efficiency of this procedure and thus be able to predict the number of observations one needs to obtain a predetermined squared error. This number will, of course, be smaller than that required by the Classical Procedure.

No doubt there are optimal designs for the Empirical Bayes procedures. Since Empirical Bayes is more efficient than the Classical approach using the classically optimal design it makes good sense to hope for an even better efficiency when we find the Empirical Bayes optimal design. This question is, as yet, unanswered, but we are working on it.

REFERENCES

1. Clemmer, B. A. and Krutchkoff, R. G. (1968). The use of Empirical Bayes estimators in a linear regression model, BIOMETRIKA, Vol. 55, pp. 525-534.

2. Gabbert, J. T. and Krutchkoff, R. G. (1970). Supplementary Sample Non-parametric Empirical Bayes in a Quality Control Situation, BIOMETRIKA, to appear.

3. Krutchkoff, R. G. (1967). A supplementary Sample non-parametric Empirical Bayes approach to some statistical decision problems, BIOMETRIKA, Vol. 54, pp. 451-458.

4. Lemon, Glen H. "Empriical Bayes Analysis of Some Sequential Experiments." Virginia Polytechnic Institute Ph.D. Dissertation, December, 1968.

5. Lemon, G. H. and Krutchkoff, R. G. (1969). "An Empirical Bayes Smoothing Technique." BIOMETRIKA, Vol. 56, pp. 361-365.

6. Martz, H. F. and Krutchkoff, R. G. (1969). Empirical Bayes estimators in a multiple Linear regression model. BIOMETRIKA, Vol. 56, pp. 367-374.

7. Philpot, J. W. and Krutchkoff, R. G. (1969). Probability Forecasts of 30-Day Precipitation, Published by the Water Resources Research Center of Virginia, Bulletin #2, 145 pages.

8. Rencher, Alvin C. "The Empirical Bayes Approach to the Analysis of Variance and Linear Regression." Virginia Polytechnic Institute Ph.D. dissertation, August, 1968.

9. Rutherford, J. R. and Krutchkoff, R. G. (1967). The Empirical Bayes approach: Estimating the prior distribution, BIOMETRIKA, Vol. 54, pp. 326-328.

10. Rutherford, J. R. and Krutchkoff, R. G. (1969). & asymptotic optimality of Empirical Bayes estimators, BIOMETRIKA, Vol. 56, pp. 220-222.

11. Rutherford, J. R. and Krutchkoff, R. G. (1969). Some Empirical Bayes techniques in point Estimation, BIOMETRIKA, Vol. 56, pp. 133-137.

TABLE 1

NIKE MISSILE RESULTS

MAXIMUM PRESSURE VS. AGE:

	FOR α:		FOR β:	
n	Z	R	Z	R
2	.43	.98	1.34	.93
3	.51	.95	1.72	.88
4	.45	.95	1.08	.87
5	.23	.97	.53	.90
6	.45	.91	.89	.85
7	.42	.91	.98	.83
8	.74	.86	1.38	.75
9	.64	.88	1.30	.75
10	.59	.86	1.23	.72
11	1.04	.77	5.01	.58
12	.95	.77	3.69	.56
13	.30	.91	.71	.81
14	.69	.81	1.72	.80
15	.54	.86	1.36	.70
16	.54	.86	1.57	.69
17	.41	.89	2.51	.61
18	.52	.86	1.57	.69
19	.17	.96	.45	.87

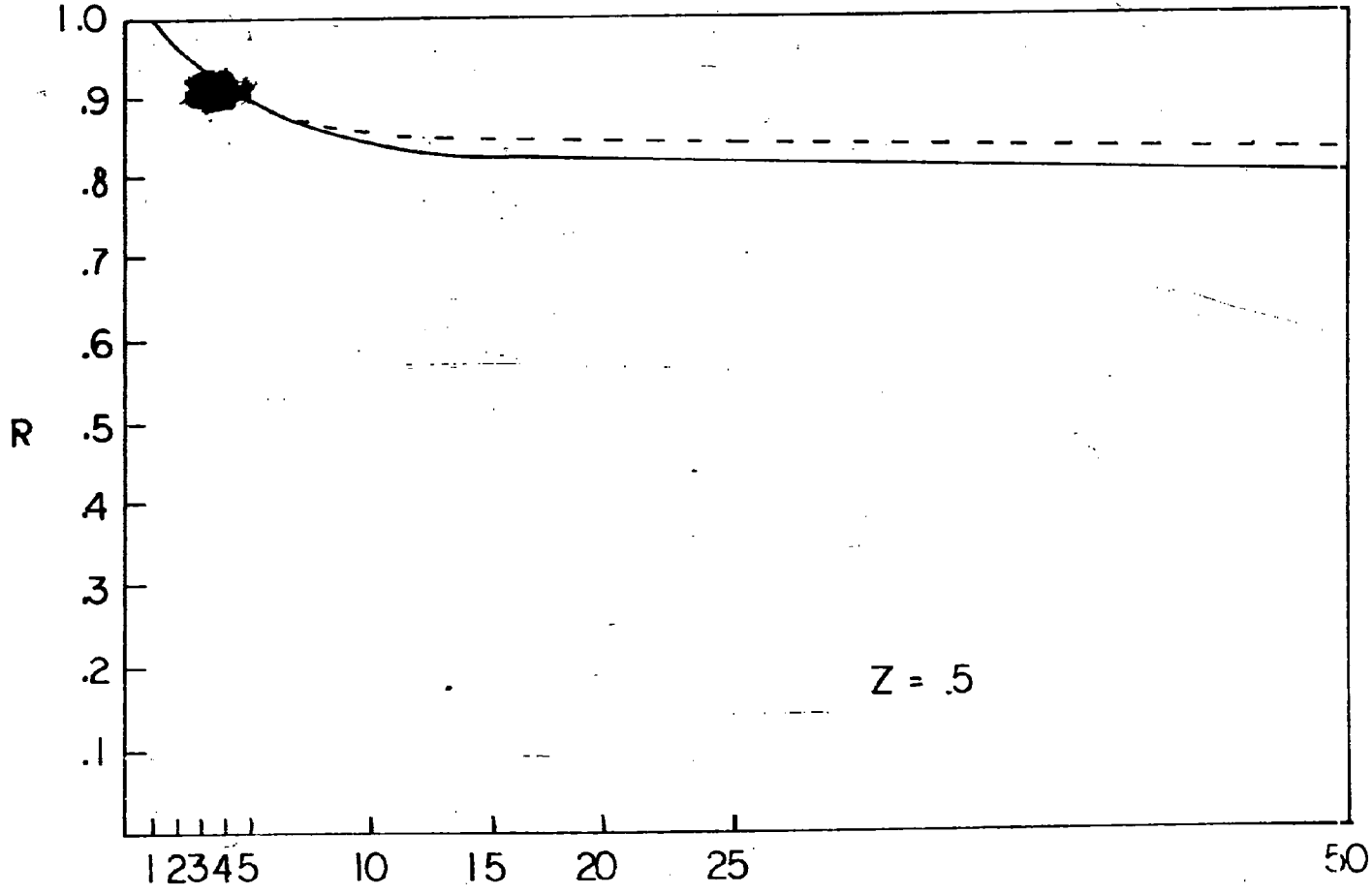


FIGURE 1

Ratio of mean squared error of the Empirical Bayes estimator to the MSE of the least squares estimator for Z = .5

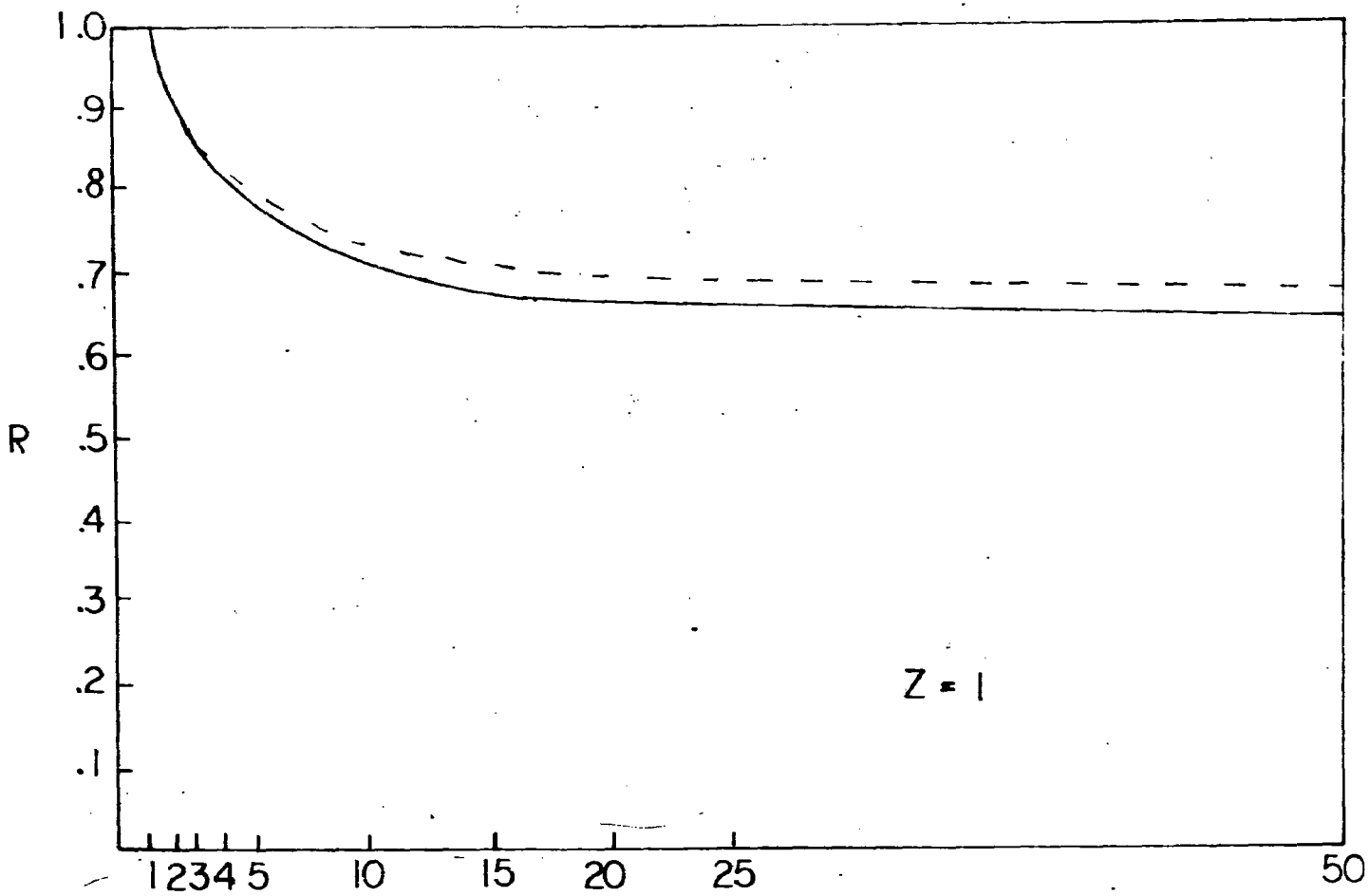


Figure 2

Ratio of the mean squared error of the Empirical Bayes estimator to the mean squared error of the least squares estimator for Z = 1

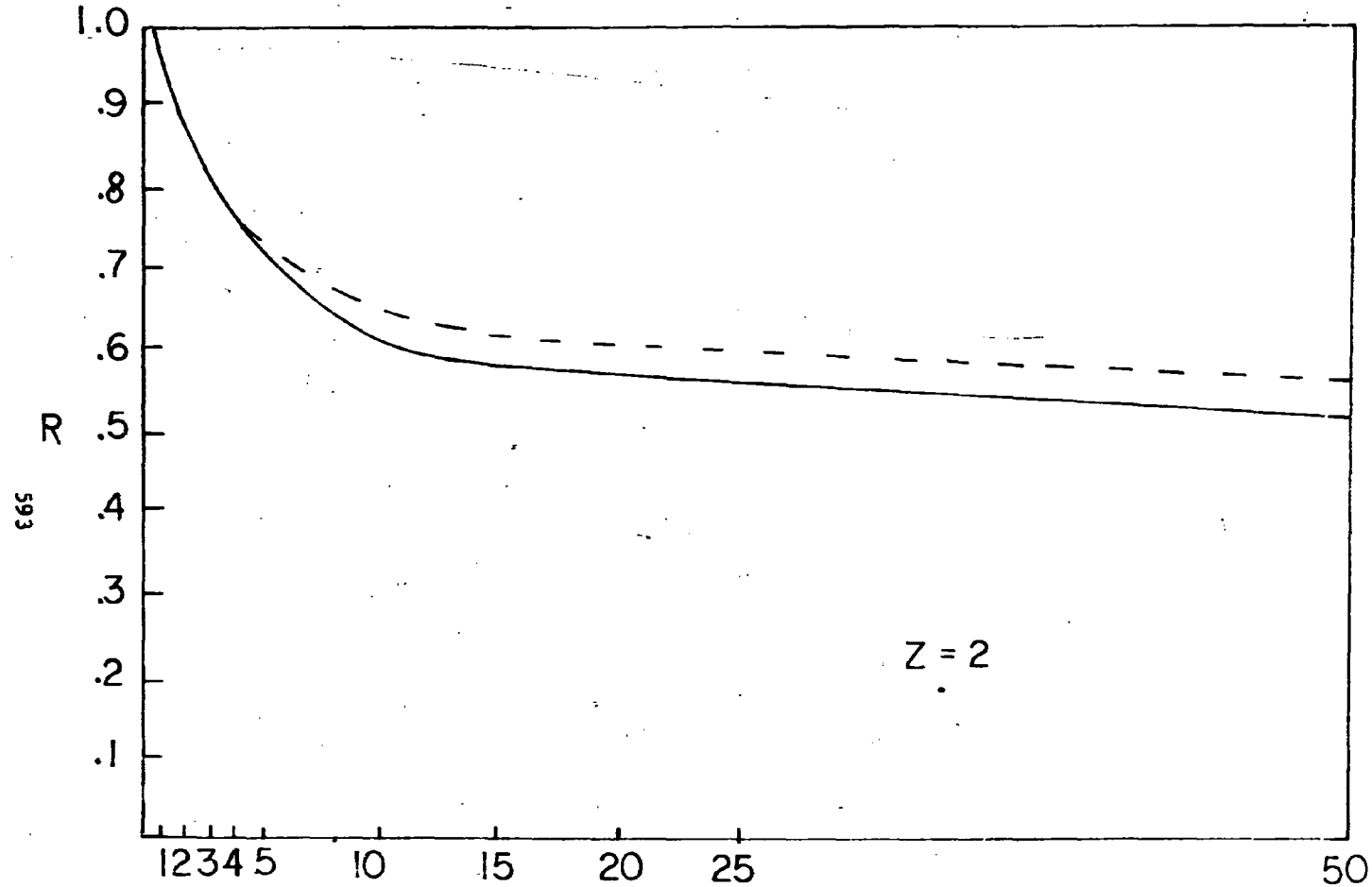


Figure 3

Ratio of the mean squared error of the Empirical Bayes estimator to the mean squared error of the least squared estimator for Z = 2

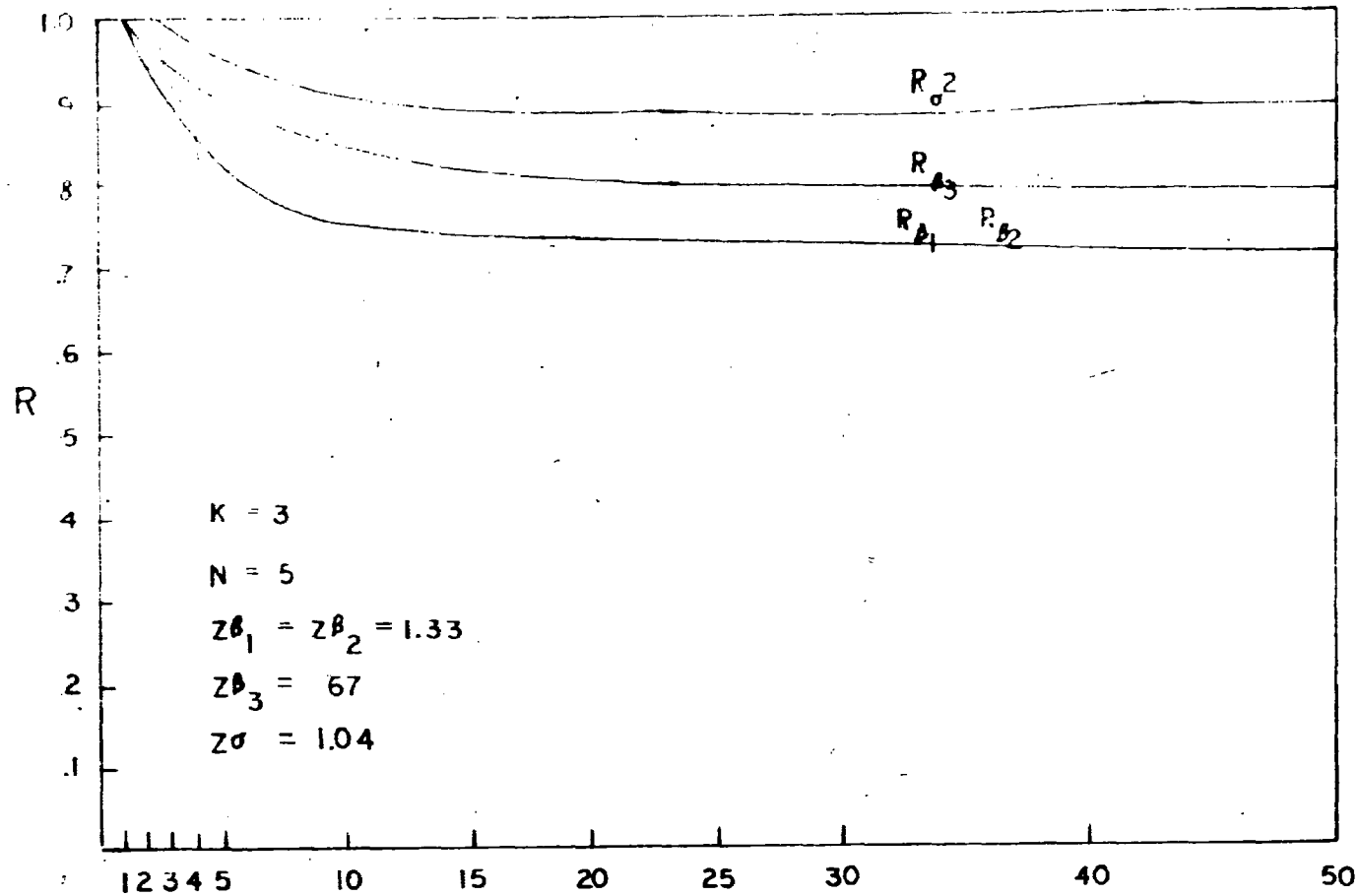


Figure 4

Ratio of mean squared error of the Empirical Bayes Estimators to the MSE of the maximum likelihood estimators for $_1$, $_2$, $_3$ and ^{2}

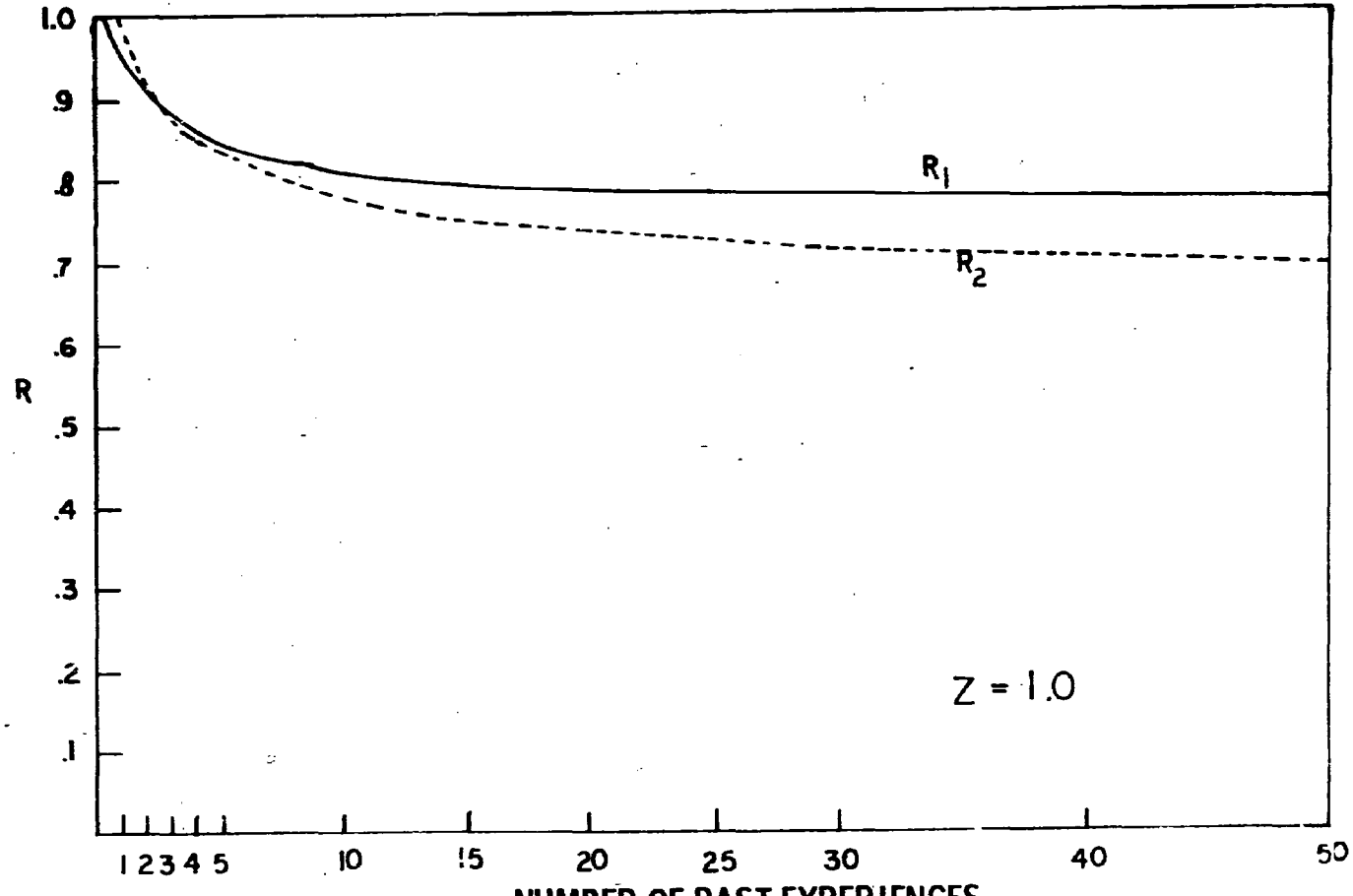


Figure 5

Ratio of cost of Empirical Bayes procedure to the cost of the Classical Procedure in sequential estimation

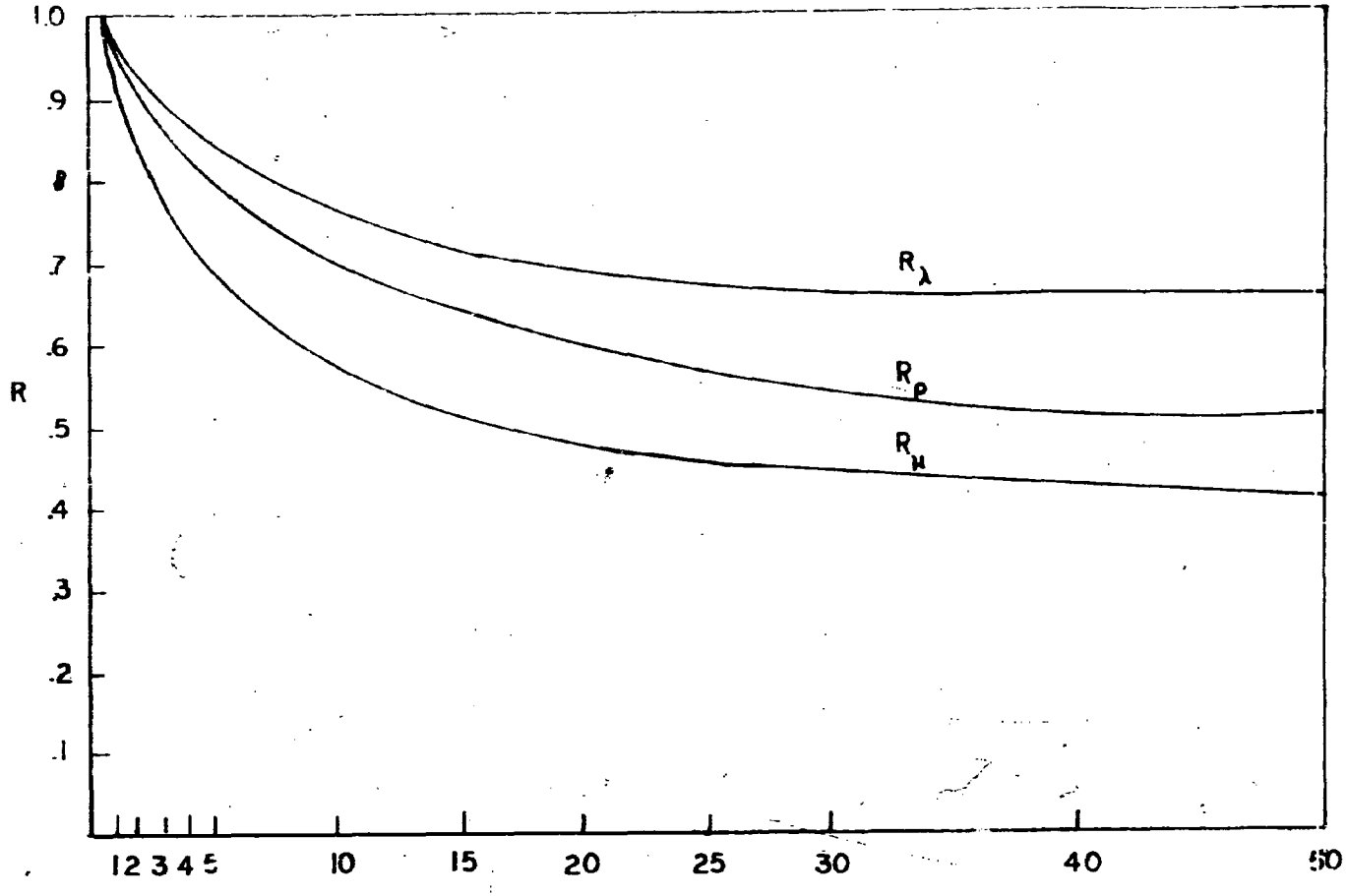
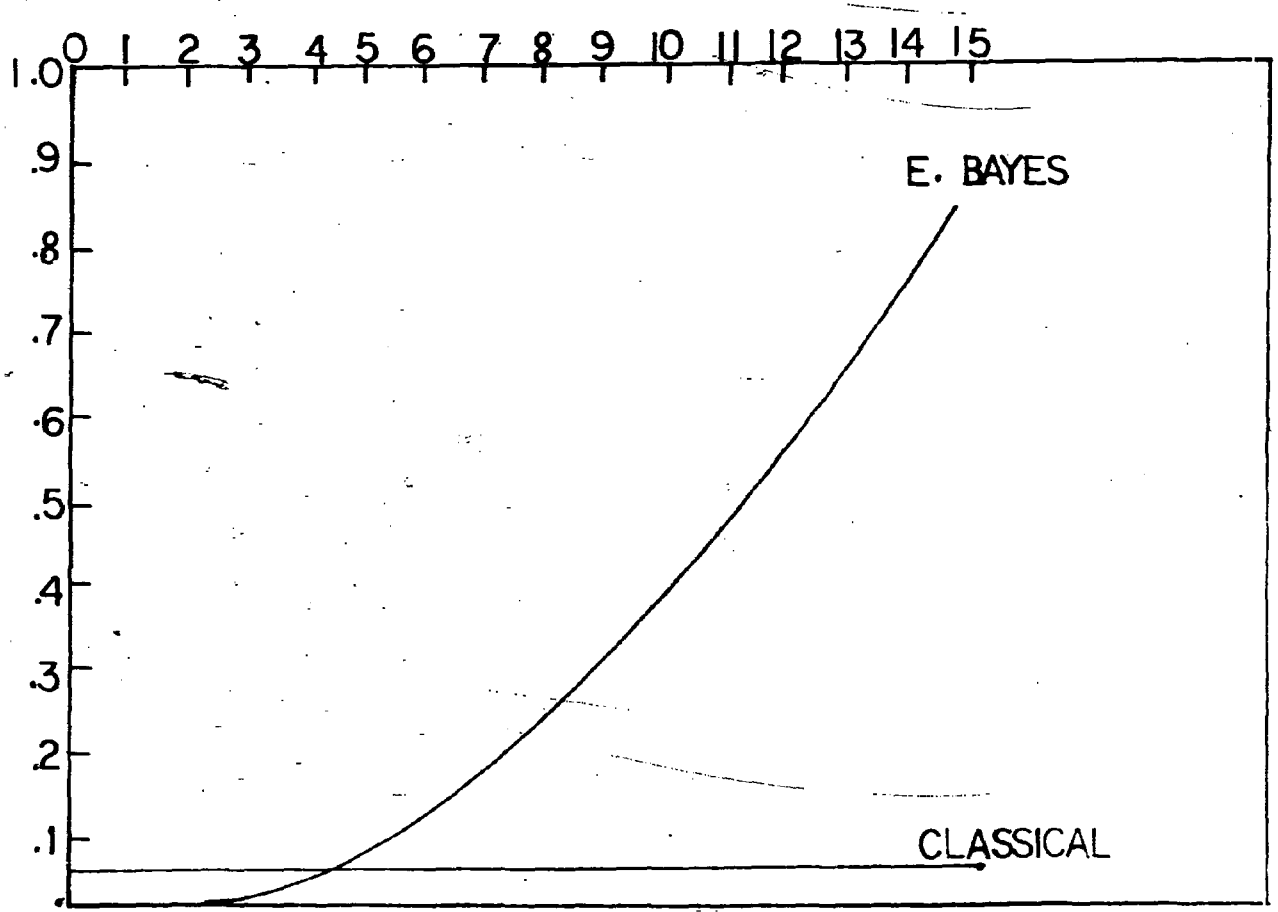


Figure 6

Ratio of mean squared error of the Empirical Bayes Estimation to the MSE of the usual estimation in for an M/M/1 Que



NUMBER PAST EXPERIENCES WITH MEAN FRACTION DEFECTIVES P_1

Figure 7

Power curves for 3 control limits when $P_0 = .10$ and $P_1 = .15$

FIFTH SAMUEL S. WILKS AWARD

Presentation made by

Dr. Frank E. Grubbs

Preceding page blank



650

YOUDEN AWARDED THE 1969 SAMUEL S. WILKS MEMORIAL MEDAL

Dr. W. J. Youden, now retired from the National Bureau of Standards, has been awarded the Samuel S. Wilks Memorial Medal for 1969. The announcement of Dr. Youden's selection for the 1969 Wilks Award was one of the highlights of the Fifteenth Annual Conference on the Design of Experiments in Army Research, Development and Testing, which was held at the U.S. Army Missile Command, Huntsville, Alabama, 22-24 October 1969. Dr. Youden has long been recognized as one of the outstanding applied statisticians by both the U. S. A. and countries abroad, as well, having made many fundamental contributions to the design and analysis of statistical experiments and methodology. The citation for Dr. Youden reads as follows:

> To Dr. W. J. Youden, father of 'Youden Squares' and the 'Youden Diagram,' for his extensive contributions to the art and practice of experimentation in the sciences and engineering, through conception and lucid exposition of novel, yet rather elementary, techniques of statistical analysis and crafty application of standard methods; and through his exceptional productivity as an author, indefatigable energy and phenomenal effectiveness as a speaker, by which he has inspired a whole generation of scientists and engineers to greater achievements through application of his unique statistical precepts.

Previous recipients of the Samuel S. Wilks Memorial Medal include: John W. Tukey, of Princeton University (1965); Major General Leslie E. Simon (1966); William G. Cochran of Harvard University (1967); and, Jerzy Neyman of the University of California (1968).

The Samuel S. Wilks Memorial Medal Award is administered by the American Statistical Association, a non-profit, educational and scientific society founded in 1839. The Wilks Award is given each year to a statistician and is based primarily on his contributions to the advancement of scientific or technical knowledge in Army statistics, ingenious application of such knowledge, or successful activity in the fostering of cooperative scientific matters which coincidentally benefit the Army, the Department of Defense, the U.S. Government, and our country generally.

The Award consists of a medal, with a profile of Professor Wilks and the name of the Award on one side, the seal of the American Statistical Association and name of the recipient on the reverse, and a citation and honorarium related to the magnitude of the Award funds. The annual Army Design of Experiments Conferences, at which the Award is given each year, are sponsored by the Army Mathematics Steering Committee on behalf of the Office of the Chief of Research and Development, Department of the Army.

The funds for the S. S. Wilks Memorial Award were donated by Philip G. Rust, Thomasville, Georgia.

With the approval of President A. Ross Eckler of the American Statistical Association, the Wilks Memorial Medal Committee for 1969 consists of the following:

Professor Robert E. Bechhofer - Cornell University
Professor William G. Cochran - Harvard University
Dr. Francis G. Dressel - Duke University and the Army Research Office-Durham
Dr. Churchill Eisenhart - National Bureau of Standards
Professor Oscar Kempthorne - Iowa State University
Dr. Alexander M. Mood - University of California
Major General Leslie E. Simon - Retired
Dr. John W. Tukey - Princeton University
Dr. Frank E. Grubbs, Chairman - U. S. Army Aberdeen Research and Development Center, Aberdeen Proving Ground, Maryland

BIOGRAPHICAL SKETCH. Dr. Youden was born in Townsville, Australia, on April 12, 1900. Two years later his father returned to his birthplace, Dover, England, with his wife and young son; and the three resided there April 1902 - June 1907. During these years a sister, Dora Alice, and brother, Harry, were born. In 1907, the family of five set out for America, and entered the United States through the Port of New York in July 1907. They lived for a while at Ivoryton, Connecticut, and at Niagara Falls, New York, where Jack attended the local public schools; then they moved to Rochester, New York, in 1916, for Jack's senior year of high school. Youden spent the years 1917-1921 at the University of Rochester, except for one brief interruption to serve his new country as a private in the U. S. Army, October 15 - December 12, 1918. At the University of Rochester, Jack was elected to the National Phi Beta Kappa honor society, and was awarded a B. S. in Chemical Engineering in June 1921. The following academic year, 1921-22, he continued at the University of Rochester as an instructor in Chemistry, then went the two succeeding years, 1922-24, to Columbia University as a graduate fellow in chemistry, earning an M. A. (Chemistry) in 1923; and a Ph.D. (Chemistry), in 1924.

Immediately following receipt of his doctorate, Dr. Youden joined the staff of the Boyce Thompson Institute for Plant Research in Yonkers, New York, as a Physical Chemist. He continued with the Institute in this capacity, with two short leaves of absence and one 3-year assignment as an Operations Analyst with the Army Air Force, until he joined the staff of the National Bureau of Standards in May 1948 as Assistant Chief of the Statistical Engineering Laboratory, which was then beginning its second year of existence.

Dr. Youden was often heard telling a "client" in consultation on statistical aspects of experimentation, or an audience at one of his well attended lectures on statistical methodology, that he is a "chemist," implying, it would appear, that he is really not a statistician. Well,

Youden may have been all chemist for his first seven years at the Boyce Thompson Institute, but by September 1931 he had already begun to dish out advice on the statistical aspects of experimentation. The evidence is to be found in his paper entitled, "A Nomogram for use in connection with Gutzeit arsenic determinations on apples," published in Vol. 3, No. 3 of the Contributions from the Boyce Thompson Institute, pp. 363-374. And from impeccable authority we learn that during the academic year 1931-32 he commuted on his own volition from Yonkers to Morningside Heights in New York City to attend Professor Harold Hotelling's lectures on "Statistical Inference" at Columbia University. He was on his way to becoming an expert on statistical aspects of experimentation. From then on he became more and more of a statistician.

The paper that was ultimately to make his name a laboratory, if not a household word, saw publication in early 1937: "Use of Incomplete Block Replications in Estimating Tobacco Mosaic Virus" (Contributions from Boyce Thompson Institute, Vol. 9, No. 1, pp. 41-48). Here he gave examples and illustrated the application of a new class of symmetrical balanced incomplete block designs that possessed the characteristic "double control" of Latin square designs, without the restriction that the number of replications of each "treatment" (or "variety") must equal the number of "treatments" (or "varieties"). This paper and its new designs led to Dr. Youden obtaining a Rockfeller Fellowship that enabled him to take his first leave of absence from Boyce Thompson, and to devote the academic year 1937-38 to further work in the field of experiment design under the direction of R. A. Fisher himself at the Galton Laboratory, University College, London. Youden's new rectangular experiment designs, termed "Youden Squares" by Fisher and Yates in the introduction to the first edition of their Statistical Tables for Biological Agricultural and Medical Research (1938), were found immediately to be of broad utility in biological and medical research generally; applicable but of less value in agricultural field trials; and with the coming of World War II, Youden Squares proved to be of great value in the scientific and engineering experimentation connected with the research development activities of the war effort of the British and their allies.

Following Pearl Harbor, Dr. Youden took a somewhat longer leave of absence from the Boyce Thompson Institute to serve as an Operations Analyst with the United States Army Air Forces, 1942-45, first as head of the Bombing Accuracy Section of the Operations Analysis Unit of the U. S. Eighth Air Force in Britain, where he directed a group of civilian scientists seeking to determine the controlling factors in bombing accuracy; then, in the latter part of World War II, he was transferred to the Pacific to conduct similar studies preparatory to the B-29 assult on Japan. Stories are legion among the members of the Operations Research Group of the U.S.A.A.F. Eighth Bomber Command about Dr. Youden's exceptional skill in the invention of novel and the adaptation of standard statistical tools of experiment design and analysis to cope with various problems arising in these studies of bombing accuracy. Some of these military applications

written up for immediate use, and embalmed for posterity in his [illegible]ure, "How to Improve Formation Bombing," Air Force Manual No. 67, [illegible], and in the front material to his "Bombing Chart," Air Force Manual No. 79, April 1945. He was awarded the Medal of Freedom in 1946 for his important contributions to the allied victory.

In 1947 Dr. Youden took his third and final leave of absence from the Boyce Thompson Institute: from May to November 1947 he was employed by Project RAND, Douglas Aircraft Company, Santa Monica, California, as a Consultant on statistical problems in design and use of military aircraft.

As stated earlier, Dr. Youden joined the staff of the National Bureau of Standards on May 10, 1948, as Assistant Chief of the Statistical Engineering Laboratory, Applied Mathematics Division. Three years later he became a Consultant (on statistical design and analysis of experiments) to the Chief, Applied Mathematics Division, a post that he held until his retirement on June 30, 1965. Since then he has enjoyed the privileges of a Guest Worker at the NBS.

During Dr. Youden's first two years at the NBS, a fraction of his salary was underwritten by the Research and Development Division, Office of the Assistant Chief of Staff, G-4, Department of the Army. This involved coordination with Dr. Merrill M. Flood and others at the headquarters office in the Pentagon; visits to Dr. Ellis Johnson's group at Ft. McNair; and, quite characteristically, Dr. Youden took a number of trips to Army research and development installations in various parts of the country, to size-up "the problems" in their actual habitats.

Dr. Youden's first decade at the National Bureau of Standards saw the invention and publication of his two-sample chart for "Graphical diagnosis of inter-laboratory test results" (Industrial Quality Control, Vol. 15, No. 11, May 1959), now called the "Youden Diagram," which has proved to be an indispensable tool in the inter-laboratory test programs on the National Conference of Standard Laboratories that provide continuing surveillance on the central calibration programs of the U. S. National Measurement System.

The early 1960's saw Dr. Youden's exploitation of a class of selected incomplete block designs of block size two for the specific purpose of identification and estimation of the effects of sources of systematic error, the central theme of his paper, "Systematic Errors in Physical Constants," (Physics Today, September 1961).

Not the least among Dr. Youden's assets are the effectiveness with which he communicates both in writing and speaking; his exceptional productivity in both areas; and the inspiration with which he amuses his readers and audiences. During Dr. Youden's almost two decades at the National Bureau of Standards he was the sole author of thirty, and co-author of fifteen published research papers; the sole author of

two books and of seven chapters in other books; and for six years (1954-1959) he authored a highly original bi-monthly column on Statistical Design in the professional journal, Industrial Engineering Chemistry. (These columns have since been brought together and issued in booklet form by the American Chemical Society under the title, "Statistical Design.") During this same period, Dr. Youden gave 211 talks around the country on topics in Statistical Methodology and Experiment Design, under 125 titles, the repetition of some talks being by demand. In addition, he made two lecture tours on behalf of the American Chemical Society, addressed the NBS Scientific Staff Meeting twice, and was called upon repeatedly by the Bureau to address special groups (e.g., high school science teachers). Almost without exception, he was the speaker most highly spoken of afterwards by such audiences.

Dr. Youden's first book, STATISTICAL METHODS FOR CHEMISTS (1951) has had a sale of well over fifteen thousand copies. Together with his "column," this book constitutes one of the best sources of real-life examples of effective applications of statistical principles and techniques in physical-sciences research and development work.

As part of the program of the National Science Teachers Association to place some of the most recent advances in science before junior and senior high school students, Dr. Youden prepared one of the NSTA's Vistas of Science Books, EXPERIMENTATION AND MEASUREMENT (1962). As of July 1969, this booklet has sold over 52,000 copies; and is continuing to sell at the rate of over 1,000 copies per year.

Dr. Youden's total contribution to the art and science of statistics in experimentation is truly impressive. A few weeks before Dr. Youden's retirement from the National Bureau of Standards on June 30, 1965, the Royal Statistical Society elected Dr. Youden to Honorary Fellowship at its annual meeting in London on June 2, 1965. There can be no question that Dr. Youden is a very derserving recipient of the Samuel S. Wilks Memorial Medal for 1969.

THE USE OF A HYBRID COMPUTER TO EVALUATE MAN-MACHINE PERFORMANCE OF COMPLEX VEHICLE CONTROL SYSTEMS

Myrna L. Toivanen, Human Factors Engineering and Simulation, Systems and Research Division, Honeywell Inc., Minneapolis, Minn.

Bernard S. Gurman, Avionics Laboratory, U.S. Army Electronics Command, Fort Monmouth, N. J.

Dr. Erwin Biser, Avionics Laboratory, U.S. Army Electronics Command, Fort Monmouth, N. J.

Summary

This paper describes the role of systematic real-time man-in-the-loop simulations in evaluating man/machine performance of complex vehicle control systems. The simulated system consists of: 1) a hybrid (digital and analog) computer system which is used to simulate the vehicle dynamics and the environmental conditions defining the system state, to drive pilot information displays, and to translate the pilot's control inputs into vehicle responses; 2) electro-mechanical flight displays and/or a computer-addressed cathode-ray-tube (CRT) display which are used to provide the pilot with the information required to perform the defined control task; and, 3) a fixed-base control station which is configured to represent control characteristics of the vehicle being studied. The important considerations in the formulation of the experimental design and schedule are discussed. The performance measures used to evaluate overall system performance are described. The limitations on the application of the simulation study results to the real-world situation are discussed. A summary of the simulation mechanization, methodology, and results of a previous study of helicopter IFR formation flight system requirements is provided as an example of the use of man-in-the-loop simulations to evaluate system performance.

This article has been reproduced photographically from the authors manuscript.

Preceding page blank

THE USE OF A HYBRID COMPUTER SYSTEM TO EVALUATE PERFORMANCE OF COMPLEX VEHICLE CONTROL SYSTEMS

Introduction

The evaluation of the performance of complex systems would be prohibitively expensive if it were necessary to wait until the system were built and only then test it under actual environmental conditions. The cost involved in redesigning and rebuilding the system at this point in its development would probably increase the cost of production systems by at least an order of magnitude. If the system is man-oriented and there is a chance that system inadequacies may result in catastrophic human injury or loss of life, the risk of system failure is too great and evaluating system performance in this manner becomes unthinkable.

Tests of aircraft or spacecraft systems involve unusually high risks, both in terms of human life and equipment loss. It has been necessary, therefore, for spacecraft and aircraft systems and design engineers to develop alternative techniques for evaluating performance of such systems prior to actual flight tests. The most comprehensive and flexible technique which has been developed has been that of computer simulation.

A computer simulation to be used for system performance analysis consists of an approximate model of both the system and its relevant relationships with its environment mechanized in the form of computer programs (see Figure 1). The system and those environmental conditions which affect system performance are defined in terms of mathematical equations. Then computer programs are written to mechanize models via appropriate computers and algorithms.

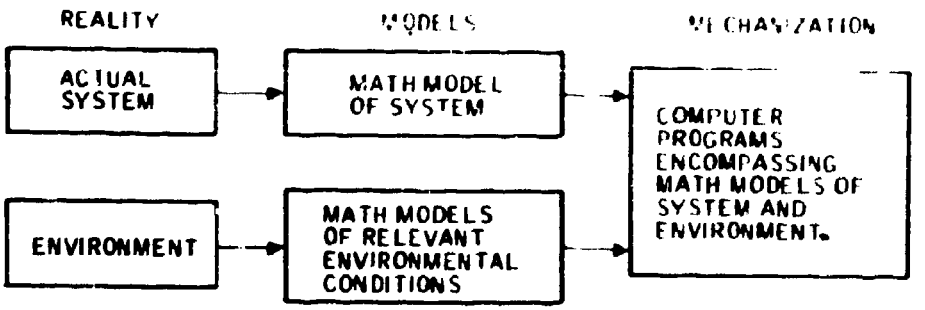


Figure 1. Computer Simulation Development

The simulation of a man-oriented system must include an additional math model for human behavior or must allow the human operator to be included as part of the simulated system. In most complex man-oriented systems the operator's function is a significant factor in total system performance and is too complex to be represented by existing human behavior models. When this is the case, the human operator is included as one of the subsystems of the simulation. Then the simulation must include sensory cues to provide the operator with necessary information and physical apparatus which allows him to perform his function in the simulated system.

In addition to being much less expensive than actual system tests, analysis by simulation is much more flexible. After a simulation is developed, it is possible to systematically investigate system performance resulting from changes in one or a number of the subsystem models or environmental conditions by merely changing the values of the selected parameters in the computer programs. The system design can be changed and tested in much less time and at relatively low cost compared to that involved in rebuilding and reinstalling hardware in the actual system. The system can be tested step by step, analyzing one subsystem while holding other system parameters constant, adding to the complexity of the model only as greater levels of detail are required.

The proper selection of the computers to be used for a simulation depends on the objectives of the study and the nature of the system being simulated. The most appropriate computer is not necessarily that which has the greatest sophistication or capability. Most digital computers are at a point of development now which makes them suitable for use in simulations (i.e., computation speed has been increased to a point where it is suitable for solution of at least the low-frequency system dynamics). The high accuracy, exact repeatability, and flexibility of the digital make it the most appropriate choice for overall simulation control and for simulation of the subsystems which are nonlinear in nature. The analog computer, although not as accurate as the digital, provides continuous solutions and is more appropriate for solving higher frequency system dynamics. It is also more appropriate for simulating characteristics of system hardware which operates in a continuous fashion. Other advantages of the analog for simulation of vehicle dynamics are more operational in nature. The analog can be programmed to represent linear vehicle dynamics more easily and debugged more quickly than the digital computer. Also, it can be changed on-line without going through the somewhat time-consuming processes of making card changes, recompiling the deck of cards, and reloading the program as required for digital programs. Because of the unique characteristics of the analog and digital computers, a combination of these two computer types, or, a "hybrid" computer system, is extremely attractive for the simulation of real-time, man-oriented systems involving vehicle dynamics.

The extensive use of computer simulation techniques by aircraft and spacecraft systems engineers has made the computer an integral part of the systems design and analysis process. This paper explains how and where the computer, specifically the hybrid computer system, fits into the design and analysis process of man-oriented systems involving vehicle dynamics. Although the discussion will be based on this specific class of systems, most of the techniques and methodology discussed would be applicable to evaluation of a wide variety of complex systems.

Manual IFR Formation Flight System Model

The specific system which will be provided as an example is a manual IFR formation flight system. This system is being investigated by Honeywell under the Joint Army-Navy Aircraft Instrumentation Research (JANAIR) Program, a research and exploratory development program to define and validate advanced concepts which may be applied to future, improved Naval and Army aircraft instrumentation systems. This system would allow a pilot to fly at a fixed position (fixed range and bearing) with respect to another aircraft under poor visibility conditions. The primary elements of the envisioned system (Figure 2a) would be: 1) an on-board special-purpose computer to calculate information regarding aircraft position in the required format for display presentation; 2) a display which will present the necessary information for position control to the pilot; 3) the pilot, who must make control inputs based on the information displayed to him; 4) the controls of the vehicle, which will move the vehicle control surfaces as dictated by the pilot's control movements; 5) the vehicle dynamics, which will result in rotational and translational movements; 6) attitude and/or rate sensors on-board the vehicle which will sense the rotational dynamics of the vehicle; 7) an autopilot, which will provide feedback into the pilot's controls and/or the vehicle dynamics; 8) navigation sensors, which will sense the aircraft's position with respect to the lead aircraft; 9) and a filter which will smooth the data obtained from the sensor. As mentioned previously, the simulation of the system must include relevant environmental conditions; two of these are shown in Figure 2a- i.e., system disturbances, such as turbulent wind conditions (10), and the leader flight profiles (11).

Figure 2b shows a block diagram of the organization of Honeywell's hybrid system in the mechanization of the formation flight system model. Honeywell's hybrid system includes a high-speed digital computer with a real-time capability, a number of analogs, a hybrid link for two-way communication between the digital and analog, a cathode-ray-tube which can be used for generation of

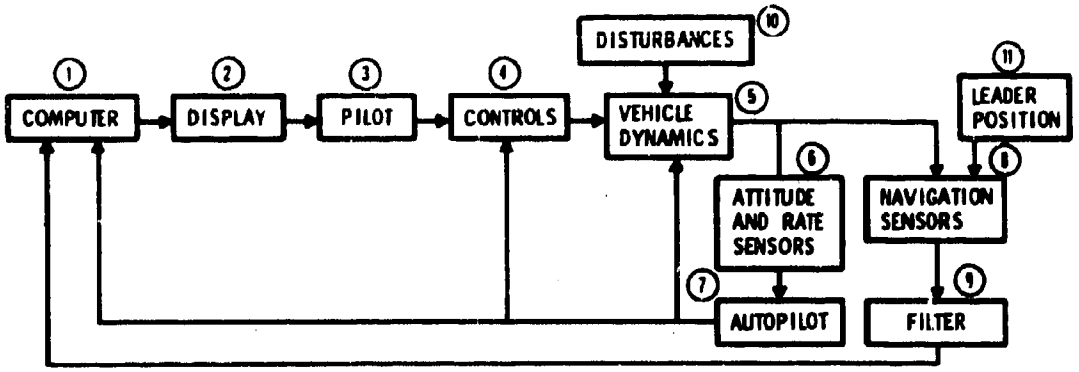


Figure 2a. Manual IFR Formation Flight System Model

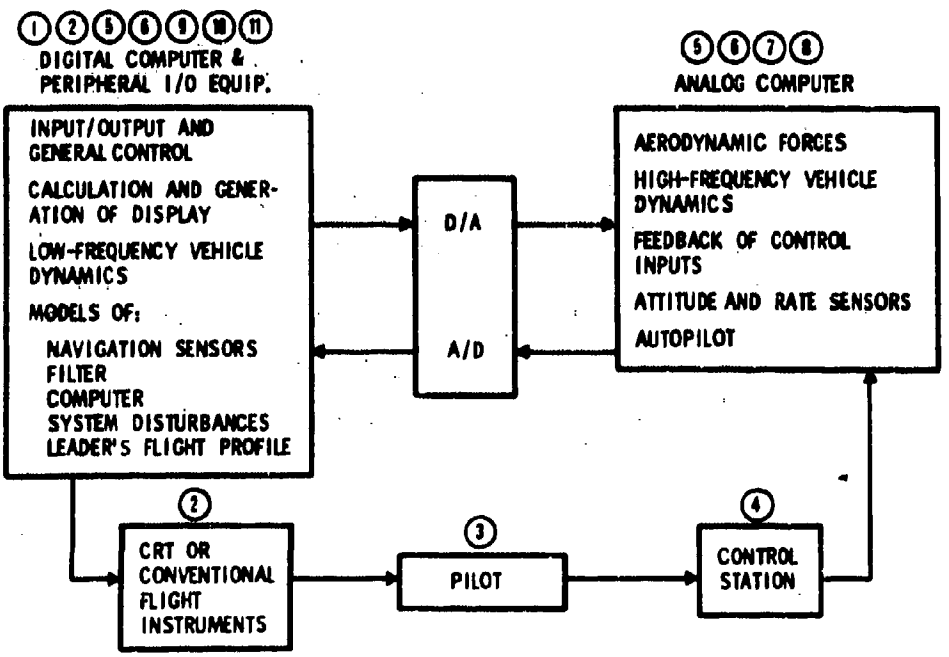


Figure 2b. Simulated IFR FF System

desired display formats (as programmed on the digital), a simulated instrument panel with a number of conventional flight instruments, a fixed-base pilot control station, and standard peripheral equipment which allows communication between the computer and the user. Figure 3 shows the control station located in front of the cathode ray tube display and Figure 4 shows the control station and the simulated aircraft instrument panel.

In the investigations of this simulated formation flight system, pilots have been included in the simulation loop. The flight control and management task which must be performed by the human operator consists of multi-axis aircraft control and three-axis control of position with respect to another aircraft. This task is much too complex to be represented by existing math models of human behavior, which are generally used to represent only single-axis control tasks.

When simulation runs are made, the pilot is seated at the control station in front of a display which provides him with information about his position in space with respect to the lead aircraft. The display can also provide him with aircraft control commands, in which case the position control task becomes primarily one of three-axis tracking. The flight profile of the lead aircraft (programmed on the digital computer) results in the sequential performance of a number of basic maneuvers. The pilot's task is to maintain his position with respect to the lead aircraft throughout these maneuvers. Since this system is to be used for low visibility conditions, the pilot is provided with no visual cues other than the CRT display or specific flight instruments in the simulated cockpit.

System Analysis and Design Process

The process of system analysis and design by simulation and empirical evaluation can be broken down into the following tasks: 1) development of the system



Figure 3. Pilot's Station at Cathode-Ray Tube Display



Figure 4. Pilot's Station in Cockpit Mockup with Electro-Mechanical Flight Instrument Display

model; 2) development of the experimental plan for systematic investigation of system performance; 3) development of the computer simulation; 4) preliminary simulation runs to optimize the simulation; 5) conduct of man-in-the-loop simulations according to the prescribed experimental plan; 6) analysis and interpretation of system performance data; 7) system recommendations and formulation of conclusions. As noted in Figure 5, these tasks are interdependent, with both the sequences and relationships between them being important factors in the effective evaluation of system performance.

One of the mistakes commonly made in the use of simulation techniques is to begin the programming of the computer prior to developing the math models of the system and the experimental plan to be followed in the evaluation of this system. Efficient organization of the program requires at least 90 percent completion of both these tasks prior to development of the simulation programs.

Development of System Model

The definition of the system to be simulated in terms of math models is probably the most difficult task in the system design and analysis process. The difficulty of this task varies with the number of subsystems of the model which have not been previously defined. If the study is a design problem rather than an evaluation of an existing system, the number of such undefined system variables is usually greater. However, even when subsystems are previously defined in terms of hardware characteristics, there are not necessarily existing math models to describe them. Thus the development of a number of math models is usually necessary for any system analysis problem.

The complexity of the simulation to be developed depends on the number of math models required to define the system adequately and the extent of detail necessary for each model. The validity of the simulation, of course, depends on the level of complexity selected for the modeling of the system. Determining the exact level of complexity suitable for a given study is one of the most

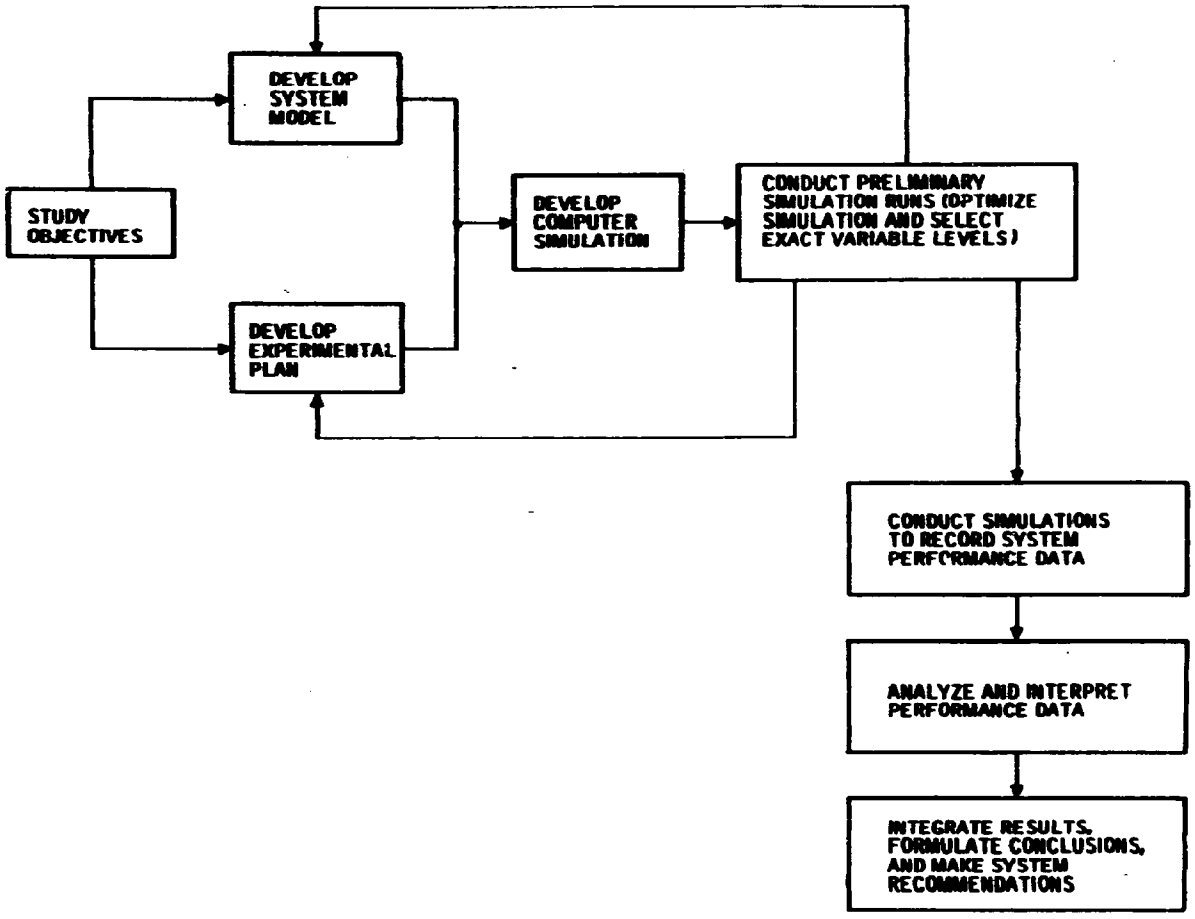


Figure 5. System Analyses and Design Process Using Computer Simulation for Performance Evaluation

difficult chores of the analyst. There must always be a tradeoff between simulation validity and the scope of the study. Determining exactly what this tradeoff should be requires a combination of good judgement and a great deal of previous simulation experience. Unfortunately, as the simulation model approaches exact duplication of the real world, the cost of simulation analysis also approaches the cost of flight testing the actual system.

The subsystems which have been defined for the manual formation flight system model are those shown in Figure 2a (subsystems numbered 1-11). Since the primary objective of this research program has been to develop appropriate cockpit displays for manual IFR formation flight with existing aircraft, the models of the vehicle, the autopilot, and vehicle controls have been predefined at the outset of the investigations. With a human operator as a part of the simulated system, of course, it has not been necessary to develop a math model for the pilot. The remaining subsystem models have been developed during the course of this research program. The greatest amount of time and effort has been devoted to the development of alternative display models. Development of the display model consists of the following tasks:

1) Identifying the information required by the pilot to perform the manual IFR formation flight control task.

2) Incorporating this required information into a total display configuration, which requires:

   a) development of display format (i.e., display symbology) of the primary display,

   b) development of display driving functions (i.e., those equations used to control motion of display elements),

   c) selection of standard cockpit instruments which should be included to provide the required information which could not be incorporated into the primary display.

Two of the display configurations investigated for use in this system are shown in Figures 6 and 7. It may be noted that both display configurations shown provide the pilot with tracking symbols which indicate how he should manipulate the controls to achieve the desired positions. The multi-axis control task required for a high-order control system such as this becomes prohibitively difficult if the pilot must wait for the system's response before receiving feedback regarding the accuracy of his control input. For example, suppose that a pilot makes a roll input to correct a lateral position error. His control input will instantaneously effect a control surface deflection, producing a roll rate, which in turn will result in a significant change in roll attitude within a second or two, which will produce a heading change and lateral rate of movement of the aircraft, which finally, after several seconds, will produce a significant change in the aircraft's lateral position with respect to the leader. In other words, the input response must pass through several integrations before the system finally responds with a change in aircraft position.

To provide immediate knowledge of the results of his control actions, lead information which is based on anticipatory knowledge of the system response must be presented to the pilot. One means of providing this information which has been used successfully in the investigations of this manual IFR formation flight system is display "quickening". Quickening refers to the display of higher-order deriviatives of the system response, which in this case would be time derivatives of the follower aircraft's position errors. Complete quickening, (i.e., presenting the sum of the position error and its derivatives in one element on the display), was utilized to drive the tracking symbols shown in the display formats.

Use of Computer in System Model Development

The computer can be used to aid in developing the system model. Either the digital, the analog, or both can be used for the configuration of a simplified

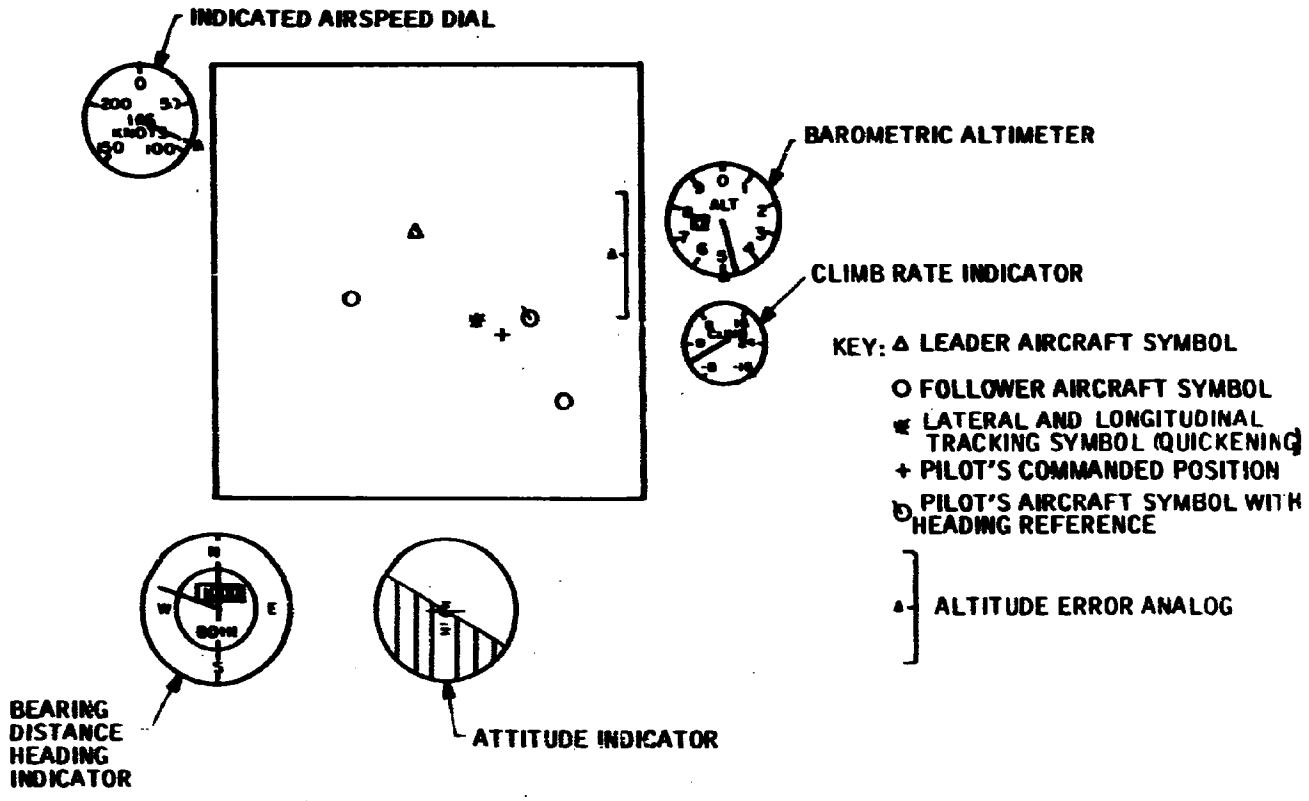
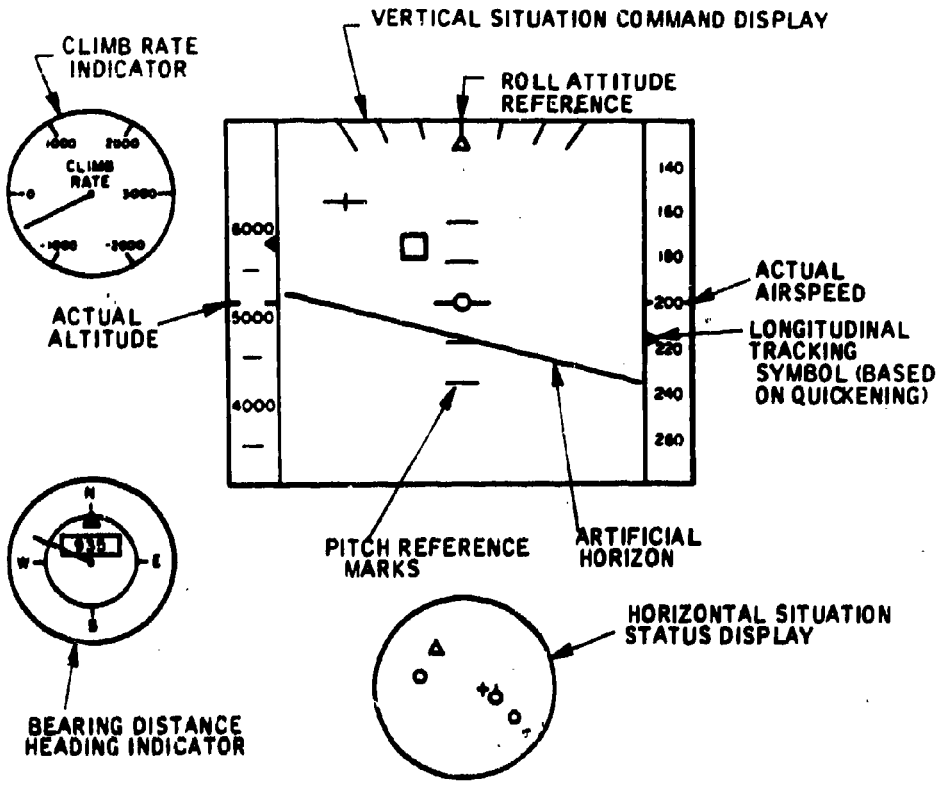


Figure 6. Horizontal Situation Formation Flight Display (Used for the Conventional Helicopter Study)



KEY

VERTICAL SITUATION DISPLAY:

- LATERAL AND VERTICAL TRACKING SYMBOL (BASED ON QUICKENING)
- PILOT'S COMMANDED VERTICAL AND LATERAL POSITION
- PILOT'S OWN AIRCRAFT SYMBOL

HORIZONTAL SITUATION DISPLAY:

- LEADER AIRCRAFT SYMBOL
- PILOT'S COMMANDED LATERAL AND LONGITUDINAL POSITION
- PILOT'S AIRCRAFT SYMBOL WITH HEADING REFERENCE

Figure 7. Vertical Situation Formation Flight Display (Used for the Advanced Rotary-Wing and Jet Fighter Study)

model of the system to empirically evaluate preliminary designs of subsystems. For example, in developing appropriate display driving functions for the display in the manual formation flight system, performance of a simplified system model can be observed to determine which combinations of feedback terms are more appropriate. This simplified model could consist of only one axis of the vehicle control dynamics, a simple model of human behavior (perhaps only a lag) in response to the specified display driving function, and a continuous (analog) recording of vehicle response in terms of position and attitude. Small programs can be written to analytically evaluate specific subsystem math models. For example, to determine the relative velocity required by the follower aircraft to maintain his position with respect to the leader during a turn at various ranges and bearings, the mathematical relationships defining this required velocity are quickly programmed with the ranges and bearings being the variables in the model. Then by merely typing in the desired ranges and bearings, the corresponding relative velocities can be calculated in a fraction of a second.

In addition to the incorporation of the system model, the digital computer programs must include the capabilities of overall simulation control and computer/user communication through appropriate input and output channels. The programs should be organized so that it is easy for the user to change the values of any of the system parameters which are independent variables in the study. They should calculate, record, and output the desired performance measures and be structured to correspond to the experimental procedure to be followed during the course of the man-in-the-loop simulation runs. These additional functions of the computer simulation programs make it necessary that the experimental plan for the study be developed prior to simulation development.

Development of Experimental Plan

The analysis of system performance by simulation, especially when the human operator is included in the simulated system, can involve many hours of simulator runs. If too little time and effort is devoted to planning the experiments, the results of all these hours of simulator runs may turn out to be completely meaningless. No matter how complex or valid the simulation, its usefulness as a tool in evaluating system performance depends on the experimental plan followed in collecting the performance data.

When human performance is one of the contributors to system performance, the system evaluation is based primarily on statistical inference from the performance data collected. The statistical techniques which can be used in interpreting the performance data are necessarily limited by the experimental designs and procedures which have been followed and the exact performance measures which have been recorded. When developing the experimental plan, therefore, it is necessary to select the desired statistical analysis techniques and system performance measures, as well as the experimental design and procedure. The tasks involved and some of the factors which should be considered in developing the experimental plan are shown in Table 1. Specific measures used to describe performance of the manual formation flight system model are shown in Table 2.

Computer Program Development

After the system model and the experimental plan are well defined, it is possible to develop the computer simulation programs in an efficient manner. The digital computer programs are written to incorporate all those subsystem models not programmed on the analog, to calculate and record performance measures, and to allow effective communication between the user and the computer. It should be organized in accordance with the planned

Table 1. Development of Experimental Plan

Identify Independent Variables (IV)

- Will depend on study objectives
- Number of IV's limited by time and money available
- In man-oriented system, human element will be an IV

Identify Dependent Variables (DV)

- Complex system performance evaluation usually requires a number of different performance measures
- When the computer is used to record, calculate, and output performance measures, additional costs associated with additional data are minimal
- Desired statistical analysis techniques should be considered.

Identify Conditions Which are to Remain Constant

- Will include all subsystems of the simulated system which are not independent variables

Determine most Appropriate Experimental Design

- Limited by scope of study
- Should be based on study objectives and nature of the selected IV's

Determine Experimental Procedure to be Followed

- Should minimize learning and order effects
- Should maximize control over experimental constants

Determine Statistical Techniques to be Utilized

- Limited by experimental design and performance measures selected
- Should relate to practical application of study results

Table 2. Performance Measures Used in Analysis of the Manual Formation Flight System

- Mean, standard deviation, and RMS (root-mean-square) error in:
  - Longitudinal position wrt leader
  - Lateral position wrt leader
  - Vertical position wrt leader
  - Range from leader

  These measures are used to indicate the pilot's level of position control precision.
- RMS control stick rates (e.g., $\sqrt{\frac{\sum_{i=1}^{N} X_i^2}{N}}$, where X_i is rate of control stick deflection in inches or degrees per second). This measure is used as an indication of the extent of pilot control activity.
- Collisions with other aircraft
- Continuous time histories of position error, aircraft attitude, and control inputs
- Proportion of time subject utilizes a given display
- Number of attention shifts between displays

experimental procedure to allow conduct of the simulation to be as efficient as possible and to free the experimenter from trivial tasks such as timing or data recording. It should be organized such that changing values of system variables can be accomplished easily and quickly.

The analog computer is wired to incorporate those subsystem models which can be more conveniently programmed and debugged on the analog (such as rotational vehicle dynamics) and the control station is configured to represent the controls of the vehicle being simulated.

Once the simulation is developed, it can be used over and over again for a long period of time. It should, therefore, be well-organized and documented so that it can be used and/or modified in the future with minimal difficulty The digital computer programs should be modular in design for ease of checkout and modification. The wiring diagram of the analog simulation should be detailed and complete. The program development should not be considered complete until the program is well documented.

After the entire system model is mechanized in the form of computer programs, the simulation is operational and simulator runs can be conducted.

Preliminary Simulations - System Optimization

Preliminary simulation runs are required to optimize all subsystem models based on man-in-the-loop system performance, to optimize simulation procedures, and to experiment with the ranges of the independent variables of the study. It is usually only after this phase of the systems analysis process that the system model is totally defined. Although the basic math models for the system must be developed prior to writing the computer programs, it is usually desirable to experiment with the parameters which define the exact characteristics of a given model after the simulation is operable. For example, the model of a digital filter may be specifically characterized

by parameters which effect differential lags and variance reductions. The optimal values of these parameters are dependent on the specific system application. For complex, man-oriented systems it is usually easier to determine the most appropriate values for such subsystem parameters empirically by means of preliminary simulator runs.

When system response depends on the performance of a human operator, there is often no analytical way to accurately estimate the effect of specific system designs or parameter variation on total system performance. When this is the case, the only way to determine whether a given system model is functional is to test it empirically with a human operator in the simulation loop.

The preliminary simulation phase of the study provides the analyst with rough estimates of system performance. As a result of this preliminary investigation, he can select optimal values for parameters of the subsystem models, select reasonable ranges for the independent variables of the study, and redesign parts of the system if necessary. For example, in the investigations of the manual formation flight system, it was found that the position control task was unreasonably difficult when the display provided only position error information. The analyst experimented with various feedback terms and found that the addition of position error rate and aircraft attitude terms to the position error information made the task much easier. Such preliminary investigations prove to be very valuable in maximizing the effectiveness of the formal experimental tests conducted in the next phase of the system analysis process.

Systematic Man-in-the-Loop Simulation for Collection of Performance Data

After the computer simulation has been completely developed and optimized through preliminary simulator runs, it is finally in the appropriate stage to

begin the systematic tests required for performance evaluation. Although conducting real-time man-in-the-loop simulator runs may be the most time-consuming portion of the system evaluation, if preceding phases of the system analysis process have been conducted with care, it is also generally the most straight-forward part of the study from the system analyst's point of view. This phase of the system analysis process consists of familiarizing the subjects with their task under the various experimental conditions and then conducting the simulated runs as prescribed by the experimental plan. For example, in the investigations of the manual formation flight system, the pilot-subjects "fly" the simulated aircraft through the programmed mission repeatedly under each of the experimental conditons until very little further improvement is noted in their position control performance. Then each subject "flies" the required number of missions for the various experimental treatments in the exact sequence prescribed for him by the experimental plan. System performance data are recorded and output for each simulator run.

The most important points to remember in this usually extensive and repetitive process of system testing is that the value of the results depends on strict adherence to the prescribed experimental plan and frequent checks on the accuracy of the simulation programs. The experimenter must guard against becoming lax in his experimental procedure and should frequently, preferably before each simulator run, perform a diagnostic check on the simulation to assure that the programs are working correctly. Although exact repeatability is a characteristic of the digital computer, it is not so for the analog or the link between the digital and the analog. The analog and digital may be improperly connected, some of the potentiometers may be set incorrectly, there may be an amplifier or integrator which is not plugged in securely or not working correctly because of some defect, etc. A diagnostic check can be performed which will usually identify these problem areas quite quickly.

After all prescribed simulator runs have been completed, the system performance data is sorted and collated for interpretation and analysis in the next phase of the system evaluation.

Statistical Analysis and Interpretation of Performance Data

There are numerous statistical techniques which can be employed in analyzing the performance data. The most appropriate techniques depend, of course, on the study objectives and the nature of the system being evaluated. As previously emphasized, the techniques to be used should be selected when the experimental plan and procedure are being defined, to assure that the performance measures recorded and the plan followed will allow valid application of the desired techniques.

The digital computer can be used to perform the lengthy calculations required for the statistical tests. Most large digital computer facilities have a number of general-purpose statistical programs available for performing the more commonly used and well-known statistical tests. A number of such programs are available at Honeywell and have been used in analyzing data obtained in tests of the formation flight system.

When the computer is used both for calculating and recording data and for performing the statistical tests, the added expense associated with recording additional performance measures and/or performing a number of different tests on this data is relatively low. Since it is often difficult for the analyst to determine which performance measures are more appropriate until after the data has been collected and analyzed, it is a good idea for him to record all those measures which seem to be relevant. If digital computer programs are available for performing the desired statistical tests, the same philosophy can be followed in determining the number of tests to be applied to the data. Even when the scope of the study limits the statistical analysis effort, it is

important that sufficient performance data be collected and saved. It is often desirable to perform further tests later as additional funding becomes available. Also, it may be that a future study with different objectives requires different methods of data analysis but could be based on the same system performance data. Some of the statistical techniques which have been used in evaluating the manual formation flight system are shown in Table 3.

After results of the statistical tests are obtained, they must be interpreted in terms of their implications in designing or developing the system. Sometimes results which are statistically significant are not significant from a practical systems application standpoint. For example, suppose that the results of a statistical comparison of two displays for the manual formation flight system showed that lateral position control performance was significantly better for one display, and that the average position errors for this display were consistently five feet lower than those for the other display. No matter what the level of statistical significance, this small difference in position error may not be of practical significance in terms of system development. If performance resulting from two such displays were this similar, the system analyst would probably recommend that display which would be less expensive to incorporate into the system. Since the only results of real significance to the system analyst are those which can be related to the design or development of the system, it is important that sufficient time be devoted to interpreting the results of the statistical tests accordingly.

After the analyst has decided what the important study results are, these results should be presented clearly in the final study report. Since the graphical form of presenting data is usually more easily interpretable for the reader than the tabular, it may be helpful to show at least the more important results graphically.

Table 3. Statistical Tests Used to Evaluate Performance of Manual IFR Formation Flight System

Statistical Procedures:	Based on Following Performance Measures:
Factorial analyses of variance	1) RMS position error 2) RMS control stick rate 3) Mean position error
Calculation of mean and/or median position errors for factorial combinations of independent experimental variables.	Mean position error for a specific subject, treatment, and maneuver.
Calculation of mean and/or median variances for factorial combinations of independent experimental variables.	Standard deviations around the mean position error for a specific subject, treatment, and maneuver.
Calculation of correlation and/or regressions between system variables.	E. g., between lateral and longitudinal position errors, or between fore-aft and right-left control stick movements, etc.

Formulation of Conclusions and System Recommendations

The final step in the system evaluation process is the careful examination of study results to formulate conclusions and make system recommendations. Sometimes in the early phases of the system investigation, recommendations regarding system implementation--such as required hardware characteristics--cannot be made until further research has been conducted. The study conclusions at this point usually relate primarily to the feasibility of system concepts. The primary system recommendations made will be suggested areas for further investigation, utilizing those effective system concepts and models which have been developed during the previous system studies as a basis for future studies. The system analysis process described in this paper would then be repeated a number of times until sufficient system aspects have been investigated to allow recommendations regarding hardware characteristics and specifications.

Results and conclusions of the formation flight system research are provided below as an example of the steps which must be completed prior to actual system development and the type of conclusions which can be drawn from system evaluation studies based on computer simulation analysis.

Results and Conclusions of Research on Manual IFR Formation Flight System

The research on the formation flight system has not yet reached the point where the exact specification of required hardware characteristics would be desirable. The philosophy followed in this system investigation has been to first determine the feasibility of basic system concepts and associated system performance, assuming no limitations imposed by system hardware. Then, one by one, the system limitations imposed by realistic hardware characteristics and expected environmental conditions

have been added and resultant system performance evaluated. Since there are so many system variables in a complex system such as this, it has not been desirable, either from a cost or operational standpoint, to investigate all these variables simultaneously. Instead, a number of different studies have been conducted and each new study has utilized results of previous studies as a basis for evaluating additional system variables. The studies which have been performed to investigate the manual IFR formation flight system and the major results of these studies are described briefly below.

The first study (Reference 1) of this program was conducted to investigate basic information requirements for the manual IFR formation flight control task for rotary-wing aircraft and to evaluate display concepts for a computer-addressed cathode-ray tube display. The effects of turbulent wind conditions and subsidiary pilot workload were also investigated in this study. This study assumed no limitations on sensor outputs (i.e., high data transmission rate and no measurement noise). Some of the study results are shown in Figures 8 through 11. The major conclusion of the study was that formation flight under IFR conditions appears to be a realizable goal with the aid of the computer-generated display formats developed.

The second study (Reference 2) was conducted to evaluate an existing helicopter formation flight system. This study assumed the sensor, computer, and display characteristics of the system being evaluated. Results of this study demonstrated the important effects of filtering techniques, data transmission rate, and display driving functions on total system performance (see Figures 12 through 14).

A third study (Reference 3) was conducted to evaluate the effectiveness of conventional flight instruments in a manual IFR helicopter formation flight system. Two state-of-the-art electro-mechanical displays, i.e., a flight director and a horizontal situation indicator, were used in conjunction to display the required information and were evaluated under alternative display formats. This study again assumed no limitations on sensor system outputs.

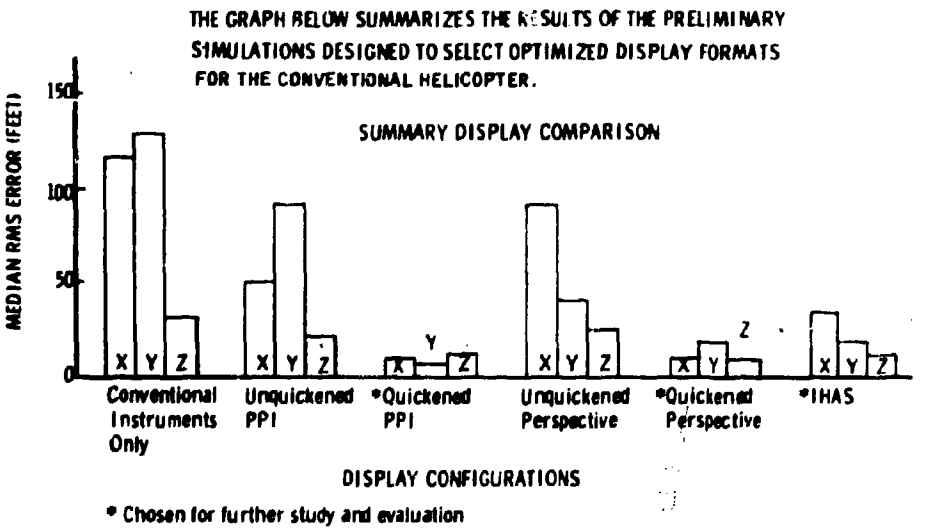


Figure 8. Original Study of Concept Feasibiltiy and Display Requirements for the Conventional Helicopter

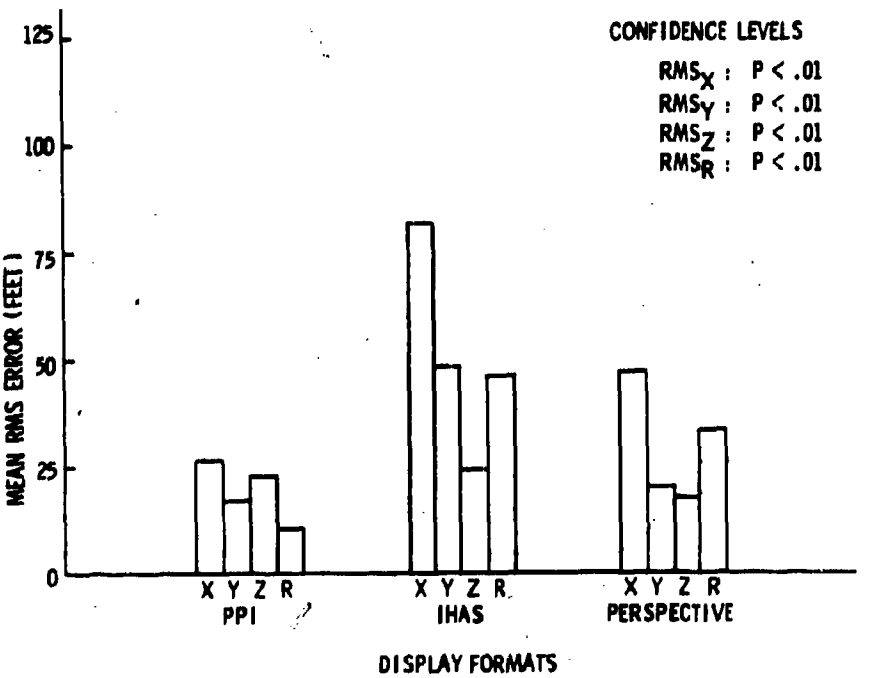


Figure 9. Comparative Evaluation of Display Formats for Conventional Helicopter Study

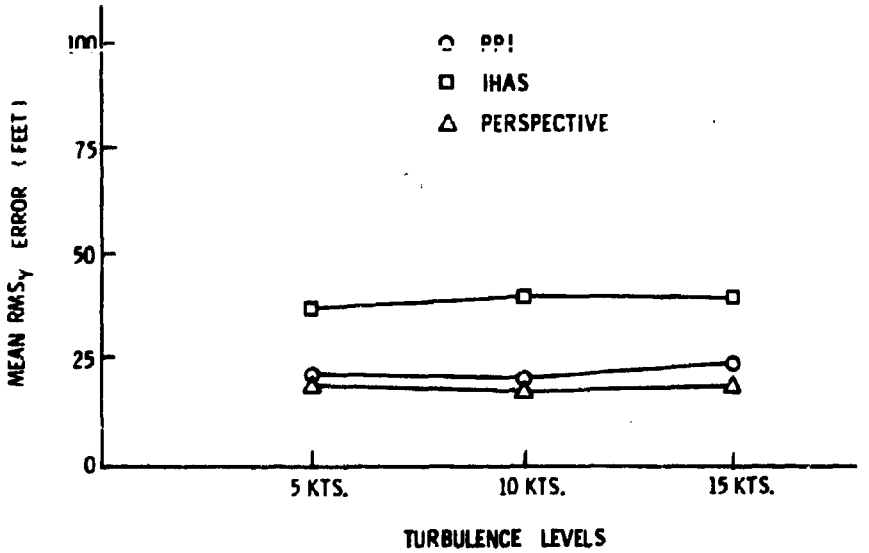


Figure 10. Effects of Turbulence on System Performance - Conventional Helicopter Study

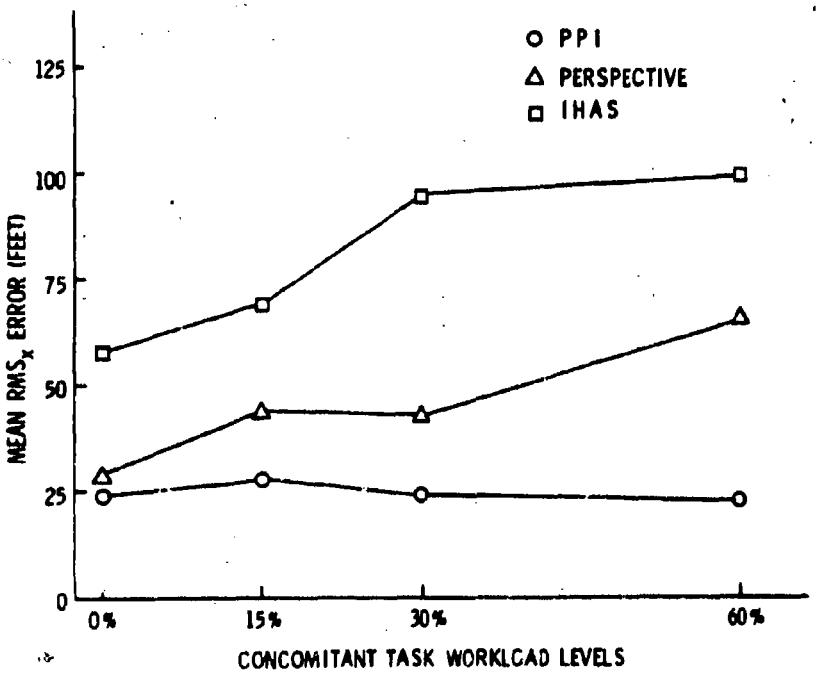


Figure 11. Effects of Subsidiary Pilot Workload on System Performance - Conventional Helicopter

THE GRAPH BELOW SHOWS MEAN RMS ERRORS PLOTTED AS FUNCTIONS OF THE SYSTEM UPDATE TIME INTERVAL. SYSTEM UPDATE RATE IS THE RECIPROCAL OF SYSTEM UPDATE TIME INTERVAL.

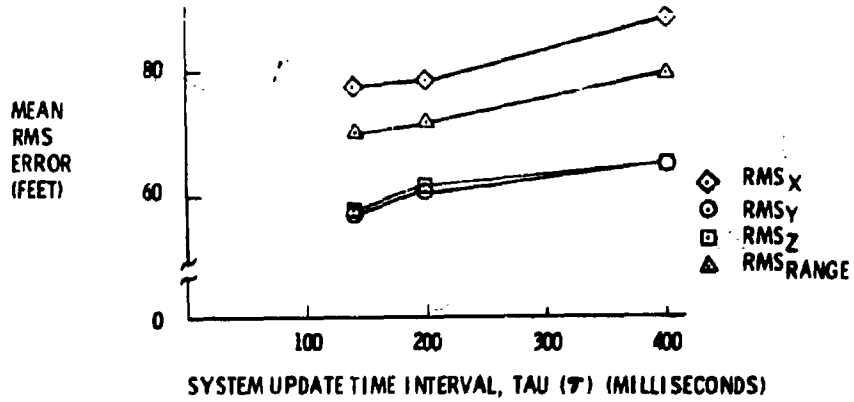


Figure 12. Sample Prototype System Evaluation - Effect of System Update Rate on System Performance - Conventional Helicopter

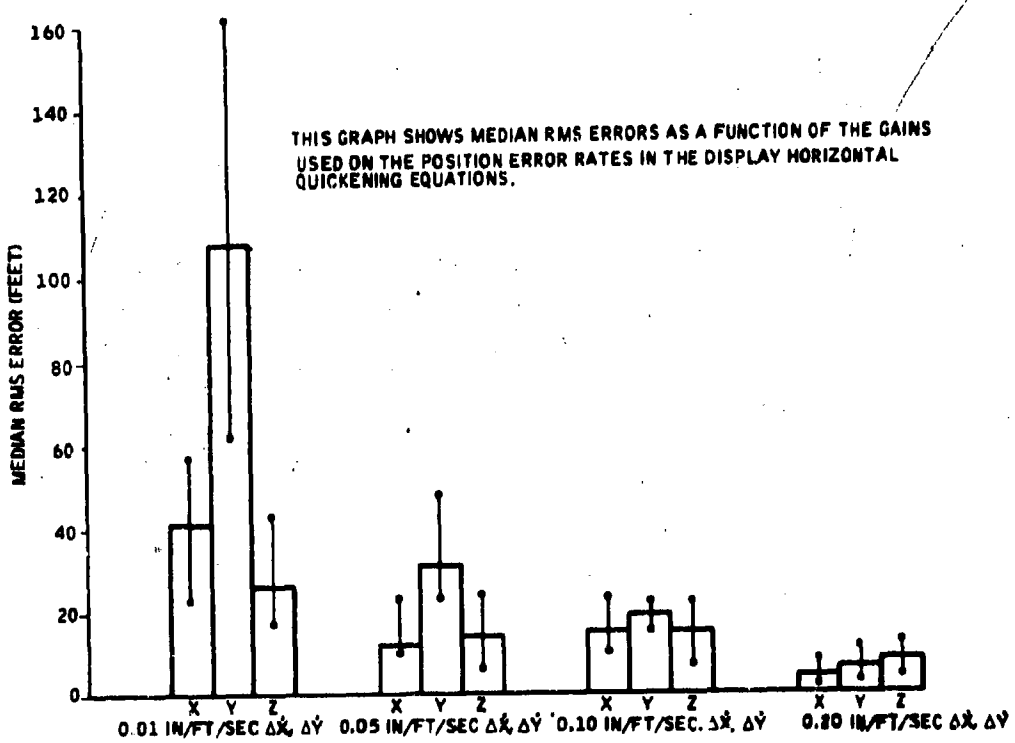


Figure 13. Sample Prototype System Evaluation - Effects of Terms in Display Quickening Equations on System Performance

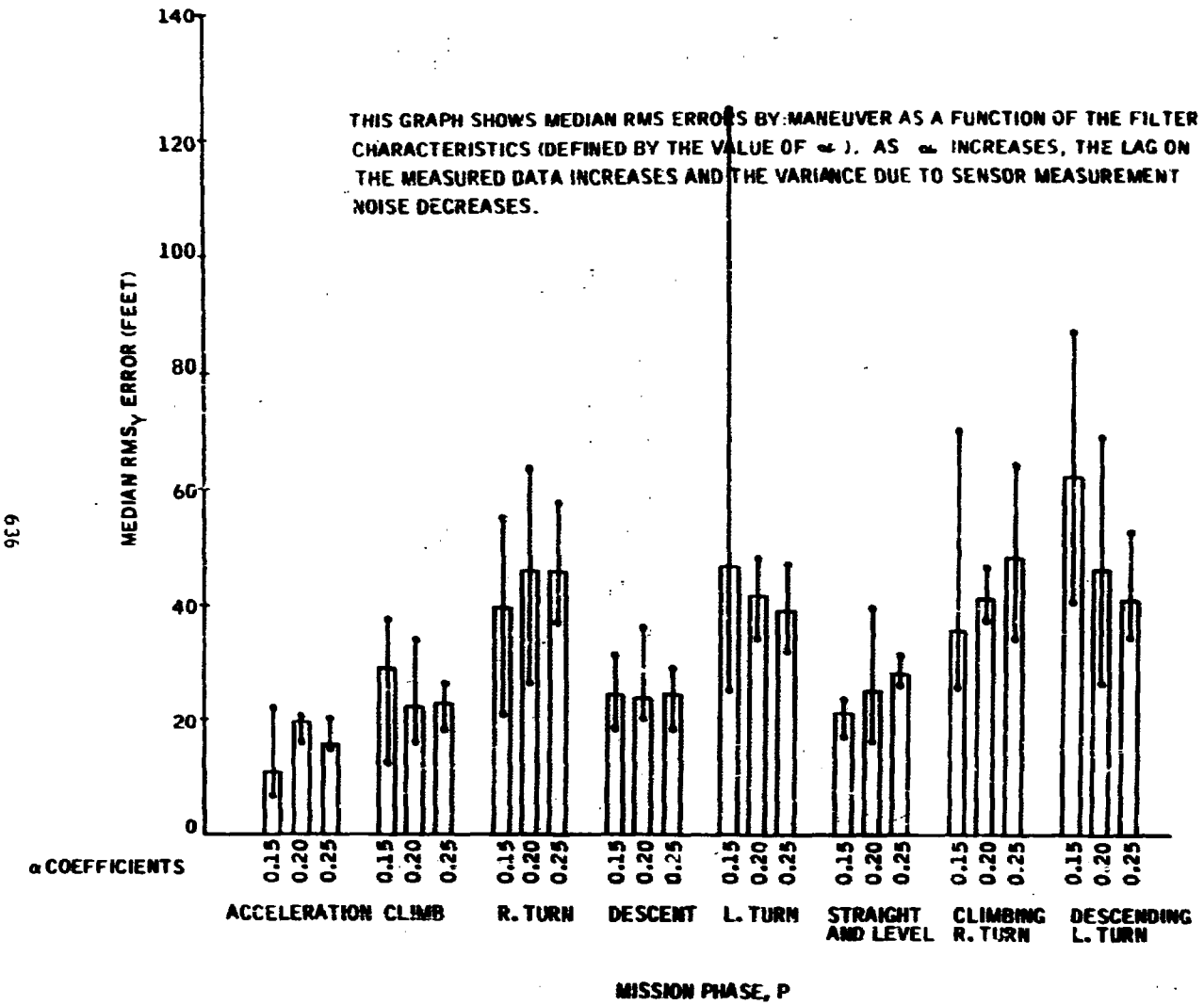


Figure 14. Sample Prototype System Evaluation - Effect of Filter Characteristics - Helicopter Study

The results of this study (see Figures 15 and 16) indicated that it was possible to obtain position control with the electro-mechanical displays comparable to that obtained with the computer-generated displays.

The fourth study (Reference 4) investigated the relationship of variations in the effective data transmission rate (i.e., defined as the update rate of the information presented on the displays) and the effective level of measurement noise (i.e., noise which appears on the display after filtering) on pilot/system performance in the helicopter IFR formation flight mode. Results (see Figures 17 through 19) showed: 1) that position control performance degraded with increasing measurement noise; 2) that position control performance did not improve significantly by increasing the update rate above 4/second; and 3) that optimal display driving functions and data filtering techniques are dependent on the data update rate and accuracy characteristics of the system.

The objective of the fifth study (Reference 5) was to evaluate the effect of varying levels of automatic control assistance on pilot/system performance in the simulated helicopter IFR formation flight mode. An information update rate of 4/second and a moderate noise level were assumed throughout this study. Results of the study (see Figures 20 and 21) indicated that increasing the level of automatic control assistance provided greater system stability and made the pilot's control task less demanding, but did not significantly improve position control performance over that obtainable manually with the aid of the quickened display.

The sixth study (Reference 6) was conducted to define information and display requirements and investigate variable sensor output characteristics for two additional vehicle classes, i.e., the advanced rotary-wing and the jet fighter aircraft. The results (see Figures 22 through 25) indicated that manual IFR formation flight with the envisioned system appears to be feasible for the advanced rotary-wing and the jet fighter aircraft and that the effects

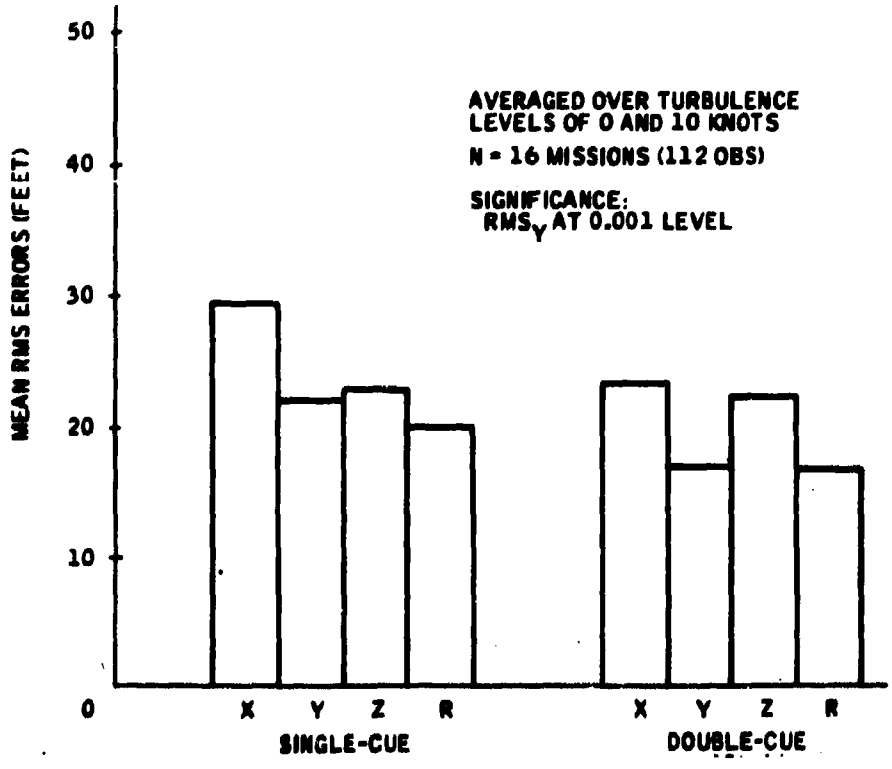


Figure 15. Comparative Evaluation of Two Flight Director Displays for Conventional Helicopter

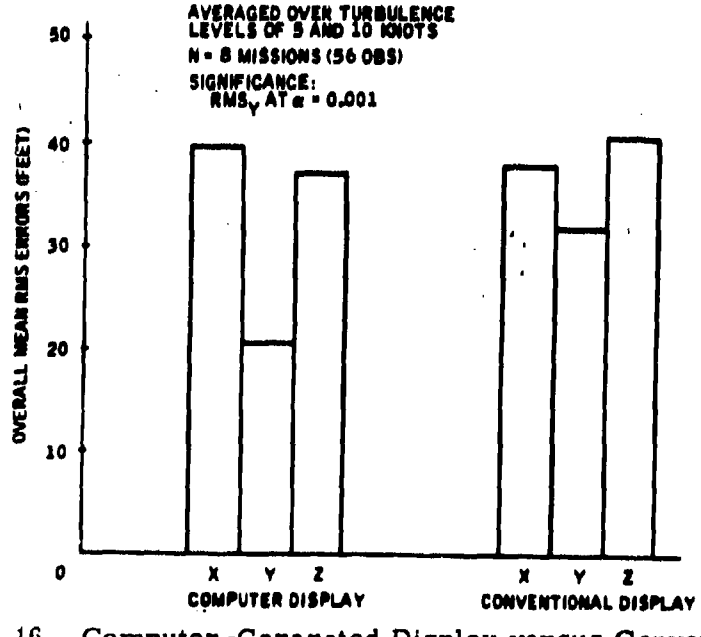


Figure 16. Computer-Generated Display versus Conventional Flight Director Display - Conventional Helicopter

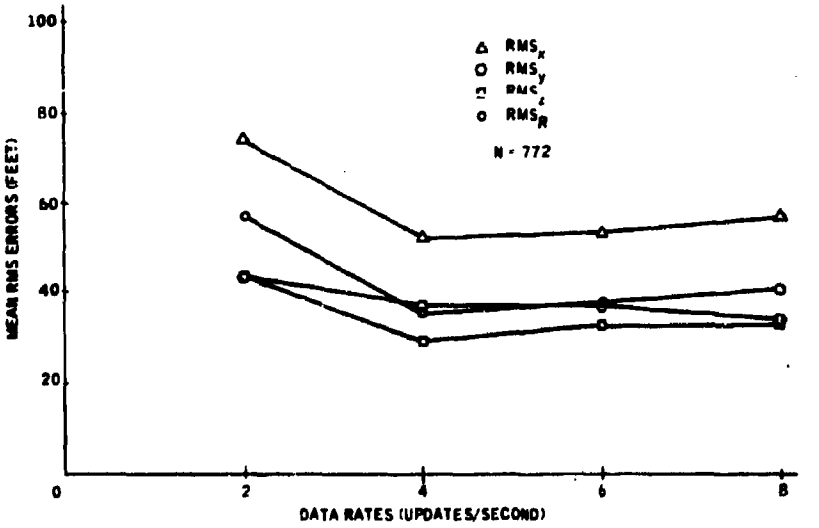


Figure 17. RMS Position Errors by Data Rate - Conventional Helicopter

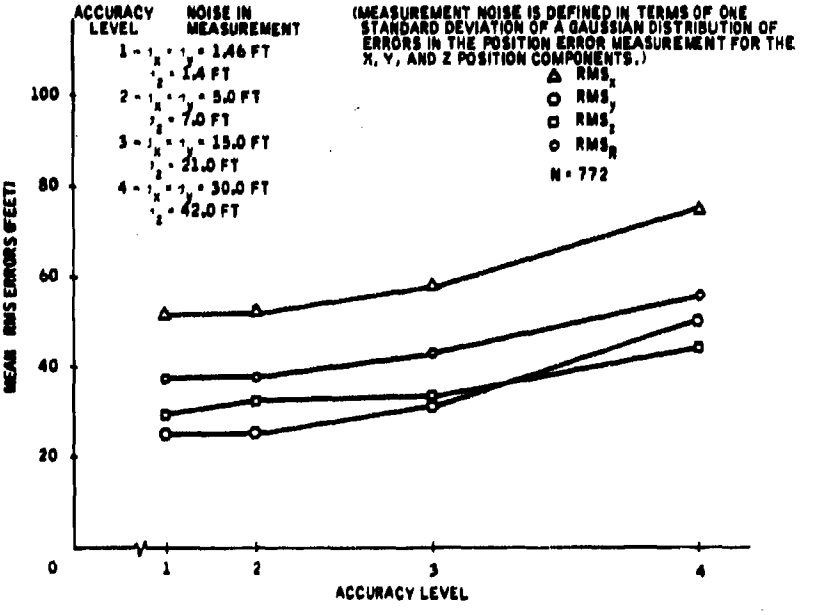


Figure 18. RMS Position Errors by Data Measurement Noise Level - Conventional Helicopter

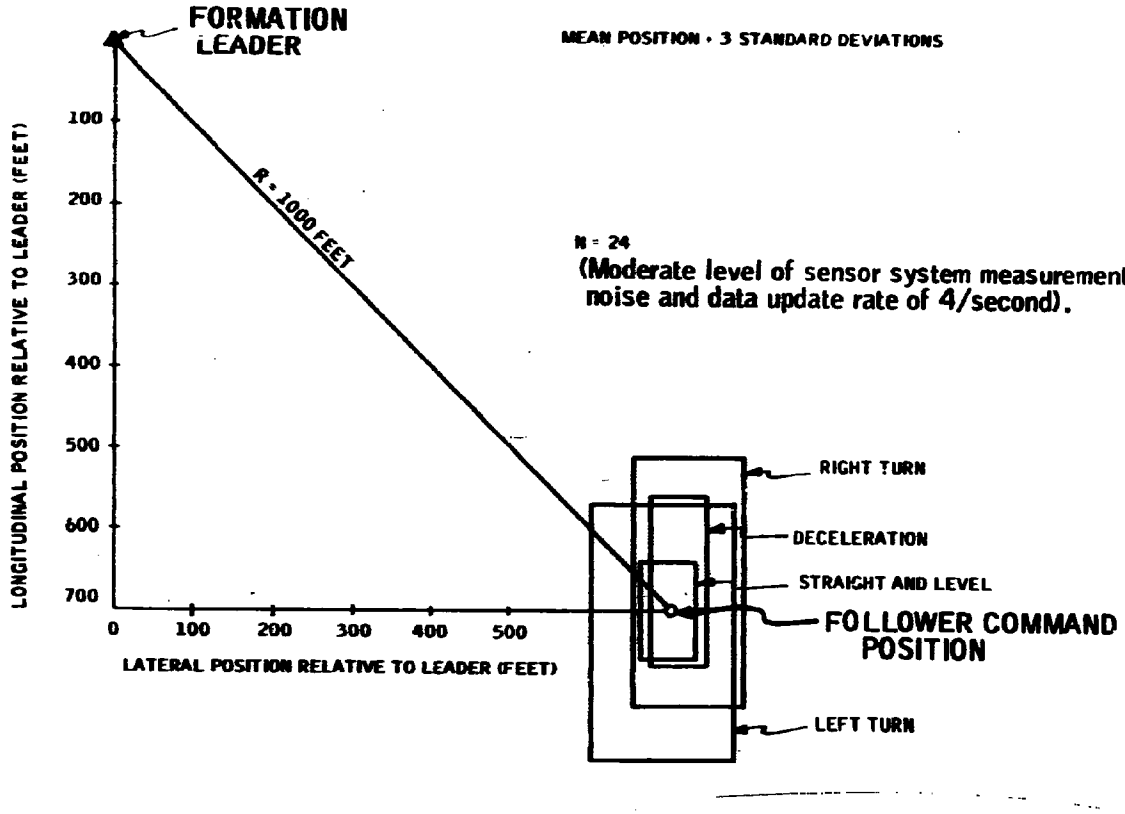


Figure 19. Horizontal Position Error Envelope for Conventional Helicopter

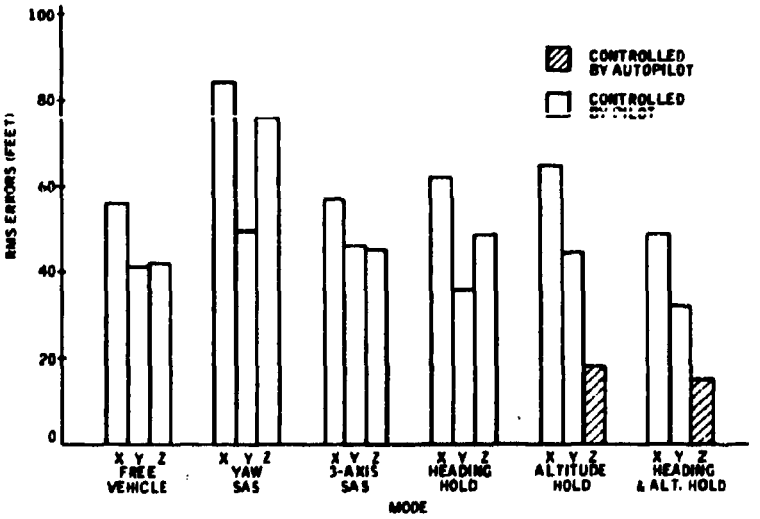


Figure 20. Mean RMS Errors by Autopilot Mode for Conventional Helicopter

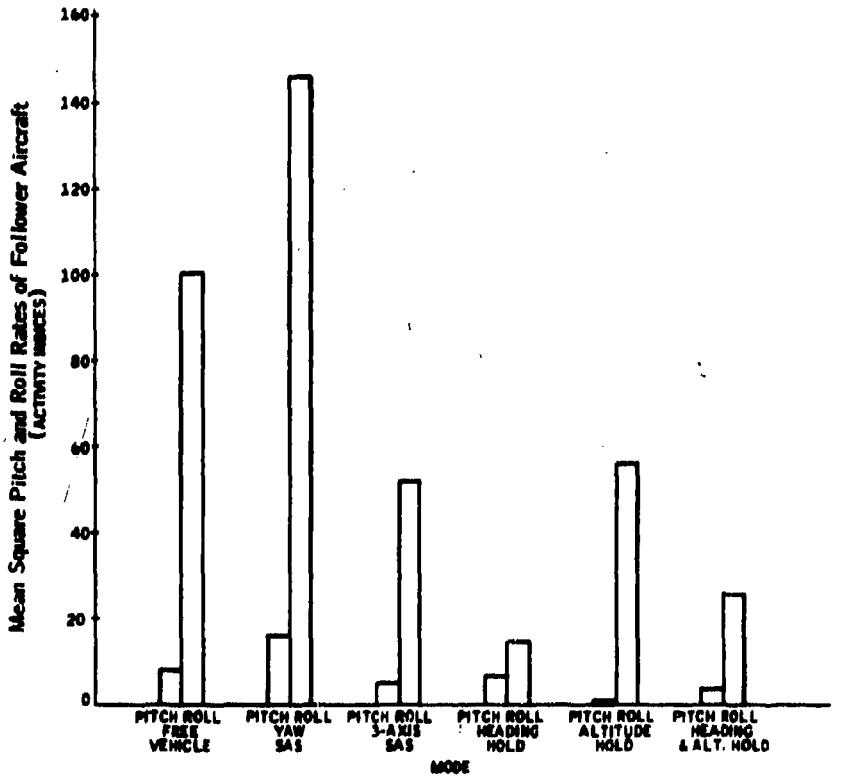


Figure 21. Aircraft Pitch and Roll Activity as a Function of Level of Automatic Assistance - Conventional Helicopter

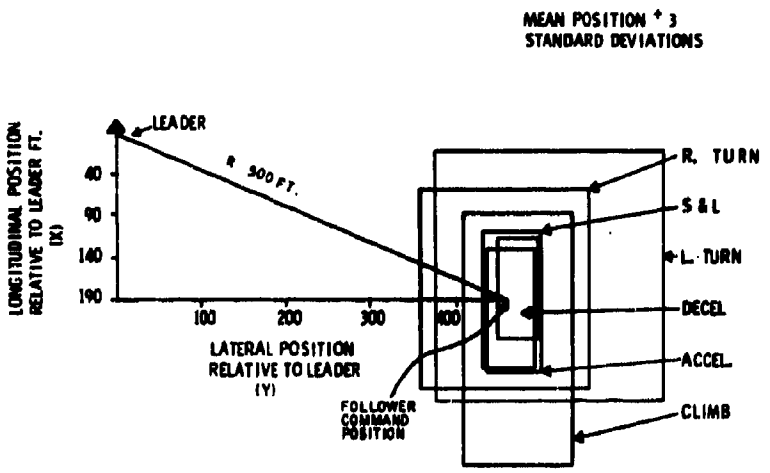


Figure 22. Horizontal Error Envelope by Maneuver for the Advanced Rotary-Wing

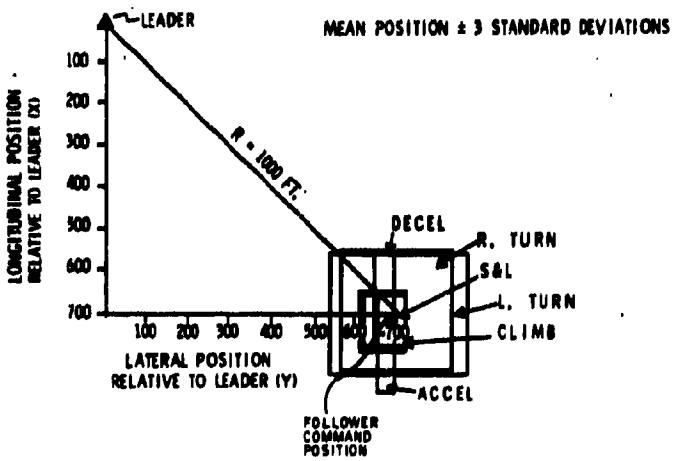


Figure 23. Horizontal Position Error Envelope by Maneuver for the Jet Fighter

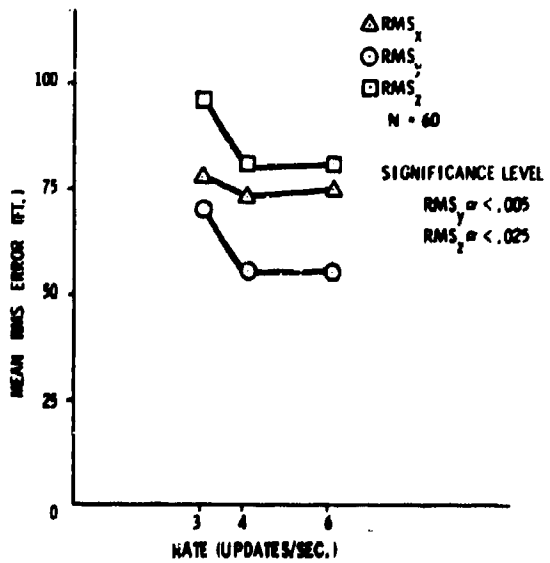


Figure 24. Effects of Data Update Rate on System Performance - Jet Fighter

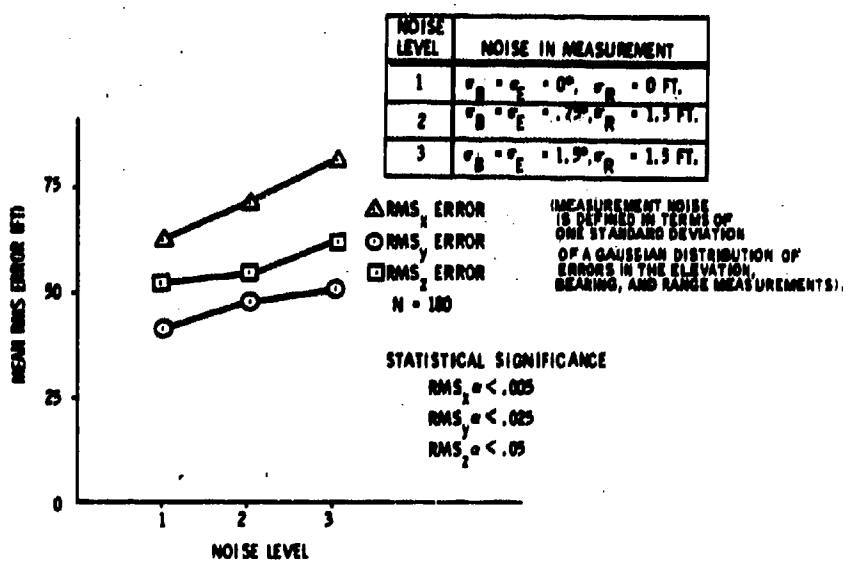


Figure 25. Effects of Sensor Measurement Noise on System Performance - Advanced Rotary-Wing

on system performance of variation in measurement noise and display update rate were the same as found previously for the helicopter.

Some of the basic conclusions which have been drawn as a result of these research studies are summarized below.

Display driving functions are more important than the display format in determining the pilot's control performance. The results of the various display evaluations suggest that as long as the basic display/control relationships are satisfied, all required information is presented, and the display formats are interpretable, a number of different display formats (in terms of specific symbology and orientation) are appropriate for the envisioned IFR formation flight system.

Manual IFR formation flight with the system as modeled for these research studies appears to be a realizable goal for the conventional helicopter, the advanced rotary-wing, and the jet fighter. The level of position-control precision to be expected for a specific aircraft class, given the presentation of all required information, the use of an appropriate display format, and optimal display quickening, will be a function of at least the following variables:

1) Rate at which new information is available for display
2) Level of measurement noise on positon information.
3) Filtering technique used to smooth the noisy data
4) Experience level and capability of the individual pilot
5) Extent of pilot workload required for subsidiary tasks.

It is suspected that other variables, such as pilot fatigue, aircraft separation distances, formation geometry, formation airspeed, command mode (i.e., fly-to or fly-from), lead aircraft perturbation, etc., also effect

postion-control performance. Current research under the JANAIR program includes investigation of a number of these additional system variables. Prior to exact specification of the hardware required for the manual IFR formation flight system, these and other seemingly relevant system variables should be investigated.

The Value of Simulation in System Performance Evaluation

The conclusions summarized above are somewhat tentative in nature and do not specify the exact level of system performance which can be expected with the developed system under actual flight conditions. It must be emphasized that simulation analysis can not be a substitute for actual system tests. Rather, simulation is a compromise method of evaluating system performance with a degree of validity which falls somewhere between preliminary pencil-and-paper analyses and tests of an actual breadboard system. Exactly where it falls in terms of validity depends on the complexity of the system model which is developed for the simulation.

The value of simulation as a system evaluation tool is greatest for those systems which cannot be easily defined analytically (such as those involving complex human behavior) and which cannot be tested under real-world conditions without extremely high risks (i.e., in terms of human life and/or system costs). For this type of system, compromise methods of system evaluation are necessary, and they are not meant to replace actual system tests, but to precede and minimize the extent of these tests.

Although there are inherent constraints and limitations in the evaluation of system performance by simulation analysis, it is extremely useful in answering questions such as the following:

- Is a given system design or concept feasible?
- How does one method or design compare to another?
- What are reasonable minimum and maximum limits on the variation of a given system parameter?
- What are the effects of varying two or more system parameters simultaneously?
- Can the human operator perform the required task?
- Which system tasks should be performed automatically?
- What is the general relationship between system performance and a given system parameter or environmental condition?

If the system analyst conducts his system investigation scientifically and if he tempers his formulation of conclusions and system recommendations by acknowledgement of the constraints and limitations of the simulation, he will find computer simulation techniques to be an invaluable tool in the system analysis and design process.

GLOSSARY

Activity indices	Mean-square pitch and/or roll rates of the simulated aircraft
A/C	Aircraft
Autopilot mode	Level of automatic control included in the aircraft control system. Ranges from simple damping in a single axis of the aircraft to control of the aircraft's heading and altitude.
Concomitant task	A class of secondary task which is performed simultaneously with the primary task and is highly quantifiable in nature. In the study referenced in this paper, the primary aircraft position control task was performed continuously and the pilot's formation flight display was intermittently blanked out. Then secondary task cues were provided, requiring that the pilot simultaneously perform the secondary task at a forced-pace (i.e., the frequency and time interval of the formation flight display interruption was controlled by the experimenter rather than left to pilot discretion). The levels of the secondary workload were defined in terms of percentage of time the pilot was required to perform the secondary task.

Data transmission and/or update rate	The rate at which new information about the follower aircraft's position in space (with respect to the lead aircraft of the formation) is available for display.
Dependent variables	Variables of an experiment which describe system performance. These performance measures are assumed to reflect changes in the levels of the independent variables of the experiment and are thus considered to be "dependent".
Double-cue flight director	Electro-mechanical flight instrument currently used in both fixed-wing and helicopter aircraft to provide information about aircraft pitch and roll attitudes. The "double-cue" flight director has two separate moving elements, one representing pitch and the other roll attitude deviations.
Filter	In the referenced studies the filter simulated was a digital α-β filter. The filter is required to smooth the aircraft position information obtained from the assumed sensor system. Sensor systems are usually characterized by a certain level of measurement noise.
IFR	Instrument Flight Rules -- This term is used in this report to refer to very low visibility conditions when a pilot would have to depend primarily on instruments for visual cues.

IHAS display format	Integrated Helicopter Avionics System - The vertical situation display format which was intended as a part of this avionics system was configured for use in the IFR formation flight mode in one of the referenced studies.
Independent variables	Parameters of a system which are varied to investigate their effect on system performance.
Mean position error	$\sum_{i=1}^{N} X_i/N$, where X_i was the measured position in a specific axis during a specific maneuver and N was the total number of position measurements recorded during the maneuver. For the referenced studies N was equal to the total maneuver length times the display update rate.
Measurement noise	Used to refer to the error in the measurement of the follower aircraft's position with respect to the leader. The sensor system was assumed to include noise with a normal (Gaussian) distribution described by the standard deviation. In the referenced studies these standard deviations were defined in terms of the bearing (σ_B), elevation (σ_E), and range (σ_R) measurements (in degrees), or in terms of the longitudinal (σ_X), lateral (σ_Y), and vertical (σ_Z) position (in feet).

Perspective display format	A display format configured for the formation flight mode which provided the pilot with a three-dimensional representation (in two dimensions) of his formation position with respect to the leader.
PPI display	Plan Position Indicator - A display format configured for the formation flight mode which presented a horizontal view of all aircraft in the formation.
Quickening	A method of providing lead or anticipatory information regarding the system's response. As used in the referenced studies, it consisted of adding higher order derivatives (i.e., rate of change of position) of the system's response (change in aircraft position) to the actual position error. The resultant sum was used to drive one element on the display.
RMS	Root-Mean-Square--$\sqrt{\sum_{i=1}^{N} X_i^2/N}$, where X_i was the measured position in a specific axis during a specific maneuver and N was the total number of position measurements recorded during the maneuver. The RMS measure was also used to represent the levels of pilot and aircraft control activity during a given maneuver, in which case X_i represented respectively the rate of movement of the control stick or the aircraft's attitude.

SAS	Stability Augmentation System.
Single-cue flight director	An electro-mechanical flight display which has only one moving element to represent both pitch and roll attitudes of the aircraft. The element represents both aircraft axes by movement in two different axes on the display.
SK/FF	Stationkeeping/Formation Flight.
Standard deviation	$\sqrt{\sum_{i=1}^{N} (\bar{X} - X_i)^2 / N}$, where $\bar{X}$ is the mean value of the observations (X_i) and N is the total number of observations.
Subsidiary pilot workload	The pilot's workload on tasks other than his primary task (aircraft position control in the referenced studies). See the description of "concomitant task" for more detail on how this subsidiary workload was simulated.
System update time interval	Reciprocal of the rate at which the pilot received new information about his position with respect to the formation leader for display.
α	Coefficient of α-β digital filtering model which determines weight of current raw position measurement versus average over old measurements.
β	Coefficient of α-β digital filtering model which determines weight of current velocity measurement versus average over old measurements.

REFERENCES

1) "Display Requirements Study for Helicopter IFR Formation Flight", Honeywell Document No. 12543-FR1, AD 665-277, JANAIR Report NR 213-054, January 1968.

2) "Experimental Evaluation of a Formation Flight System", Honeywell Document No. 12543-IR5, JANAIR Report 6804002, July 1968.

3) "Evaluation of Conventional Flight Instruments for Use in Helicopter IFR Formation Flight", Honeywell Document No. 12543-FR2, JANAIR Report 680204, February 1968.

4) "Effect of Variations in System Data Rates and Measurement Accuracies on Pilot Performance in the Helicopter IFR Formation Flight Mode", Honeywell Document No. 12543-FR3, JANAIR Report 680408, April, 1969.

5) "Effects of Varying Levels of Autopilot Assistance on Pilot Performance in the IFR Formation Flight Mode", Honeywell Document No. 12543-FR4, Preliminary JANAIR Report, June, 1968.

6) "IFR Formation Flight Display/System Requirements for Advanced Rotary-Wing and Jet Fighter Aircraft", Honeywell Document No. 12543-FR5, Preliminary JANAIR Report, June, 1969.

EXPERIMENTAL DESIGN CONSIDERATIONS IN VALIDATING A METHOD OF MODELING A MAN-ORGANIZED SYSTEM

Bob B. Lukens and Robert A. Brown
University of Alabama
Huntsville, Alabama

ABSTRACT. During the past decade, significant advancements in business organization modeling have been achieved. The man-organized system is a system whose elements consist of people, material, money, and information. Various methods and procedures have been developed to model this business organization. The purpose of the research project is to investigate the theoretical aspects of validating a method of modeling a man-organized system.

A new approach to modeling the organization is a method called Dynamic Organization Network Analysis (DONA). The basis for this method is the use of state variable equations. A question arises concerning the validity of the DONA model. This is answered through an experimental system designed to test the model. The experimental system incorporates a computer simulation with known characteristics. The simulation is used as a standard for comparing the performance characteristics of the DONA model.

The remainder of this article has been reproduced photographically from the authors' manuscript.

EXPERIMENTAL DESIGN CONSIDERATIONS IN VALIDATING A METHOD OF MODELING A MAN-ORGANIZED SYSTEM

Introduction

During the past decade, significant advancements in business organization modeling have been achieved. Various procedures and methods have been developed to model systems including man-organized systems. The man-organized system is characterized by elements such as people, material, money, and information. A method widely used in modeling large organizations is computer simulation. One of the most extensive works on computer simulation of large organizations is by Jay W. Forrester.[3] In his book "Industrial Dynamics" he introduces his concept of modeling organizations and he describes a new computer language, DYNAMO, to implement the modeling technique. Forrester's method is suited to a project which requires modeling of fine details of an organization. However, the method may require years to fully model a large organization. The need for a method of modeling to be accomplished in a timely manner caused a search for other approaches.

Within the last few years, state variables have gained attention as a tool to be used in systems analysis.[1,4,6,7] The state variable equations have found use in mechanical, hydraulic and electrical systems. Herman E. Koenig, et. al. in the book "Analysis of Discrete Physical Systems"[7] proposed that the analysis may be carried into socio-economic systems. The strength of this approach lies in the fact that the systems is considered from a control-system point of view. The basic postulate is that the system is regarded as a conservative system. The matching of inputs to outputs requires for-

mulation according to energy considerations. This approach differs from the economists approach in that the economist views the system as a black box with only inputs and outputs. Other quantities such as internal performance characteristics are not given detailed consideration by the economists. By using the control theory point of view, the analyst is capable of manipulating those quantities which control the performance of the organization.

The Dynamic Organizational Network

Analysis Model

The state variable approach was used in the modeling of a socio-economic system by Koenig.[5,8] Koenig's procedure paralleled Forrester's in that mathematical relationships have to be developed for all significant operations. This is very time consuming for a large organization. The fact that only two forms of equations are needed in the state variable approach caused further considerations of this approach. The state variable equations are as follows:

$$\dot{\Psi}(t) = P\,\Psi(t) + Q\,X(t)$$

$$Y(t) = M\,\Psi(t) + N\,X(t)$$

where P, Q, M, and N are matrices characteristic of the system, X(t) is the vector of inputs, Y(t) is the vector of outputs, and the $\Psi(t)$ are state variables.

Rewriting these equations in matrix multiplication form yields:

$$\begin{bmatrix} \dot{\Psi}(t) \\ Y(t) \end{bmatrix} = \begin{bmatrix} P & Q \\ M & N \end{bmatrix} \cdot \begin{bmatrix} \Psi(t) \\ X(t) \end{bmatrix}$$

This form suggests a form appearing in multiple regression analysis.[2] With these concepts in mind, a modeling method called Dynamic Organizational Network Analysis (DONA) evolved.

The DONA method of modeling is a self-generating procedure in that the matrices characterizing the system are generated from the system through regression analysis. Regression analysis requires data from the system and this data will certainly consist of discrete quantities or measurements taken at some time interval. Since the state variable equations describe continuous functions, a discrete form may be derived by writing the equations in difference-equation form.

$$\begin{bmatrix} \Psi(t+h) \\ Y(t) \end{bmatrix} = \begin{bmatrix} (hP+I) & hQ \\ M & N \end{bmatrix} \cdot \begin{bmatrix} \Psi(t) \\ X(t) \end{bmatrix}$$

Taking first differences to high-pass filter the records, thus eliminating trends, yields

$$\begin{bmatrix} \Delta(\Psi) \\ \nabla(Y) \end{bmatrix} = \begin{bmatrix} (P+I) & Q \\ M & N \end{bmatrix} \cdot \begin{bmatrix} \nabla(\Psi) \\ \nabla(X) \end{bmatrix} \quad \text{for } h = 1$$

where

$$\Delta(\Psi) = \Psi(t+1) - \Psi(t)$$

$$\nabla(\Psi) = \Psi(t) - \Psi(t-1)$$

Letting the matrix $[S] = \begin{bmatrix} (P+I) & Q \\ M & N \end{bmatrix}$ and substituting yields

$$\begin{bmatrix} \Delta(\Psi) \\ \nabla(Y) \end{bmatrix} = [S] \cdot \begin{bmatrix} \nabla(\Psi) \\ \nabla(X) \end{bmatrix}$$

The ∇X, ∇Y, $\nabla \Psi$, and $\Delta \Psi$ are determined from calculations using actual data produced by the real system to be modeled. The time interval may be one hour, one week, etc. By the use of a computer program for multiple regression analysis, the characteristic matrix, S, can be determined.

The Laboratory Concept for Validation

A question arises concerning the validity of the DONA model and the method which produced the model. The usual procedure for validating such a model is by using information from the real system and determining if the model predicts in an acceptable manner when compared with the performance of the real system. It was felt that a better method could be used to validate the DONA model. A laboratory concept was developed to validate the method, which in turn validates the model produced by the method. In this concept the paramaters can be controlled to determine the range and responsiveness of the model. The experimental system shown in Figure 1 is used for the validation procedure. The block marked "SIMCO" is a computer simulation of a sales company. This simulation was developed by C. McMillan and Richard F. Gonzales.[9] As orginally written SIMCO was a distributor operation for a single product. Stochastic demand and lead times were incorporated. SIMCO was modified to handle a second product and to simulate personnel actions; i.e. hires, fires, and transfers. These modifications were made to widen the scope of operations through the addition of the second product and to have some interaction of elements; e.g. the transfer of personnel from one product line to the other. All of the characteristics of SIMCO are known. Since SIMCO is a subroutine, it can easily be replaced to study the validity of DONA as

applied to any other simulated activity.

The DONA methodology block in Figure 1 contains the multiple regression analysis procedure for developing the "S" matrix. In the laboratory system the "S" matrix will characterize the SIMCO Sales Company. The DONA model block in Figure 1 represents the following matrix equation.

$$\begin{bmatrix} \Psi(t+1) \\ Y(t) \end{bmatrix} = \begin{bmatrix} S \end{bmatrix} \cdot \begin{bmatrix} \Psi(t) \\ X(t) \end{bmatrix}$$

The left side of the matrix equation is the DONA output.

The outputs of SIMCO and the DONA model are finally compared as shown in Figure 1. The comparison is made on all parameters desired or deemed appropriate for consideration. For example, the DONA model not only produces the system outputs but also predicts for the next time interval the value of the state variables.

Experimental Design Considerations in the Validation Procedure

The problems revealed in this project have provided some valuable insights into modeling of a man-organized system using this method. The laboratory concept described above has been fully implemented. The entire concept has been written into a computer program and the program has been run and debugged.

Since this is a laboratory, the generation of stimulus data was the first big problem. Even though the computer program could handle a total of 40 input and state variables the problem was not the number of variables. The problems centered around the behavior of the variables.

It was learned that a man-organized system as simulated produces variables some of which cause redundancies in the equations. This of course causes singularities in the matrix. A particular variable of the stimulus data, though time-varying over the long term, may have a constant value for the simulated time period; and when the forward and backward differences are taken, the value of the difference is zero. This same problem can occur with a parameter whose first derivative is a constant. The forward and backward differences will be constant but in the multiple regression analysis, a zero variance is computed. A possible solution to the problems of zero values for the variances is the use of a "dithering signal," much the same as the dithering signal used in control systems. The redundancies in the equations can be overcome by careful selection of the system parameters.

After taking careful note of the above conditions, attention is then directed to the types and levels of stimulus data. The prediction capabilities of the DONA model can be tested through its ability to "track" the real system or, as in the case of the laboratory concept, to "track" the SIMCO simulation. The inputs to SIMCO and DONA may be any one or a combination of signals composed of random noise, sinusoids, impluses or step functions. A particular characteristic to observe is the frequency response of the DONA model. The interactions designed into the SIMCO simulation were part of this test to determine the "worth" of the DONA model and the method to produce the model.

Problems can also arise in the comparison phase of the laboratory concept. The criterion for agreement in the outputs is based on the quadratic form of the covariance matrix. A recent book by Jenkins and Watts[10] describes this procedure when an analysis

is required of a multivariate system. The equation of the quadratic form of the covariance matrix is as follows:

$$\text{Probability}\{[\vec{X}-\vec{\mu}]^T\ V^{-1}[\vec{X}-\vec{\mu}]\} = \text{Probability}\ (\chi^2_m)$$

$$\text{where } V = [\text{cov}_T\{y(j,t), y(k,t)\}]\ ,\ t = 1\ldots,T$$

$$m = \text{degrees of freedom} = \text{order of } V \text{ matrix}$$

This equation is used in two different statistical tests. First, the equation is used to test the prediction capability of the model as a function of time. It can be determined when in the time domain the model ceases to predict with the specified confidence. The equation then may be used to test the significance of each dimension (variable) in the model. Instead of measuring the variability as a function of time the variability is measured as a function of a variable in the model with the time fixed. Again it can be observed when the model ceases to predict with the specified confidence. Using this method for comparison of the outputs quantitative information is generated which gives the performance characteristics of the DONA model and the method used to produce the DONA model.

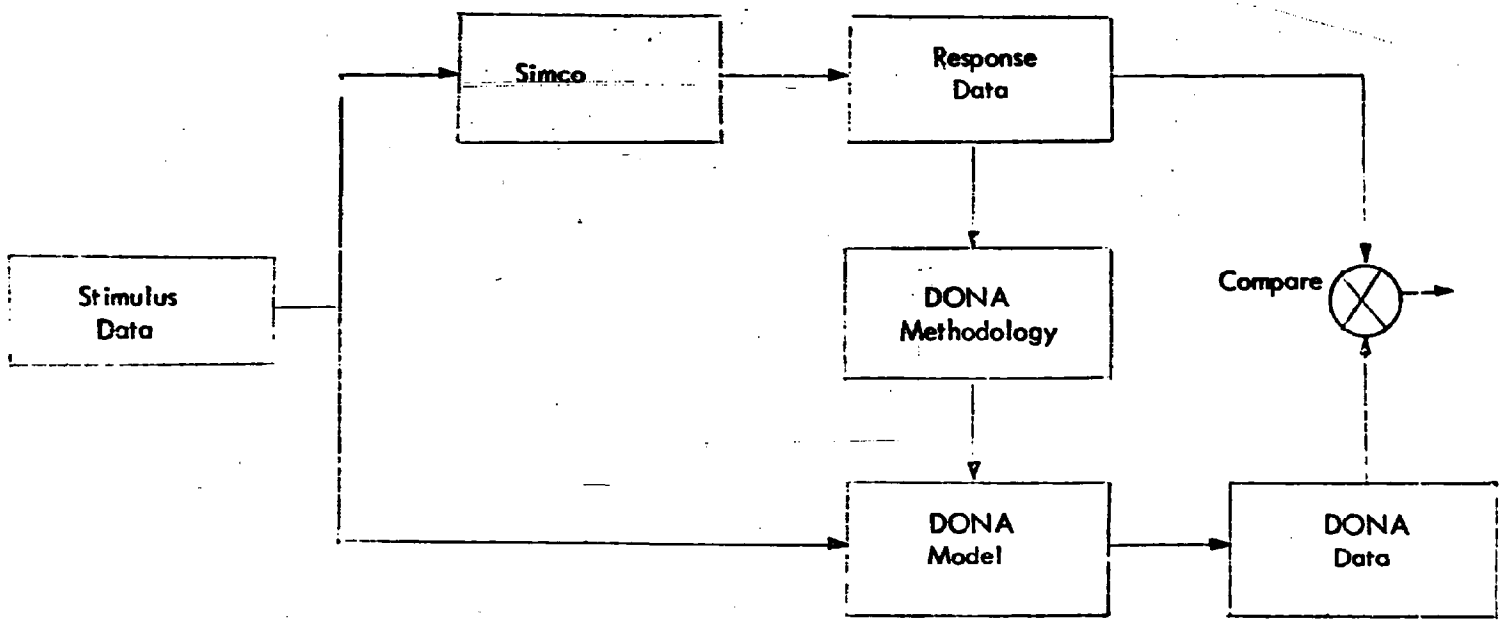


Figure 1 The Laboratory System

BIBLIOGRAPHY

1. Bashkow, Theodore , R., "The A Matrix, New Network Description." IRE Transactions on Circuit Theory, Vol. CT-4, No. 3, (September, 1957), pp. 117-119

2. Duncan, Acheson J., Quality Control and Industrial Statistics. Homewood, Ill., Richard D. Irwin, Inc., 1965.

3. Forrester, Jay W., Industrial Dynamics. New York, the M.I.T. Press and John Wiley, 1961.

4. Frame, J. S. and Koenig, H. E., "Matrix Functions and Applications. Part III-Applications of Matrices to Systems Analysis." IEEE Spectrum, May, 1964, pp. 100-109.

5. Koenig, H. E., "Mathematical Models of Socio-Economic Systems: An Example." IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-1, No. 1 (Nov., 1965), pp. 41-45.

6. Koenig, Herman E., Tokad, Yilmaz, and Kesavan, Analysis of Discrete Physical Systems, New York, McGraw-Hill, 1967.

7. Koenig, H. E., and Tokad, Y., "State Models of Systems of Multi-Terminal Linear Components." IEEE 1964 International Convention Record, Part 1, Vol. 12, pp. 318-329.

8. Koenig, H. E., Kenney, M. G., and Zenach R., A System Model for Management, Planning, and Resource Allocation in Institutions of Higher Education. Final Report to the National Science Foundation, Washington, D. C. Project C-518. September 30, 1968.

9. McMillan, Claude, and Gonzalez, Richard F., Systems Analysis: A Computer Approach to Decision Models. Homewood, Ill., Richard D. Irwin, Inc., 1965.

10. Jenkins, Gwilym M., and Watts, Donald G., Spectral Analysis and Its Applications. San Francisco, Calif., Holden-Day, 1969.

An Investigation of the Effect of Some Prior Distributions on Bayesian Confidence Intervals For Attribute Data

Alan W. Benton
Aberdeen Research and Development Center
Aberdeen Proving Ground, Maryland

ABSTRACT

One method of obtaining confidence intervals on the reliability of a system or component whose sample outcomes are either a success or failure is founded on Bayes theorem. In the Bayesian formulation of the problem one must assume a prior distribution on the random variable of interest, namely, the reliability. The purpose of this investigation was to determine the effect of some prior distributions on Bayesian confidence intervals in which it was assumed that the prior distribution may be represented by the beta distribution. It was my intent to restrict attention to alternative priors one might use when no previous data or experience exists on a system.

INTRODUCTION

Of primary importance to Bayesian statistics is the attachment of probabilities to various possible hypotheses, or in this case probable reliabilities. The mechanism that performs this role is the prior distribution. The plausible values which the random variable might take on do not necessarily imply that one believes the probability exactly, they are only a measure or rough indication of what one tends to believe are the most likely values.

In Bayesian analysis the prior distribution is combined with the test data to yield a modified distribution of reliability, namely, the posterior distribution. That is, the information that we have on a component by way of the prior distribution is updated with the latest test results. If we let

This article has been reproduced photographically from the author's manuscript.

W = Reliability, $0 \leq W \leq 1$

n = Number of tests

r = Number of successes

then by Bayes theorem

$$f(W/n,r) = \frac{g(n,r/W)f(W)}{\int_W g(n,r/W)f(W)dW}$$

where

f(W) = Prior distribution of W

g(n,r/W) = Probability of observing r successes in n trials, given W.

f(W/n,r) = Posterior distribution of W.

In practice it is quite common to assume a beta for the prior distribution. There are several reasons for suggesting the beta, among them are (1) its random variable has range [0, 1], (2) the beta can fit almost any unimodal or nonmodal distribution for a r.v. over the unit interval, and (3) its ease in computations. The beta density is given by

$$f(W) = \frac{(a+b+1)!}{a!\,b!} W^a(1-W)^b, \quad 0 \leq W \leq 1$$
$$= 0 \qquad \text{elsewhere}$$

where $a,b > -1$. It may be shown that by altering a and b the shape of the distribution is changed. See Figure 1. These distributions indicate the plausible values which the random variable might take on.

Assuming f(W) to be beta distributed and g(n,r/W) to be a binomial distribution, the posterior distribution is given by

$$f(W/n,r) = \frac{g(n,r/W)f(W)}{\int_W g(n,r/W)f(W)dW}$$
$$= \frac{\binom{n}{r} W^r (1-W)^{n-r} \frac{(a+b+1)!}{a!b!} W^a(1-W)^b}{\int_0^1 \binom{n}{r} W^r(1-W)^{n-r} \frac{(a+b+1)!}{a!b!} W^a(1-W)^b dW}$$

$$= \frac{(a+b+n+1)!}{(a+r)!(b+n-r)!} W^{a+r} (1-W)^{b+n-r}$$

which is also beta distributed.

CONSTRUCTION OF CONFIDENCE INTERVALS

Since the posterior distribution was found to be beta distributed, it follows that a confidence interval for a beta variate is desired. If we let $W \sim B(a,b)$, then a $100(1-\alpha)\%$ one sided confidence interval is given by

$$\Pr\left[W_L^1 < W < 1\right] = \int_{W_L}^{1} f(W)dW = 1-\alpha$$

An exact lower bound is given by

$$\frac{1}{1 + \frac{b+n-r+1}{a+r+1} F_{1-\alpha}}$$

where

$F_{1-\alpha}$ = $(1-\alpha)$ percentile of the F- distribution with $2(b+n-r+1)$ and $2(a+r+1)$ degrees of freedom.

a,b = parameters from the prior

n = sample size

r = number of successes

The use of the F- distribution may be seen from the following. From the change of variable theorem of integral calculus we may write

$$g(F) = h(W = U(F)) \frac{dW}{dF}.$$

Let

$$W = \frac{\frac{a+r+1}{b+n-r+1} F}{1 + \frac{a+r+1}{b+n-r+1} F}$$

then

$$\frac{dW}{dF} = \frac{\frac{a+r+1}{b+n-r+1}}{\left(1 + \frac{a+r+1}{b+n-r+1} F\right)^2}$$

and on substitution

$$g(F) = \frac{\Gamma(a+b+n+2)}{\Gamma(a+r+1)\,\Gamma(b+n-r+1)} \left(\frac{a+r+1}{b+n-r+1}\right)^{a+r+1} \frac{F^{a+r}}{\left(1+\frac{a+r+1}{b+n-r+1}F\right)^{a+b+n+2}}$$

which is the F- distribution with 2(a+r+1) and 2(b+n-r+1) degrees of freedom. Now recall

$$\int_{W_L^1}^{1} f(W)\,dW = \int_{\frac{RF}{1+RF}} f(W)\,dW = 1-\alpha$$

where

$$R = \frac{a+r+1}{b+n-r+1}$$

Noting that $F_b^a(\alpha) = 1/F_a^b(1-\alpha)$, the lower limit is given by

$$W_L^1 = \frac{1}{1+\frac{b+n-r+1}{a+r+1}F_{1-\alpha}}$$

One of the reasons for using the F- distribution is that tables of this distribution are frequently more readily available than those of the incomplete beta. Also, they are convenient for those values of α most often used, e.g., 0.10, 0.05, 0.01.

For sufficiently large sample sizes the normal approximation may be employed, the approximation being best in the vicinity of W equal to one-half. For the posterior distribution the expected value and variance of W are:

$$E(W) = \frac{a+r+1}{b+n-r+1}$$

and

$$\sigma^2 = \frac{(a+r+1)(b+n-r+1)}{(a+b+n+2)^2(a+b+n+3)}$$

Thus, the $100(1-\alpha)\%$ lower limit is given by $E(W) - z_{1-\alpha}\sigma$

where $z_{1-\alpha}$ is such that

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_{1-\alpha}} e^{-t^2/2}\, dt = 1-\alpha.$$

DISCUSSTION

As indicated in the previous section, the lower confidence limit (LCL) was computed for illustrative purposes. A confidence level of 95% was chosen and the following prior distributions were selected:

a	-1 -1/2 0 1 6
b	-1 -1/2 0 0 1

References [1] and [4] give discussion on the reasons for assuming prior distributions when a and b are both set equal to -1, -1/2, and 0. By letting a=6 and b=1 we give more credence to moderately high reliabilities, which would appear to be a likely area of interest. For large a and/or b one gets into the problem of the prior distribution far outweighing the most recent evidence or sample.

In Figure 2 the lower 95% limit is plotted against the number of failures observed in a sample of size twenty. In general, the results are quite similar for the other sample sizes investigated (n=10(5)25). It may be noted that the number of failures begins at one. This was done since when using the (+1, -1) prior one must assume the occurrence of at least one success and at least one failure in the sample.

An examination of the lower limits plotted on Figure 2 yields the following general results:

1. The (-1, -1) prior results in shorter confidence intervals than a (-1/2, -1/2) prior which in turn gives a shorter confidence interval than the (0,0) prior. This holds true for moderate to high reliabilities. The

order is reversed for low reliabilities.

2. The differences in confidence interval length become smaller as n, the sample size, increases.

3. All three priors [(-1,-1), (-1/2, -1/2), (0,0)] result in shorter confidence intervals than those obtained with classical methods.

4. The (6,1) prior does not result in a uniformly shorter confidence interval for moderate to high reliabilities.

Now, let r/n (# success/sample) be an indicator of reliability and rank the priors according to their interval length. Thus the shortest confidence interval indicates the most optimistic result and the longest the most pessiminstic result for sample estimates of reliability. The following table presents a summary of these rankings.

Summary of Length of Confidence Intervals for
(-1, -1), (-1/2, -1/2), (0, 0), (1, 0) Priors

r/n	Shortest	——	——	Longest
$\geq$ 0.85	(-1, -1)	(-1/2, -1/2)	(1,0)	(0,0)
0.75-0.85	(-1, -1)	(1, 0)	(1/2, - 1/2)	(0, 0)
0.65-0.75	(1, 0)	(-1, -1)	(-1/2, -1/2)	(0, 0)

Thus, for r/n greater than 0.75 the (-1, -1) prior gives the shortest interval while the uniform prior (o, 0) gives the longest interval.

References

1. Lindley, D. V., Introduction to Probability and Statistics from a Bayesian Viewpoint, Part 2, Cambridge University Press, New York, 1965.

2. Mood, A. M., Graybill, F. A. Introduction to the Theory of Statistics, McGraw-Hill Book Co., Inc., New York, 1963.

3. Savage, L. J., The Foundations of Statistics, John Wiley and Sons, Inc., New York, 1954.

4. "Planning Guide for Demonstration and Assessment of Reliability and Durability", AMC Pamphlet AMCP 702-5.

Figure I

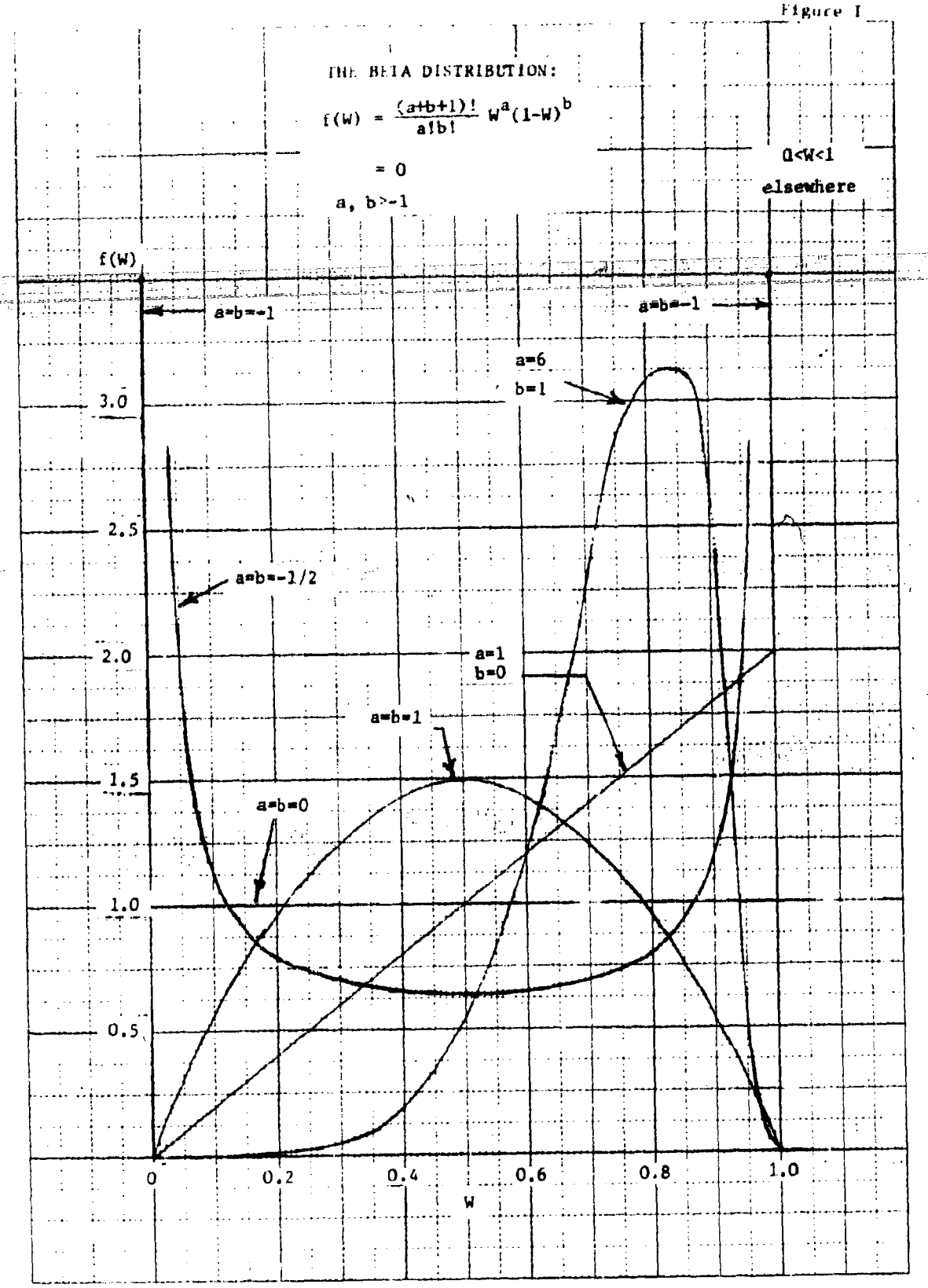
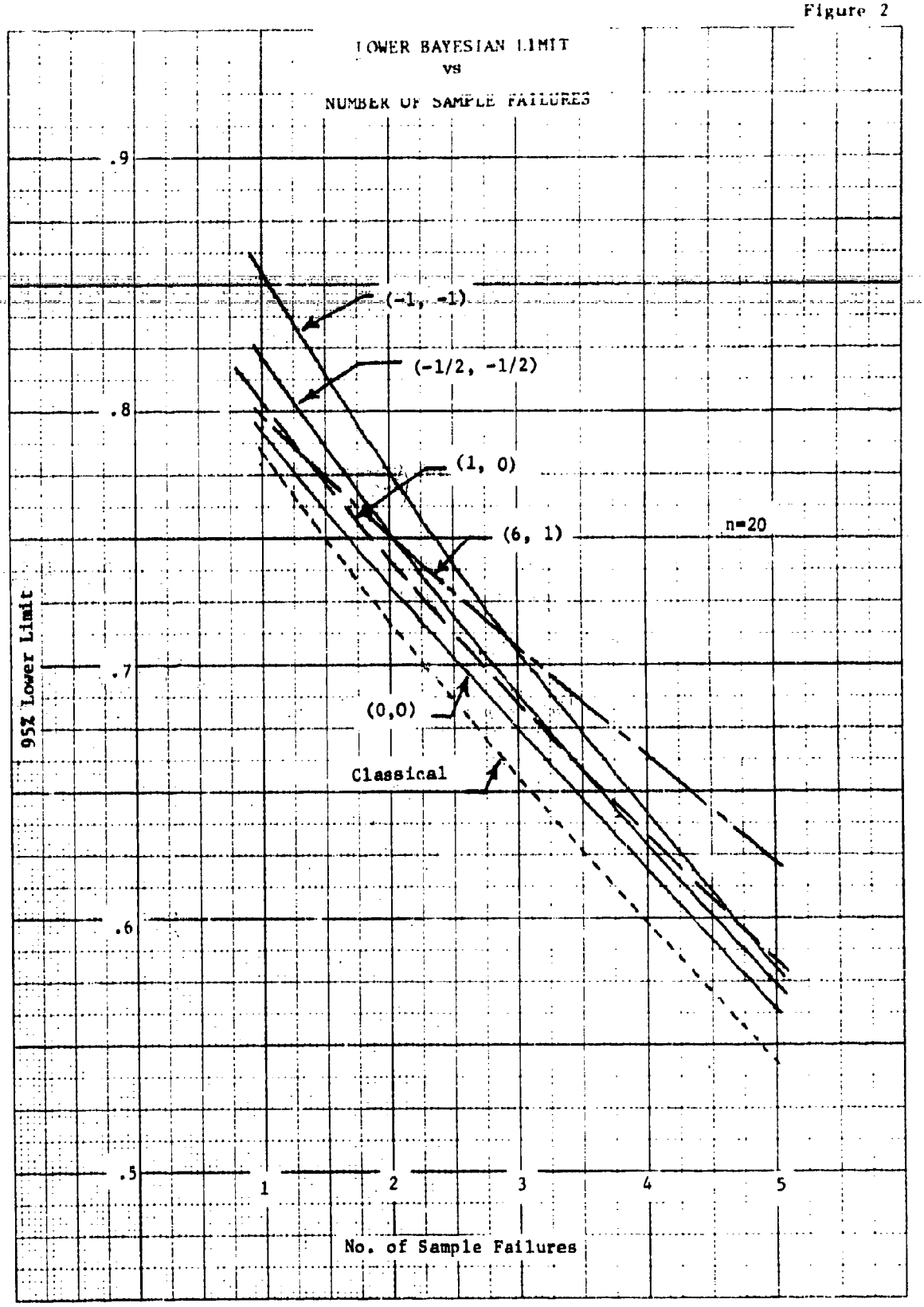

THE BETA DISTRIBUTION:
$f(W) = \frac{(a+b+1)!}{a!b!} W^a(1-W)^b$
= 0
a, b>-1
0<W<1
elsewhere
f(W)
a=b=-1
a=b=-1
a=6
b=1
3.0
2.5
a=b=-1/2
2.0
a=1
b=0
a=b=1
1.5
a=b=0
1.0
0.5
0
0.2
0.4
0.6
0.8
1.0
W

Figure 2


LOWER BAYESIAN LIMIT
vs
NUMBER OF SAMPLE FAILURES
.9
.8
.7
.6
.5
95% Lower Limit
(-1, -1)
(-1/2, -1/2)
(1, 0)
(6, 1)
n=20
(0,0)
Classical
1
2
3
4
5
No. of Sample Failures

SOME TECHNIQUES FOR CONSTRUCTING MUTUALLY ORTHOGONAL LATIN SQUARES*

W. T. Federer, A. Hedayat, E. T. Parker
B. L. Raktoe, Esther Seiden, and R. J. Turyn
Mathematics Research Center, U. S. Army
Madison, Wisconsin

ABSTRACT. Various methods of constructing a set of mutually orthogonal latin squares are presented and the theoretical aspects of various methods are discussed. Illustrative examples of constructing latin squares and sets of mutually orthogonal latin squares are given. The methods of constructing latin squares and sets of orthogonal latin squares are complete and partial confounding, fractional replication, analysis of variance, group, projecting diagonals, orthomorphism, pairwise balanced design, oval, code, product composition, and sum composition. The methods of construction designated as partial confounding, fractional replication, analysis of variance, and sum composition appear not to have been discussed previously in the literature. The methods of complete confounding and of projecting diagonals have been given only a passing reference with no indication as to the actual construction procedure. The sum composition method has interesting consequences in combinatorial theory as well as in the construction of orthogonal latin squares. Lastly, equivalences of fourteen combinatorial systems to orthogonality in latin squares has been investigated and described.

*This article has appeared as Technical Summary Report No. 1030 of The Mathematics Research Center, U. S. Army, The University of Wisconsin, Madison, sponsored under Contract No.: DA-31-ARO-D-462.
The remainder of this article has been reproduced photographically from the authors' manuscript.

Preceding page blank

TABLE OF CONTENTS

Section and Title

I. Introduction and Some Terminology . . . . W. T. Federer and A. Hedayat

II. Factorial Confounding Construction of O(n,t) Sets

II.1 Complete Confounding . . . . . W. T. Federer and B. L. Raktoe

II.2 Partial Confounding . . . . . W. T. Federer

III. Fractional Replication Construction of O(n,t) Sets . . . W. T. Federer

IV. ANOVA Construction of O(n,t) Sets . . . . . W. T. Federer

V. Group Construction of O(n,t) Sets . . . . . A. Hedayat

VI. Projecting Diagonals Construction of O(n,t) Sets . . . W. T. Federer

VII. Relations Between Complete Confounding and Simple Orthomorphisms . . . . . B. L. Raktoe

VIII. Some Remarks on "Orthomorphism" Construction of O(n,t) Sets . . . . . E. T. Parker

IX. Oval Construction of O(n,t) Sets . . . . . Esther Seiden

X. Code Construction of O(n,t) Sets . . . R. J. Turyn and W. T. Federer

XI. Pairwise Balanced Design Construction of O(n,t) Sets E. T. Parker

XII. Product Composition of O(n,t) Sets . . . . . A. Hedayat

XIII. Sum Composition of O(n,t) Sets . . . . . A. Hedayat and Esther Seiden

XIV. Computer Construction of O(10,t) Sets . . . . . E. T. Parker

XV. Equivalences of O(n,t) Sets with other Combinatorial Systems . . . . . A. Hedayat

XVI. Acknowledgements

XVII. Literature Cited

SOME TECHNIQUES FOR CONSTRUCTING MUTUALLY ORTHOGONAL LATIN SQUARES

W. T. Federer[1], A. Hedayat[2], E. T. Parker[3]
B. L. Raktoe[4], Esther Seiden[5], and R. J. Turyn[6]

1. Introduction and Some Terminology

The purpose of this paper is to present a set of methods for constructing mutually orthogonal latin squares and to exhibit some squares produced by each of the methods. The set of methods presented herein was discussed in a series of informal seminars held during the weeks of July 14-18 and 21-25, 1969, by the authors at Cornell University. The motivation for these discussions was derived from results obtained by Hedayat [1969] and from the optimism of the authors. New procedures for constructing a set of mutually orthogonal latin squares and new views of present methods of construction were desired in order to advance the theory of mutual orthogonality in latin squares.

---

[1] Professor of Biological Statistics, Cornell University and Visiting Professor, Mathematics Research Center, U.S. Army, and University of Wisconsin (on sabbatical leave 1969-70).

[2] Assistant Professor, Cornell University.

[3] Professor Mathematics, University of Illinois, and Visiting Professor, Cornell University (July, 1969).

[4] Associate Professor, University of Guelph and Visiting Associate Professor, Cornell University (January to August, 1969).

[5] Professor, Michigan State University, and Visiting Professor, Cornell University (June, July, August, 1969).

[6] Mathematician, Raytheon Corporation, and Visiting Professor, Cornell University (July, 1969).

As may be noted from the table of contents, the different sections were written by different authors. An attempt was made to have a consistent notation and a uniform style. Although much more work is required to finalize the method in several of the sections enough is known about the method to use it to construct a latin square of any order or to construct a set of two or more mutually orthogonal latin squares. Also, a number of equivalences may be noted for some of the methods.

The theory of mutual orthogonality in latin squares has application in the construction of many classes of experiment designs and in many combinatorial systems. The latter subject is discussed in section XV where the equivalences of various combinatorial systems are presented. With regard to the former subject, there is an ever present need for new experiment designs for new experimental situations in order for the experimenter not to have to conduct his experiment to fit known experiment designs.

Some of the notation and terminology that will be utilized is presented below.

Definition 1.1. A latin square of order n on a set $\underline{\Sigma}$ with n distinct elements is an $n \times n$ matrix each of whose rows and columns is a permutation of the set $\underline{\Sigma}$.

Example:

1	2	3
2	3	1
3	1	2

is a latin square of order 3 on $\underline{\Sigma} = \{1, 2, 3\}$.

Definition 1.2. Two latin squares $L_1 = (a_{ij})$ and $L_2 = (b_{ij})$ of order n are said to be orthogonal if the n^2 ordered pairs (a_{ij}, b_{ij}) $(i,j) = 1,2,\ldots,n)$ are all distinct. Note that L_1 and L_2 need not be defined on the same set.

Example:

1	2	3
2	3	1
3	1	2

and

A	B	C
C	A	B
B	C	A

.

Definition 1.3. The members of a set of t latin squares $L_1, L_2, \ldots, L_t$ of order n are said to be mutually (pairwise) orthogonal if L_i is orthogonal to L_j, $i \neq j$, $i,j = 1,2,\ldots,t$. Hereafter by an (n,t) set we mean a set consisting of t mutually orthogonal latin squares of order n.

Example:

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

,

1	2	3	4
4	3	2	1
2	1	4	3
3	4	1	2

, and

1	2	3	4
3	4	1	2
4	3	2	1
2	1	4	3

.

Latin squares and orthogonal latin squares have at least 187 years of history. Hedayat [1969-section IX] has presented a reasonably good picture of this history which will not be repeated here. It is planned to prepare a historical account of developments related to orthogonality in latin squares and to publish this material together with a bibliography elsewhere.

II. Factorial Confounding Construction of $Q(n,t)$ Sets

II.1. Complete Confounding

The n^2 row-column intersections may be related to the treatment combinations in a n^2 factorial treatment design. To illustrate let us consider the 4^2 factorial and the latin square of order 4. The levels of the main effects, A and B, in the factorial will be used to designate the rows and the columns of the latin square of order 4 as follows:

latin squares of order 4

	column 1 = $(B)_0$	column 2 = $(B)_1$	column 3 = $(B)_2$	column 4 = $(B)_3$
row 1 = $(A)_0$	00	01	02	03
row 2 = $(A)_1$	10	11	12	13
row 3 = $(A)_2$	20	21	22	23
row 4 = $(A)_3$	30	31	32	33

Thus, four combinations out of the 16 which have $i = 0$ in the subscript ij, i.e., 00, 01, 02, and 03, are designated as $(A)_0$ and are put in row 1. Continuing this procedure the remainder of the ij combinations are allocated to the remaining rows and to the columns as shown above.

Now, three other effects with 4 levels each can be set up from the projective geometry $PG(1,2^2)$; the effects are $(AB^{u_1})_{u_1 \cdot u_1 u_j}$, $(AB^{u_2})_{u_1 \cdot u_2 u_j}$, and $(AB^{u_3})_{u_1 \cdot u_3 u_j}$. The levels of these effects and the corresponding latin square produced by letting all combinations of the level of an effect be a symbol in the latin square are (see page 337 of Kempthorne [1952], e.g.):

$$(AB^{u_1})_{u_i+u_1u_j} = \begin{cases} 0 & 00 + 11 + 22 + 33 \cdot \text{I} \\ 1 & 01 + 10 + 23 + 32 \cdot \text{II} \\ 2 & 02 + 13 + 20 + 31 \cdot \text{III} \\ 3 & 03 + 12 + 21 + 30 \cdot \text{IV} \end{cases}$$

00 = I	01 = II	02 = III	03 = IV
10 = II	11 = I	12 = IV	13 = III
20 = III	21 = IV	22 = I	23 = II
30 = IV	31 = III	32 = II	33 = I

$$(AB^{u_2})_{u_i+u_2u_j} = \begin{cases} 0 & 00 + 13 + 21 + 32 \rightarrow \alpha \\ 1 & 03 + 10 + 22 + 31 \rightarrow \beta \\ 2 & 01 + 12 + 20 + 33 \rightarrow \gamma \\ 3 & 02 + 11 + 23 + 30 \cdot \delta \end{cases}$$

α	γ	δ	β
β	δ	γ	α
γ	α	β	δ
δ	β	α	γ

$$(AB^{u_3})_{u_i+u_3u_j} = \begin{cases} 0 & 00 + 12 + 23 + 31 \cdot \text{W} \\ 1 & 02 + 10 + 21 + 33 \rightarrow \text{X} \\ 2 & 03 + 11 + 20 + 32 \cdot \text{Y} \\ 3 & 01 + 13 + 22 + 30 \cdot \text{Z} \end{cases}$$

W	Z	X	Y
X	Y	W	Z
Y	X	Z	W
Z	W	Y	X

In the above the complete confounding scheme of sources of variation in the O(4, 3) set and the effects in the factorial may be illustrated in the following analysis of variance table wherein the sums of squares in the lines of the analysis of variance are orthogonal to each other:

Source of variation	Degrees of freedom
Correction for mean	1
Rows = A effect	3
Columns = B effect	3
Roman numbers = (AB^{u_1}) effect	3
Greek letters = (AB^{u_2}) effect	3
Latin letters = (AB^{u_3}) effect	3
Total	16

Instead of relating the mutually orthogonal latin squares of order 4 to a 4^2 factorial we may relate them to a 2^4 factorial in the following manner. Let the 16 row-column intersections be numbered as follows:

	column			
row	1	2	3	4
1	0000	0001	0010	0011
2	0100	0101	0110	0111
3	1000	1001	1010	1011
4	1100	1101	1110	1111

Let the factors be a, b, c, and d with two levels (0 and 1) each. The rows correspond to factorial effects A, B, and AB and the columns correspond to factorial effects C, D, and CD. (This form of constructing latin squares has been used by Fisher and Yates [1957] for latin squares of order 8.) Then, let the symbols in the 3 latin squares be represented by the following scheme:

factorial generators	combinations	latin squares
$(AC)_0$, $(BD)_0$, $(ABCD)_0$	0000 + 0101 + 1010 + 1111 = I	I II III IV
$(AC)_0$, $(BD)_1$, $(ABCD)_1$	0001 + 0100 + 1011 + 1110 = II	II I IV III
$(AC)_1$, $(BD)_0$, $(ABCD)_1$	0010 + 0111 + 1000 + 1101 = III	III IV I II
$(AC)_1$, $(BD)_1$, $(ABCD)_0$	0011 + 0110 + 1001 + 1100 = IV	IV III II I
$(AD)_0$, $(ABC)_0$, $(BCD)_0$	0000 + 0110 + 1011 + 1101 = W	W Z X Y
$(AD)_0$, $(ABC)_1$, $(BCD)_1$	0010 + 0100 + 1001 + 1111 = X	X Y W Z
$(AD)_1$, $(ABC)_0$, $(BCD)_1$	0001 + 0111 + 1010 + 1100 = Z	Y X Z W
$(AD)_1$, $(ABC)_1$, $(BCD)_0$	0101 + 0011 + 1000 + 1110 = Y	Z W Y X

$(ACD)_0$, $(BC)_0$, $(ABD)_0$	0000 + 0111 + 1110 + 1001 = α	
$(ACD)_0$, $(BC)_1$, $(ABD)_1$	1010 + 0100 + 0011 + 1101 = β	
$(ACD)_1$, $(BC)_0$, $(ABD)_1$	1000 + 0110 + 1111 + 0001 = γ	
$(ACD)_1$, $(BC)_1$, $(ABD)_0$	0010 + 0101 + 1011 + 1100 = δ	

α	γ	δ	β
β	δ	γ	α
γ	α	β	δ
δ	β	α	γ

The correspondence of the latin squares obtained from complete confounding considering a 4^2 factorial and considering a 2^4 factorial is demonstrated in the following analysis of variance table:

Source of variation		degrees of freedom
Correction for mean		1
Rows	= A effect in 4^2 factorial	3
	A effect in 2^4 factorial	1
	B " " 2^4 "	1
	AB " " 2^4 "	1
Columns	= B effect in 4^2 factorial	3
	C effect in 2^4 factorial	1
	D " " 2^4 "	1
	CD " " 2^4 "	1
Roman numbers	= AB^{u_1} effect in 4^2 factorial	3
	AC effect in 2^4 factorial	1
	BD " " 2^4 "	1
	ABCD " " 2^4 "	1
Greek letters	= AB^{u_2} effect in 4^2 factorial	3
	ACD effect in 2^4 factorial	1
	BC " " 2^4 "	1
	ABD " " 2^4 "	1

Latin letters = AB^u,	effect in 4^2 factorial	3
	AD effect in 2^4 factorial	1
	ABC " " 2^4 "	1
	BCD " " 2^4 "	1
Total		16

It should be noted here that the effects in the 2^4 map directly into the 4^2 projective geometry or $PG(1, 2^2)$. Likewise, even though one more set of generators is available, viz.

	Generators	interaction
Roman numbers	= AD, BC	ABCD
Greek letters	= AC, ABD	BCD
Latin letters	= BD, ABC	ACD

the three orthogonal latin squares produced are the same ones. Since the third effect above is obtained as the product of the two generators, mod 2, we need consider only the generators. Multiplying these by CD (mod 2) we obtain the generators of the preceding scheme. Hence, even though two different complete confounding schemes are available there is a simple one-to-one mapping of one set into the other set. Although nothing interesting turns up here, it would be interesting to study the various complete confounding schemes in the latin square of order 9 as related to the 3^4 factorial.

As a second illustration of the use of complete confounding to construct latin squares, let us consider the latin square of order 6. Using the notation

and concepts of Raktoe [1969] we designate the 6^2 as a $2^2(3)^2$ factorial and represent a combination by ghij where g, h are members of I (3) and i, j are members of I (4). The effects in the 2^2 and in the 3^2 factorials are denoted respectively by:

A^3 C^4

B^3 D^4

A^3B^3 C^4D^4

C^4D^2

The remaining interactions are given below in the analysis of variance table:

Source of variation	Degrees of freedom
Correction for mean	1
Rows = A^3C^4	5
A^3	1
C^4	2
$A^3 \times C^4$	2
Columns = B^3D^4	5
B^3	1
D^4	2
$B^3 \times D^4$	2
Treatments or symbols = $A^3B^3C^4D^4$	5
A^3B^3	1
C^4D^4	2
$A^3B^3 \times C^4D^4$	2

Remainder	20	
C^4D^2		2
$A^3 \times D^4$		2
$A^3 \times C^4D^4$		2
$A^3 \times C^4D^2$		2
$B^3 \times C^4$		2
$B^3 \times C^4D^4$		2
$B^3 \times C^4D^2$		2
$A^3B^3 \times C^4$		2
$A^3B^3 \times D^4$		2
$A^3B^3 \times C^4D^2$		2
Total	36	

Let us now set up the 6 rows and the 6 columns of a latin square of order 6 with the corresponding designation of the 36 combinations as follows:

Columns

Rows	$(B^3D^4)_0$	$(B^3D^4)_1$	$(B^3D^4)_2$	$(B^3D^4)_3$	$(B^3D^4)_4$	$(B^3D^4)_5$
$(A^3C^4)_0$	0000	0304	0002	0300	0004	0302
$(A^3C^4)_1$	3040	3344	3042	3340	3044	3342
$(A^3C^4)_2$	0020	0324	0022	0320	0024	0322
$(A^3C^4)_3$	3000	3304	3002	3300	3004	3302
$(A^3C^4)_4$	0040	0344	0042	0340	0044	0342
$(A^3C^4)_5$	3020	3324	3022	3320	3024	3322

Now let the levels of $A^3B^3C^4D^4$ correspond to the symbols in a latin square of order 6 as follows:

Levels	Combination for which $3g+3h+4i+4j$, mod 6, is constant	Symbol
$(A^3B^3C^4D^4)_0$	0000 + 3342 + 0024 + 3300 + 0042 + 3324	0
$(A^3B^3C^4D^4)_1$	0304 + 3040 + 0322 + 3004 + 0340 + 3022	1
$(A^3B^3C^4D^4)_2$	0002 + 3344 + 0020 + 3302 + 0044 + 3320	2
$(A^3B^3C^4D^4)_3$	0300 + 3042 + 0324 + 3000 + 0342 + 3024	3
$(A^3B^3C^4D^4)_4$	0004 + 3340 + 0022 + 3304 + 0040 + 3322	4
$(A^3B^3C^4D^4)_5$	0302 + 3044 + 0320 + 3002 + 0344 + 3020	5

This produces the following latin square of order 6:

0	1	2	3	4	5
1	2	3	4	5	0
2	3	4	5	0	1
3	4	5	0	1	2
4	5	0	1	2	3
5	0	1	2	3	4

Alternatively we could have used levels of $A^3B^3C^4D^2$ to construct the following latin square of order 6:

Levels	Combinations for which $3g+3h+4i+2j$, mod 6, is constant	Symbol
$(A^3B^3C^4D^2)_0$	0000 + 3344 + 0022 + 3300 + 0044 + 3322	0
$(A^3B^3C^4D^2)_1$	0302 + 3040 + 0324 + 3002 + 0340 + 3024	1
$(A^3B^3C^4D^2)_2$	0004 + 3342 + 0020 + 3304 + 0042 + 3320	2
$(A^3B^3C^4D^2)_3$	0300 + 3044 + 0322 + 3000 + 0344 + 3022	3
$(A^3B^3C^4D^2)_4$	0002 + 3340 + 0024 + 3302 + 0040 + 3324	4
$(A^3B^3C^4D^2)_5$	0304 + 3042 + 0320 + 3004 + 0342 + 3020	5

latin square of order 6

0	5	4	3	2	1
1	0	5	4	3	2
2	1	0	5	4	3
3	2	1	0	5	4
4	3	2	1	0	5
5	4	3	2	1	0

Thus, the above square is simply a column permutation of the previous one. As there are no other sets of 5 degrees of freedom leading to a latin square of order 6 (i.e. A^3, B^3, and A^3B^3 exhaust the three single degrees of freedom from the 2^2 factorial and C^4, D^4, C^4D^4, and C^4D^2 exhaust all sets of 2 degrees of freedom from the 3^2 factorial), it is not possible to obtain a latin square of order 6 orthogonal to either of the preceding ones using complete confounding schemes.

For a latin square of order 10 we may use levels of $A^5B^5C^6D^6$, $A^5B^5C^6D^2$, $A^5B^5C^6D^8$, or $A^5B^5C^6D^4$ to form four different latin squares of order 10.

II. 2. Partial Confounding

In the last section use was made of complete confounding of effects in a factorial with the rows, columns, and symbols in a latin square. Instead of completely confounding an effect, it could be partially confounded. For example, the latin square of order 4 could be considered as a 2^4 factorial as in the preceding section, with the following scheme of confounding:

Rows		1 = $(C)_0$	2 = $(C)_1$	3 = $(D)_0$	4 = $(D)_1$
1	$(A)_0, (B)_0$	0000	0011	0010	0001
2	$(A)_0, (B)_1$	0101	0110	0100	0111
3	$(A)_1, (B)_0$	1000	1011	1010	1001
4	$(A)_1, (B)_1$	1101	1110	1100	1111

If we set up the latin square symbols for the above as

α	β	γ	δ
β	α	δ	γ
γ	δ	β	α
δ	γ	α	β

then

the symbols correspond to the following combinations:

α: 0000 + 0110 + 1001 + 1100 = $(ABCD)_0$ + other effects

β: 0011 + 0101 + 1010 + 1111 = $(ABCD)_0$ + " "

γ: 1000 + 1110 + 0010 + 0111 = $(ABCD)_1$ + " "

δ: 0001 + 0100 + 1011 + 1101 = $(ABCD)_1$ + " "

It is known that this latin square has no orthogonal mate (Hedayat [1969]). This means that no orthogonal partition of the remaining sum of squares can be made which forms a latin square.

If on the other hand, the latin square used is

α	β	γ	δ
β	α	δ	γ
γ	δ	α	β
δ	γ	β	α

, the combinations

corresponding to the Greek letters are:

α: $0000 + 0110 + 1010 + 1111 = (ABCD)_0$ + other effects

β: $0011 + 0101 + 1001 + 1100 = (ABCD)_0$ + " "

γ: $0010 + 0111 + 1000 + 1110 = (ABCD)_1 + (AC)_1$ + other effects

δ: $0001 + 0100 + 1011 + 1101 = (ABCD)_1$ + other effects

This square has two mutually orthogonal mates and hence there must be partitions of the sums of squares into orthogonal components which correspond to the symbols of a latin square.

Instead of inserting symbols in the latin square of order 4, denote the symbols in the latin square by the following partial confounding scheme:

(i) add the two 1/8 replicates generated by $((A)_0, (D)_0, (BC)_0)$ and $((A)_1, (C)_1, (ABD)_1)$ to obtain the 4 combinations $(0000 + 0110) + (1010 + 1111)$ and denote these 4 combinations as symbol α,

(ii) add the two 1/8 replicates generated by $((D)_1, (AB)_1, (AC)_0)$ and $((AB)_0, (C)_0, (AD)_1)$ to obtain combinations $(0101 + 1011) + (1100 + 0001)$ and denote these 4 combinations as symbol β,

(iii) add the two 1/8 replicates generated by $((A)_1, (D)_0, (ABC)_1)$ and $((A)_0, (C)_1, (BD)_0)$ to obtain combinations $(1000 + 1110) + (0010 + 0111)$ and denote these 4 as symbol γ,

(iv) add the two 1/8 replicates generated by $((AB)_0, (AC)_1, (D)_1)$ and $((AB)_1, (C)_0, (BD)_1)$ to obtain the combinations $(1101 + 0011) + (0100 + 1001)$ and denote these 4 as symbol δ.

This procedure results in the following latin square of order 4:

α	δ	γ	β
β	α	δ	γ
γ	β	α	δ
δ	γ	β	α

Obviously, one could take any pair of 1/8 replicates such that the 4 combinations are in different rows and in different columns to form the combinations for a given symbol.

The above type of partial confounding results in the class of latin squares denoted as half-plaid latin squares. If partial confounding were utilized in rows as well as in columns the resulting square would be denoted as a plaid latin square (so-called because of its resemblence to plaid cloth if the effects confounded were of different colors). The three types of squares are illustrated below for a latin square of order 6:

Complete confounding of effects

	Columns					
Rows	1 = $(A)_0,(C)_0$	2 = $(A)_0,(C)_1$	3 = $(A)_0,(C)_2$	4 = $(A)_1,(C)_0$	5 = $(A)_1,(C)_1$	6 = $(A)_2,(C)_2$
1 = $(B)_0,(D)_0$	0000	0010	0020	1000	1010	1020
2 = $(B)_0,(D)_1$	0001	0011	0021	1001	1011	1021
3 = $(B)_0,(D)_2$	0002	0012	0022	1002	1012	1022
4 = $(B)_1,(D)_0$	0100	0110	0120	1100	1110	1120
5 = $(B)_1,(D)_1$	0101	0111	0121	1101	1111	1121
6 = $(B)_1,(D)_2$	0102	0112	0122	1102	1112	1122

Partial confounding of effects with columns

Rows	Columns 1 = $(C)_0$	2 = $(C)_1$	3 = $(C)_2$	4 = $(CD)_0$	5 = $(CD)_1$	6 = $(CD)_2$
1 = $(B)_0,(D)_0$	0000	0010	0020	1000	1010	1020
2 = $(B)_0,(D)_1$	0001	0011	0021	1021	1001	1011
3 = $(B)_0,(D)_2$	0002	0012	0022	1012	1022	1002
4 = $(B)_1,(D)_0$	1100	1110	1120	0100	0110	0120
5 = $(B)_1,(D)_1$	1101	1111	1121	0121	0101	0111
6 = $(B)_1,(D)_2$	1102	1112	1122	0112	0122	0102

Partial confounding in both rows and columns

Rows	Columns 1 = $(C)_0$	2 = $(C)_1$	3 = $(C)_2$	4 = $(CD)_0$	5 = $(CD)_1$	6 = $(CD)_2$
1 = $(D)_0$	00	10	20	00	10	20
2 = $(D)_1$	01	11	21	21	01	11
3 = $(D)_2$	02	12	22	12	22	02
4 = $(CD^2)_0$	00	11	22	00	22	11
5 = $(CD^2)_1$	02	10	21	21	10	02
6 = $(CD^2)_2$	01	12	20	12	01	20

In the last table above only the subscripts for combinations of factors c and d have been inserted. There is some difficulty in inserting subscripts for factors a and b such that these effects are orthogonal to both rows and columns. In any event, this problem requires further study to determine if half-plaid latin squares and plaid latin squares lead to latin squares not of the same type as given by complete confounding. If the three types of latin squares of order 6 can be produced by partial and complete confounding, this would be an interesting result.

III. Fractional Replication Construction of O(n,t) Sets

Any latin square may be considered as an n^{-1} fraction of an n^3 factorial where the rows represent levels of one factor, the columns represent the levels of the second factor, and the symbols in the latin square represent the levels of the third factor. As an illustration, consider the latin square of order 3 where the 9 combinations represent the 1/3 fraction of a 3^3 factorial as follows:

Columns

Rows	0	1	2
0	000	012	021
1	102	111	120
2	201	210	222

The above is the 1/3 fraction of a 3^3 corresponding to $(ABC)_{h+i+j=0,\ \mathrm{mod}\ 3}$. Since this is a regular fraction we may write out the aliasing structure in this fraction as follows:

$$M + ABC$$
$$A + AB^2C^2 + BC$$
$$B + AB^2C + AC$$
$$C + ABC^2 + AB$$
$$AB^2 + AC^2 + BC^2$$

where the effects connected with a plus sign are completely confounded with each other. In the above latin square the symbols 0,1,2 correspond to the levels of the third factor, c. Now if we set up a second latin square in which the symbols, say α,β,γ, correspond to the levels of AB^2, the resulting square will be orthogonal to the first one. The square corresponding to levels of $(AB^2)_{i+2j,\ \mathrm{mod}\ 3}$ is

$000 + 111 + 222 = \alpha$

$021 + 210 + 102 = \beta$

$201 + 012 + 120 = \gamma$

α	γ	β
β	α	γ
γ	β	α

The class of fractional replicates constituted as an n^{-1} fraction of an n^3 factorial becomes an important one to study as it relates to construction of mutually orthogonal latin squares. In particular, all 2^{-3} fractions of a 2^9 and all 3^{-2} fractions of a 3^6 with all possible aliasing structures could produce several sets of mutually orthogonal latin squares. This could have interesting consequences in finite geometry.

The structure of the left-hand set of parameters in an aliasing structure will have a pattern; for example, for n = 4, 5, and 7, the patterns are:

n = 4	n = 5	n = 7
M + ABC	M + ABC	M + ABC
A	A	A
B	B	B
C	C	C
AB^2	AB^2	AB^2
AB^3	AB^3	AB^3
	AB^4	AB^4
		AB^5
		AB^6

Note that although ABC was completely confounded with the mean, any one of the other three-factor interaction components AB^uC^v, $u,v = 1,2,\ldots,n-1$ could have been utilized equally well. Also, note that the levels of C corresponding to symbols produce a latin square, and that the levels of effects below the factor B produce a set of $n-1$ mutually orthogonal latin squares.

In general we want to look at all possible n^{-1} fractions of an n^3 factorial, i.e., the subset of $\binom{n^3}{n^2}$ combinations for which the levels of C are the symbols

in a latin square and to study their patterns especially for $n = 7, 8$, and 9. All possible fractions, or rather all forms of the aliasing structure, could be classified into all types of t mutually orthogonal latin squares, $O(n,t)$ for $t = 1, 2, \ldots, n-1$. Perhaps this is the manner in which the geometries of various values of n can be exhaustively studied. In fractional factorial notation we want to study all possible patterns of $X_{11}^{-1}X_{12}$ in the following matrix equation:

$$\begin{pmatrix} M \\ A \\ B \\ C \\ \vdots \end{pmatrix} + X_{11}^{-1}X_{12}\beta_0 = Y_r$$

where the form of the first vector below the letter C will be determined by the values in $X_{11}^{-1}X_{12}$, the candidates for entry in the vector β_0 are the remaining two- and three-factor interactions, and Y_r is the particular set of n^2 out of n^3 combinations for which the levels of any fourth effect in the first vector form a latin square. Thus, it becomes important to study the properties of $X_{11}^{-1}X_{12}$ even for the 2^m system. The irregular fractions would appear to be the most interesting for $n = 7, 8$, and 9 since regular fractions can be related to complete confounding in section II.1 and to flats and points in the projective geometry.

We now wish to illustrate the use of fractional replication procedures to construct latin squares which are mateless and which have orthogonal mates. To illustrate let us consider the four standard latin squares of order 4 which are (Fisher and Yates [1957]):

A	B	C	D
B	C	D	A
C	D	A	B
D	A	B	C

Square I

A	B	C	D
B	D	A	C
C	A	D	B
D	C	B	A

Square II

A	B	C	D
B	A	D	C
C	D	B	A
D	C	A	B

Square III

A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

Square IV

It is known (Hedayat [1969]) that the first three squares are mateless and that the last square belongs to an $O(4, 3)$ set.

Now number the rows as $0,1,2,3$ and denote these as levels of the factor $\underline{a}$; number the columns as $0,1,2,3$, and as $0,1,2,3$ for A,B,C,D, respectively, and denote these as levels of the factor $\underline{c}$. Then, in factorial notation the above 16 combinations form a one-fourth fraction of a 4^3 factorial treatment design. The aliasing scheme for the fractional replicate given as square IV is

$$M_{..} + ABC$$
$$A + BC + AB^2C^2 + AB^3C^3$$
$$B + AC + AB^2C + AB^3C$$
$$C + AB + ABC^2 + ABC^3$$

where the effects connected with a plus sign are completely confounded with [illegible]. The completion of the remaining two aliasing structures results in the complete aliasing structures for this 4^{-1} fraction of the 4^3 factorial; these two are:

$$AB^2 + A[illegible] + B[illegible]^2 + AB^3C^2$$
$$AB^3 + AC^2 + B[illegible]^3 + AB^2C^3$$

If we use the levels of AB^2 and of AB^3 to form two latin squares, these two with square IV is the an $O(4, 3)$ set of mutually orthogonal latin squares.

Now, let us return to the set of four standard squares given above and we note that only four combinations in square IV are replaced to obtain squares I, II, and III. These are:

	additional combinations	combinations replaced in IV
Square I	112, 130, 310, 332	110, 132, 312, 330
" II	113, 120, 210, 223	110, 123, 213, 220
" III	213, 230, 320, 331	220, 231, 321, 330

The aliasing structure (without the coefficients) is given on the following page for all four standard latin squares of order 4. The 1/4 replicate given by square IV forms a regular fraction. The remaining three fractional replicates are such that none of the additional effects are unconfounded with the effects A, B, or C of the original latin squares of order 4. Since this is true no linear combinations of these effects will be unconfounded. In order to form a latin square which is orthogonal to the given one it is necessary that there be a set of effects which is unconfounded with the effects in the given square. This is impossible for the three squares I, II, and III and hence the squares are mateless as is well-known.

It would be interesting to ascertain the aliasing structures for the six standard latin squares of order 5 belonging to the $O(5,4)$ set and for the fifty standard latin squares of order 5 for which are known to be mateless (Hedayat [1969]). After a study of these fractions, one should continue such a study for $n = 7$, 8, and 9. It is suggested that one consider a 2^{6-2} fraction instead of a 4^{3-1} fraction for $n = 4$ and a 2^{9-3} fraction instead of an 8^{3-1} fraction for $n = 8$. The reason for this is that there is much more theory available for the

Aliasing structure of effects in the four 1/4 fractional replicates of a 4^3 factorial for four standard latin squares of order 4

Effect	Square I				Square II				Square III				Square IV			
	Effect				Effect				Effect				Effect			
	mean = M	rows = A	cols. = B	letters = C	mean = M	rows = A	cols. = B	letters = C	mean = M	rows = A	cols. = B	letters = C	mean = M	rows = A	cols. = B	letters = C
M	-				-				-				-			
A		-				-				-				-		
B			-				-				-				-	
C				-				-				-				-
AB	P	P	P	P				P				P				C
AB^2				P				P				P				
AB^3				P				P				P				
AC			P				P				P				C	
AC^2			P				P				P					
AC^3			P				P				P					
BC	P	P	P			P				P				C		
BC^2	P	P	P			P				P						
BC^3	P	P	P			P				P						
ABC	P	P	P	P	P	P	P	P		P	P	P	C			
ABC^2	P	P	P	P	P	P	P	P	P	P	P	P				C
ABC^3	P	P	P	P	P	P	P	P	P	P	P	P				C
AB^2C	P	P	P	P	P	P	P	P	P	P	P	P			C	
AB^2C^2	P	P	P	P	P	P	P	P	P	P	P	P		C		
AB^2C^3	P	P	P	P		P	P	P		P	P	P				
AB^3C	P	P	P	P	P	P	P	P	P	P	P	P			C	
AB^3C^2		P	P	P		P	P	P		P	P	P				
AB^3C^3	P	P	P	P	P	P	P	P	P	P	P	P		C		
No. of effects confounded with	12	13	16	12	7	12	12	12	6	12	12	12	1	3	3	

- means identical effect C means complete confounding
P means partial confounding blank means unconfounded

$s = 2$ in the s^m series than for any other value of s. Also, one may use the generalized defining contrast which has been developed by Raktoe and Federer [1969] to a considerable advantage in writing out aliasing structures in these cases. Investigation of the regular and irregular fractional replicates obtainable for various values of n could lead to considerable advances in the theory of mutually orthogonal latin squares.

IV. ANOVA Construction of O(n,t) Sets

There should be some procedure which would utilize the orthogonality of single degree of freedom contrasts in the analysis of variance (ANOVA) and which could be utilized to construct orthogonal latin squares. For example, one could make use of orthogonal polynomial coefficients for row and column contrasts and then construct mutually orthogonal latin squares from these. To illustrate, consider the latin square of order 4 used previously wherein the row-column intersections are numbered as a 2^4 factorial, i.e.:

	Column			
Row	1	2	3	4
1	0000	0001	0010	0011
2	0100	0101	0110	0111
3	1000	1001	1010	1011
4	1100	1101	1110	1111

The relation between the 16 contrasts using orthogonal polynomial coefficients and the 2^4 factorial is given below:

Source of variation	df		
C. F. M.	1		
Row contrasts	3		
$A = -R_L - 2R_C$	1	1	Rows linear = $R_L = A + 2B$
$B = -2R_L + R_C$	1	1	" quadratic = $R_Q = AB$
$AB = R_Q$	1	1	" cubic = $R_C = 2A - B$
Column contrasts	3		
$C = -C_L - 2C_C$	1	1	Columns linear = $C_L = C + 2D$
$D = -2C_L + C_C$	1	1	" quadratic = $C_Q = CD$
$CD = C_Q$	1	1	" cubic = $C_C = 2C - D$
Roman numbers = (AB^{u_1})	3		
$AC = R_L C_L + 4R_C C_C$	1	1	$R_L C_L$
$BD = 4R_L C_L + R_C C_C$	1	1	$R_C C_C$
$ABCD = R_Q C_Q$	1	1	$R_Q C_Q$
Greek letters = (AB^{u_2})	3		
$ABD = -2R_Q C_L + R_Q C_C$	1	1	$R_L C_Q$
$BC = 2R_L C_L - 2R_C C_C + 4R_L C_C - R_C C_L$	1	1	$R_Q C_C$
$ACD = (-R_L - 2R_C) C_Q$	1	1	$R_C C_L$
Latin letters = (AB^{u_3})	3		
$AD = 2R_L C_L - 2R_C C_C - R_L C_C + 4R_C C_L$	1	1	$R_L C_C$
$ABC = R_Q(-C_L - 2C_C)$	1	1	$R_Q C_L$
$BCD = (-2R_L + R_C) C_Q$	1	1	$R_C C_Q$
Total	16		

The individual degree of freedom contrast matrix for the above 16 combinations is:

	Combination															
Contrast	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
R_L	-3	-3	-3	-3	-	-	-	-	+	+	+	+	3	3	3	3
R_Q	+	+	+	+	-	-	-	-	-	-	-	-	+	+	+	+
R_C	+	+	+	+	-3	-3	-3	-3	3	3	3	3	-	-	-	-
C_L	-3	-	+	3	-3	-	+	3	-3	-	+	3	-3	-	+	3
C_Q	+	-	-	+	+	-	-	+	+	-	-	+	+	-	-	+
C_C	+	-3	3	-	+	-3	3	-	+	-3	3	-	+	-3	3	-
$R_L C_L$	9	3	-3	-9	3	+	-	-3	-3	-	+	3	-9	-3	3	9
$R_L C_Q$	-3	3	3	-3	-	+	+	-	+	-	-	+	3	-3	-3	3
$R_L C_C$	-3	9	-9	3	-	3	-3	+	+	-3	3	-	3	-9	9	-3
$R_Q C_L$	-3	-	+	3	3	+	-	-3	3	+	-	-3	-3	-	+	3
$R_Q C_Q$	+	-	-	+	-	+	+	-	-	+	+	-	+	-	-	+
$R_Q C_C$	+	-3	3	-	-	3	-3	+	-	3	-3	+	+	-3	3	-
$R_C C_L$	-3	-	+	3	9	3	-3	-9	-9	-3	3	9	3	+	-	-3
$R_C C_Q$	+	-	-	+	-3	3	3	-3	3	-3	-3	3	-	+	+	-
$R_C C_C$	+	-3	3	-	-3	9	-9	3	3	-9	9	-3	-	3	-3	+

The corresponding single degree of freedom contrast matrix for the 2^4 factorial is:

Combination

Contrast	(1)	a	b	ab	c	ac	bc	abc	d	ad	bd	abd	cd	acd	bcd	abcd
Mean	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
A	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+
B	−	−	−	−	+	+	+	+	−	−	−	−	+	+	+	+
AB	+	+	+	+	−	−	−	−	−	−	−	−	+	+	+	+
C	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+	+
D	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+
CD	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+
AC	+	+	−	−	+	+	−	−	−	−	+	+	−	−	+	+
BD	+	−	+	−	−	+	−	+	+	−	+	−	−	+	−	+
ABCD	+	−	−	+	−	+	+	−	−	+	+	−	−	+	+	−
ABD	+	−	+	−	−	+	−	+	+	−	+	−	−	+	−	+
BC	+	+	−	−	−	−	+	+	+	+	−	−	−	−	+	+
ACD	−	+	+	−	−	+	+	−	+	−	−	+	+	−	−	+
ABC	−	−	+	+	+	+	−	−	+	+	−	−	−	−	+	+
AD	+	−	+	−	+	−	+	−	−	+	−	+	−	+	−	+
BCD	−	+	+	−	+	−	−	+	−	+	+	−	+	−	−	+

The particular contrast matrix utilized is not unique as has been demonstrated above. All orthogonal contrast matrices resulting in latin squares could be considered. For example, other sets of contrasts among rows (or columns) could be:

	1	2	3	4
Mean	+	+	+	+
R_1	-	-	0	0
R_2	0	0	-	+
R_3	+	+	-	-

or

	1	2	3	4
Mean	+	+	+	+
R_1	-	+	0	0
R_2	+	+	-2	0
R_3	+	+	+	-3

The interaction of row and column contrasts possibly could be utilized to allocate the symbols in the latin square.

We wish to illustrate the method of constructing latin squares using orthogonal polynomial coefficients. We shall first consider the construction of three mutually orthogonal latin squares of order 4 and then we shall consider the construction of a single latin square of order 6. In the preceding table on orthogonal polynomials for $n = 4$ denote all combinations with a plus sign as belonging to $(R_L C_L)_1$ and those with a minus sign as belonging to $(R_L C_L)_0$. Do likewise for the $R_Q C_Q$ and $R_C C_C$ effects. Then, the four latin square symbols are obtained as follows:

$$(R_L C_L)_1, (R_Q C_Q)_1, (R_C C_C)_1 = 0000 + 0101 + 1010 + 1111 = A$$

$$(R_L C_L)_1, (R_Q C_Q)_0, (R_C C_C)_0 = 0001 + 0100 + 1011 + 1110 = B$$

$$(R_L C_L)_0, (R_Q C_Q)_0, (R_C C_C)_1 = 0010 + 0111 + 1000 + 1101 = C$$

$$(R_L C_L)_0, (R_Q C_Q)_1, (R_C C_C)_0 = 0011 + 0110 + 1001 + 1100 = D$$

This results in the following latin square of order 4

A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

Likewise, if we use the following polynomial contrasts we obtain the two mutually orthogonal mates of the above square:

$$(R_L C_Q)_1, (R_Q C_C)_0, (R_C C_L)_0 = 0001 + 0110 + 1000 + 1111 = \alpha$$

$$(R_L C_Q)_1, (R_Q C_C)_1, (R_C C_L)_1 = 0010 + 0101 + 1011 + 1100 = \beta$$

$$(R_L C_Q)_0, (R_Q C_C)_0, (R_C C_L)_1 = 0011 + 0100 + 1010 + 1101 = \gamma$$

$$(R_L C_Q)_0, (R_Q C_C)_1, (R_C C_L)_0 = 0111 + 1001 + 1110 + 0000 = \delta$$

and

$$(R_L C_C)_1, (R_Q C_L)_1, (R_C C_Q)_1 = 0011 + 0101 + 1000 + 1110 = I$$

$$(R_L C_C)_1, (R_Q C_L)_0, (R_C C_Q)_0 = 0001 + 0111 + 1010 + 1100 = II$$

$$(R_L C_C)_0, (R_Q C_L)_0, (R_C C_Q)_1 = 0000 + 0110 + 1011 + 1101 = III$$

$$(R_L C_C)_0, (R_Q C_L)_1, (R_C C_Q)_0 = 0010 + 0100 + 1001 + 1111 = IV$$

The above results in the following two latin squares of order 4

δ	α	β	γ
γ	β	α	δ
α	δ	γ	β
β	γ	δ	α

III	II	IV	I
IV	I	III	II
I	IV	II	III
II	III	I	IV

The above method of constructing mutually orthogonal latin squares using polynomial coefficients works for latin squares of order n where $n = 2^p$. We need another procedure for other values of n and shall now construct a latin square of order 6 from the orthogonal polynomial coefficients in the table of single degree of freedom contrasts for 36 combinations. If we observe only the signs of contrasts we note that the 36 combinations may be classified into six sets of four with like signs and two additional sets of six. The latter two sets will be used to build up the six sets of four into six sets of six as follows where all combinations with a plus sign go in the one level and all those with a minus sign go in the zero level:

$(R_2C_2)_1$, $(R_3C_3)_1$, $(R_4C_4)_0$, $(R_5C_5)_0$ + 2 from $(R_1C_1)_1, (R_2C_2)_1 (R_3C_3)_1, (R_4C_4)_1, (R_5C_5)_1$

$(R_2C_2)_0$, $(R_3C_3)_0$, $(R_4C_4)_1$, $(R_5C_5)_0$ + "

$(R_2C_2)_0$, $(R_3C_3)_0$, $(R_4C_4)_0$, $(R_5C_5)_0$ + "

$(R_2C_2)_0$, $(R_3C_3)_1$, $(R_4C_4)_0$, $(R_5C_5)_1$ + 2 from $(R_1C_1)_0, (R_2C_2)_1, (R_3C_3)_0, (R_4C_4)_1, (R_5C_5)_0$

$(R_2C_2)_0$, $(R_3C_3)_1$, $(R_4C_4)_1$, $(R_5C_5)_0$ + "

$(R_2C_2)_1$, $(R_3C_3)_0$, $(R_4C_4)_0$, $(R_5C_5)_1$ + "

From these sets we obtain

$$(12 + 21 + 34 + 43) + (00 + 55) = A$$
$$(02 + 20 + 35 + 53) + (11 + 44) = B$$
$$(01 + 10 + 45 + 54) + (22 + 33) = C$$
$$(04 + 15 + 40 + 51) + (23 + 32) = D$$
$$(03 + 25 + 30 + 52) + (14 + 41) = E$$
$$(13 + 24 + 31 + 42) + (05 + 50) = F$$

This results in the following latin square of order 6:

00 A	10 C	20 B	30 E	40 D	50 F
01 C	11 B	21 A	31 F	41 E	51 D
02 B	12 A	22 C	32 D	42 F	52 E
03 E	13 F	23 D	33 C	43 A	53 B
04 D	14 E	24 F	34 A	44 B	54 C
05 F	15 D	25 E	35 B	45 C	55 A

The pair of treatments in the second set of parentheses, e.g. (00 + 55), were picked from the set of six in such a manner as to have i and j in the combination ij, contain 0, 1, 2, 3, 4, and 5 since each letter must appear once in each row and once in each column.

It would be interesting and perhaps enlightening to carry out the above procedure for $n = 10$ and 12 and to exhaustively study the complete set of 35 contrasts for $n = 6$.

	00	01	02	03	04	05	10	11	12	13	14	15	20	21	22	23	24	25	30	31	32	33	34	35	40	41	42	43	44	45	50	51	52	53	54	55
R_1	-5	-5	-5	-5	-5	-5	-3	-3	-3	-3	-3	-3	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	3	3	3	3	3	3	5	5	5	5	5	5
R_2	5	5	5	5	5	5	-1	-1	-1	-1	-1	-1	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-4	-1	-1	-1	-1	-1	-1	5	5	5	5	5	5
R_3	-5	-5	-5	-5	-5	-5	7	7	7	7	7	7	4	4	4	4	4	4	-4	-4	-4	-4	-4	-4	-7	-7	-7	-7	-7	-7	5	5	5	5	5	5
R_4	1	1	1	1	1	1	-3	-3	-3	-3	-3	-3	2	2	2	2	2	2	2	2	2	2	2	2	-3	-3	-3	-3	-3	-3	1	1	1	1	1	1
R_5	-1	-1	-1	-1	-1	-1	5	5	5	5	5	5	-10	-10	-10	-10	-10	-10	10	10	10	10	10	10	-5	-5	-5	-5	-5	-5	1	1	1	1	1	1
C_1	-5	-3	-1	1	3	5	-5	-3	-1	1	3	5	-5	-3	-1	1	3	5	-5	-3	-1	1	3	5	-5	-3	-1	1	3	5	-5	-3	-1	1	3	5
C_2	5	-1	-4	-4	-1	5	5	-1	-4	-4	-1	5	5	-1	-4	-4	-1	5	5	-1	-4	-4	-1	5	5	-1	-4	-4	-1	5	5	-1	-4	-4	-1	5
C_3	-5	7	4	-4	-7	5	-5	7	4	-4	-7	5	-5	7	4	-4	-7	5	-5	7	4	-4	-7	5	-5	7	4	-4	-7	5	-5	7	4	-4	-7	5
C_4	1	-3	2	2	-3	1	1	-3	2	2	-3	1	1	-3	2	2	-3	1	1	-3	2	2	-3	1	1	-3	2	2	3	1	1	-3	2	2	-3	1
C_5	-1	5	-10	10	-5	1	-1	5	-10	10	-5	1	-1	5	-10	10	-5	1	-1	5	-10	10	-5	1	-1	5	-10	10	-5	1	-1	5	-10	10	-5	1
R_1C_1	25	15	5	-5	-15	-25	15	9	3	-3	-9	-15	5	3	1	-1	-3	-5	-5	-3	-1	1	3	5	-15	-9	-3	3	9	15	-25	-15	-5	5	15	25
R_2C_2	25	-5	-20	-20	-5	25	-5	1	4	4	1	-5	-20	4	16	16	4	-20	-20	4	16	16	4	-20	-5	1	4	4	1	-5	25	-5	-20	-20	-5	25
R_3C_3	35	-35	-20	20	35	-35	-35	49	28	-28	-49	35	-20	28	16	-16	-28	20	-20	28	16	-16	-28	20	-35	49	28	-28	-49	35	-35	49	28	-28	-49	35
R_4C_4	1	-3	2	2	-3	1	-3	9	-6	-6	9	-3	2	-6	4	4	-6	2	2	-6	4	4	-6	2	-3	9	-6	-6	9	-3	1	-3	2	2	-3	1
R_5C_5	1	-5	10	-10	5	-1	-5	25	-50	50	-25	5	10	-50	100	-100	50	-10	-10	50	-100	100	-50	10	5	-25	50	-50	25	-5	-1	5	-10	10	-5	1

V. Group Construction of O(n,t) Sets

V.0. Introduction. The construction of $O(n,t)$ sets based on groups and their associated mappings such as automorphism, complete mapping, and orthomorphism is the oldest and still the most popular method for n not of the form $4t + 2$, Euler [1782] implicitly utilized some properties of finite groups of order $2t + 1$ and $4t$ for his construction of $O(2t+1,2)$ and $O(4t,2)$ sets, respectively. It was MacNeish [1922] who, for the first time, explicitly (however, not rigorously) utilized group properties for his construction of $O(q^m, q^m-1)$ sets and $O(n, \lambda)$ sets, where q is a prime, m is a positive integer and if $n = q_1^{f_1} q_2^{f_2} \ldots q_r^{f_r}$ is the prime power decomposition of n then $\lambda = \min(q_1^{f_1}, q_2^{f_2}, \ldots, q_r^{f_r})$. The field construction of $O(q^m, q^m - 1)$ sets found independently by Bose [1938] and Stevens [1939] is based on the additive group of $GF(q^m)$ and its related cyclic group of automorphisms. The $O(n,n-1)$ sets for $n = 3,4,5,7,8$ and 9 exhibited by Fisher and Yates [1957] are based on cyclic group and abelian groups. Several beautiful applications of group theory to the existence and non-existence of $O(n,t)$ sets have been found by Mann [1942, 1943, 1944]. The $O(12,5)$ sets found by Johnson et al. [1961] and Bose et al. [1960] are based on abelian groups of order 12. Hedayat [1969] and Hedayat and Federer [1969] have found a series of results on the existence and non-existence of $O(n,t)$ sets through the group theory approach. The interested reader on this subject will find the following references together with the references given to these papers very useful: Page [1951], Page-Hall [1955], Singer [1961], Bruck [1951], and Sade [1958].

The author has no doubt that the reader can find many more interesting papers directly or indirectly related this rich subject.

V. 1. Definitions and Notations.

There are several forms of definitions of latin squares and orthogonal latin squares. The following forms are useful for the results which will follow:

Definition V. 1. 1. A latin square of order n on an n-set $\sum$ is an $n \times n$ matrix whose rows and columns are each a permutation of the set $\sum$. Every latin square of order n may therefore be identified with a set of n permutations $(p_1, p_2, \ldots, p_n)$ where p_i is the permutation associated with the ith row.

Definition V. 1. 2. Let L_i be a latin square of order n on an n-set $\sum_i$, $i = 1, 2, \ldots, t$. Then, the set $S = \{L_1, L_2, \ldots, L_t\}$ is said to be a mutually orthogonal set of t latin squares if the projection of the superimposed form of the t latin squares on any two n-sets $\sum_i$ and $\sum_j$, $i \neq j$, forms a permutation of the cartesian product set of $\sum_i$ and $\sum_j$. Such a set is denoted as an $O(n,t)$ set.

Definition V. 1. 3. If $L_1 = (P_{11}, P_{12}, \ldots, P_{1n})$ and $L_2 = (P_{21}, P_{22}, \ldots, P_{2n})$ are two latin squares of order n on an n-set $\sum$, then we may define $L_1 L_2$ to be $L_3 = (P_{11}P_{21}, P_{12}P_{22}, \ldots, P_{1n}P_{2n})$. The generalization to the product of $t > 2$ latin squares follows immediately.

V. 2. 2. Construction of $O(n,t)$ Sets Based on a Group:

We shall divide the problem into three parts based on whether n is a prime, or a mixture of prime powers. The proof of the subsequent results can be found in the references related to this section.

V. 2. 1. $n = q$ a prime. Recall that any prime ordered group is cyclic.

Theorem V. 2. 1. 1. Let $G = \{P_1, P_2, \ldots, P_q\}$ be a cyclic permutation group of degree q and order q. Then, $S_{11} = \{L_1, L_2, \ldots, L_{q-1}\}$ is an $O(q, q-1)$ set, where $L_i = (P_1^i, P_2^i, \ldots, P_q^i)$.

Demonstration V. 2. 1. 1. Let $q = 5$. Select any arbitrary generator such as $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 2 & 1 & 4 \end{pmatrix}$ which generates a cyclic permutation group G and, hence, a latin square L_1. Then,

$L_1 =$

3	5	2	1	4
2	4	5	3	1
5	1	4	2	3
4	3	1	5	2
1	2	3	4	5

, $L_2 =$

2	4	5	3	1
4	3	1	5	2
3	5	2	1	4
5	1	4	2	3
1	2	3	4	5

, $L_3 =$

5	1	4	2	3
3	5	2	1	4
4	3	1	5	2
2	4	5	3	1
1	2	3	4	5

, $L_4 =$

4	3	1	5	2
5	1	4	2	3
2	4	5	3	1
3	5	2	1	4
1	2	3	4	5

For those who do not like to work with permutation groups we present the following theorem:

Theorem V. 2. 1. 2. Let $L(r)$ be an $n \times n$ square with $ri + j \pmod q$ in its (i, j)th cell. $i, j = 0, 1, \ldots, q-1$. Then, $S_{12} = \{L(1), L(2), \ldots, L(q-1)\}$ is an $O(q, q-1)$ set if q is a prime.

Demonstration V. 2. 1. 2. Let $q = 5$; then,

$L(1) =$

0	1	2	3	4
1	2	3	4	0
2	3	4	0	1
3	4	0	1	2
4	0	1	2	3

, $L(2) =$

0	1	2	3	4
2	3	4	0	1
4	0	1	2	3
1	2	3	4	0
3	4	0	1	2

, $L(3) =$

0	1	2	3	4
3	4	0	1	2
1	2	3	4	0
4	0	1	2	3
2	3	4	0	1

, $L(4) =$

0	1	2	3	4
4	0	1	2	3
3	4	0	1	2
2	3	4	0	1
1	2	3	4	0

Note that L(1) in theorem V.2.1.2 is based on the cyclic permutation group generated by $\begin{pmatrix} 0 & 1 & 2 & \dots & q-1 \\ 1 & 2 & 3 & \dots & 0 \end{pmatrix}$ and $L(i) = L^i(1)$, $i = 2, 3, \dots, q-1$. Hence theorem V.2.1.2 is a special case of theorem V.2.1.1.

V.2.2. $n = q^m$ where q is a prime and m any positive integer. Note that this case in particular for $m = 1$ includes case 1. We shall present three theorems for this case. The first two are based on cyclic groups and the third one is based on any group which admits an automorphism of order t.

Theorem V.2.2.1. Let $G = \{P_1, P_2, \dots, P_n\}$ be a cyclic permutation group of degree n and order n. Then, $S_{21} = \{L_1, L_2, \dots, L_\lambda\}$ is an $O(n, \lambda)$ set where $n = q^m$ and $\lambda = q-1$.

Demonstration V.2.2.1. Let $n = 3^2 = 9$. Select any arbitrary generator such as $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 4 & 5 & 1 & 6 & 7 & 8 & 9 & 2 \end{pmatrix}$ which generates a cyclic permutation group G and hence, a latin square L. Then, since $\lambda = 2$,

$L_1 =$

3	4	5	1	6	7	8	9	2
5	1	6	3	7	8	9	2	4
6	3	7	5	8	9	2	4	1
7	5	8	6	9	2	4	1	3
8	6	9	7	2	4	1	3	5
9	7	2	8	4	1	3	5	6
2	8	4	9	1	3	5	6	7
4	9	1	2	3	5	6	7	8
1	2	3	4	5	6	7	8	9

and $L_2 =$

5	1	6	3	7	8	9	2	4
7	5	8	6	9	2	4	1	3
9	7	2	8	4	1	3	5	6
4	9	1	2	3	5	6	7	8
3	4	5	1	6	7	8	9	2
6	3	7	5	8	9	2	4	1
8	6	9	7	2	4	1	3	5
2	8	4	9	1	3	5	6	7
1	2	3	4	5	6	7	8	9

is an $O(9, 2)$ set.

Conjecture. The set S_{21} is orthogonally locked, meaning that there does not exist a latin square L^* such that $S_{21} \cup \{L^*\}$ is an $O(n, \lambda+1)$ set if n is not

a prime. Note that for n even this conjecture is correct since any latin square of even order based on cyclic permutation group is orthogonally mateless.

An analogous theorem to theorem V.2.1.2 for this case is:

Theorem V.2.2.2. Let L(r) be an $n \times n$ square with $ri + j \pmod n$ in its (i,j) cell, $i = 0,1,2,\ldots,n-1$. Then $S_{22} = \{L(1), L(2), \ldots, L(\lambda)\}$ is an $O(n,\lambda)$ set if $n = q^m$ and $\lambda = q-1$.

Demonstration V.2.2.2. Let $n = q = 3^2$ then,

$L(1) =$

0	1	2	3	4	5	6	7	8
1	2	3	4	5	6	7	8	0
2	3	4	5	6	7	8	0	1
3	4	5	6	7	8	0	1	2
4	5	6	7	8	0	1	2	3
5	6	7	8	0	1	2	3	4
6	7	8	0	1	2	3	4	5
7	8	0	1	2	3	4	5	6
8	0	1	2	3	4	5	6	7

and $L(2) =$

0	1	2	3	4	5	6	7	8
2	3	4	5	6	7	8	0	1
4	5	6	7	8	0	1	2	3
6	7	8	0	1	2	3	4	5
8	0	1	2	3	4	5	6	7
1	2	3	4	5	6	7	8	0
3	4	5	6	7	8	0	1	2
5	6	7	8	0	1	2	3	4
7	8	0	1	2	3	4	5	6

is an $O(9,2)$ set. Note that theorem V.2.2.2 is a special case of theorem V.2.2.1 viz., L(1) is based on the cyclic permutation group generated by $\begin{pmatrix} 0 & 1 & 2 & \ldots & n-1 \\ 1 & 2 & 3 & \ldots & 0 \end{pmatrix}$ and $L(i) = L^i(1)$, $i = 2,\ldots,\lambda$.

Theorem V.2.2.3. Let $G = \{a_1 = e$ the identity, $a_2, \ldots, a_n\}$ be a group of order n and α an automorphism of order t on G. Then,

1) $S = \{L_1, L_2, \ldots, L_t\}$ is an $O(n,t)$ set, where

$L_i =$

e	a_2	$\dots$	a_n
$\alpha^i(a_2)$	$\alpha^i(a_2)a_2$	$\dots$	$\alpha^i(a_2)a_n$
$\alpha^i(a_3)$	$\alpha^i(a_3)a_2$	$\dots$	$\alpha^i(a_3)a_n$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\alpha^i(a_n)$	$\alpha^i(a_n)a_2$	$\dots$	$\alpha^i(a_n)a_n$

$i = 1, 2, \dots, t$.

2) If in particular $t = n-1$, then one can simplify the construction of an $O(n, n-1)$ set from the following key latin square by a cyclic permutation of its last $n-1$ rows.

$L_0 =$

e	$\alpha(x)$	$\alpha^2(x)$	$\dots$	$\alpha^t(x)$
$\alpha(x)$	$\alpha(x)\alpha(x)$	$\alpha(x)\alpha^2(x)$	$\dots$	$\alpha(x)\alpha^t(x)$
$\alpha^2(x)$	$\alpha^2(x)\alpha$	$\alpha^2(x)\alpha^2(x)$	$\dots$	$\alpha^2(x)\alpha^t(x)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\alpha^t(x)$	$\alpha^t(x)\alpha(x)$	$\alpha^t(x)\alpha^2(x)$	$\dots$	$\alpha^t(x)\alpha^t(x)$

for any x in G except the identity element.

We see, therefore, that by means of theorem V. 2. 2. 3 one can construct an $O(n,t)$ set if we can find a group G and an automorphism α of order t . In particular, if $t = n-1$ the whole task of construction reduces to the construction of L_0 as described above. If $n = q^m$ then because every elementary

abelian q-group G of order n admits an automorphism α of order n-1, we can construct an $O(q^m, q^m - 1)$ set based on G and α. Here we present a general method of constructing such an automorphism for any $n = q^m$. In particular, we shall exhibit such an automorphism for the following n:

$n = 2^m$, $m = 2, 3, \ldots, 9$

$n = 3^m$, $m = 2, 3, \ldots, 6$

$n = 5^m$, $m = 2, 3, 4$

$n = 7^m$, $m = 2, 3$

$n = 11^2, 13^2, 17^2, 19^2, 23^2, 29^2$, and 31^2.

This will then perhaps be the largest table that has ever been produced so far for $O(n, n-1)$ sets.

Note that there is no loss of generality if we limit ourselves to the following elementary abelian q-group of order $n = q^m$.

$$G^* = \{(b_1 b_2 \ldots b_m),\ b_j = 0, 1, 2, \ldots, q-1,\ j = 1, 2, \ldots, m\}.$$

The binary operation on G^* is addition mod q componentwise, viz., $(b_1 b_2 \ldots b_m) + (b'_1 b'_2 \ldots b'_m) = (c_1 c_2 \ldots c_m)$ where $c_i = b_i + b'_i \pmod q$. Note that the elements of G^* are simply the treatment combinations of m factors each at q levels. The reason why we have chosen this particular elementary abelian q-group is that it has a well-known structure to those who are concerned with experiment design construction. Note also that G^* is the direct product of m Galois fields, each of order q.

The generator set for every elementary abelian q-group of order q^m consists of m elements, and for uniformity, we may choose the following ordered

generator set for G^* .

$$g = \{(100 \ldots 0), (01,00 \ldots 0), \ldots (00 \ldots, 010), (00 \ldots 01)\} .$$

Note that the structure of every automorphism α on G^* is completely defined if we know the image of each element of g under α . G^* is a vector space of dimension m over $GF(q)$.

Before proceeding further we need the following well-known result:

Theorem V, 2, 2, 4, Let G be an elementary abelian q-group of order $n = q^m$. Then, Auto G is isomorphic to the (multiplicative) group of all non-singular $m \times m$ matrices with entries in the field of integers mod q .

The construction of an automorphism of order $n-1$ for G^* is equivalent to the construction of an $m \times m$ matrix A such that $A^{n-1} = I$ but $A^t \neq I$ if t is not a multiple of $n-1$ over the field of integers mod q .

We know from linear algebra that if ϕ is a linear map on a vector space V and if $x \in V$ such that $x \neq 0$ but $\phi(x) = x$, then 1 is an eigenvalue of ϕ. Moreover, if $\{\lambda_1, \lambda_2, \ldots, \lambda_t\}$ is the set of eigenvalues of ϕ, then $\{\lambda_1^s, \lambda_2^s, \ldots, \lambda_t^s\}$ is the set of eigenvalues of ϕ^s. Therefore, for our problem we must find a linear map on G^* with a set of eigenvalues λ_i having the property that for each i, $\lambda_i^s \neq 1 \pmod q$ for all $s = 1,2,\ldots,n-2$ and $\lambda_i^{n-1} = 1$. To do so let F be a $GF(q^m)$ and let β be a generator of the multiplicative cyclic group of $GF(q^m)$, i.e. $\beta^i \neq 1$, $i = 1,2,\ldots,n-2$ while $\beta^{n-1} = 1$. Let $f(x)$ be a monic irreducible polynomial over $GF(q)$ for β. Note that $f(x)$ has degree m. β is sometimes called a primitive root or mark of F. Now, if we let A be the companion matrix for β, then it is easy to see that A has the desired property.

Example

Let us find an automorphism of order 3 for $G^* = \{(00), (01), (10), (11)\}$. It is sufficient, by previous arguments, to find a 2×2 matrix A of order 3 over the field of integer mod 2. Let $GF(2^2) = \{0, 1, \beta, \beta + 1\}$ with the following addition (+) and multiplication (·) tables

+	0	1	β	$\beta + 1$
0	0	1	β	$\beta + 1$
1	1	0	$\beta + 1$	β
β	β	$\beta + 1$	0	1
$\beta + 1$	$\beta + 1$	β	1	0

·	0	1	β	$\beta + 1$
0	0	0	0	0
1	0	1	β	$\beta + 1$
β	0	β	$\beta + 1$	1
$\beta + 1$	0	$\beta + 1$	1	β

Note that β is a primitive root for $GF(2^2)$ and $f(x) = x^2 + x + 1$ is a monic irreducible polynomial for β, since $f(\beta) \equiv 0 \pmod 2$. The companion matrix associated with $f(x)$ is

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

As a check

$$A^2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \equiv \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \text{ over } GF(2), \quad A^3 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

over $GF(2)$.

Let us now determine the image of the ordered generator set $g = \{(10), (01)\}$ under A.

$$Ag = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (10) \\ (01) \end{bmatrix} = \begin{bmatrix} 0\,(10) + 1\,(01) \\ 1\,(10) + 1\,(01) \end{bmatrix} = \begin{bmatrix} (01) \\ (11) \end{bmatrix}.$$

Therefore, $A(10) = (01)$, $A(01) = (11)$, and since $(11) = (10) + (01)$, $(00) = 2(10) + 2(01)$, we have $A(11) = (10)$, $A(00) = (00)$.

Now, we have a group G^* of order 4 and an automorphism of order 3 on G^*. We can now construct an $O(4,3)$ set. Since $e = (00)$, and if we let $x = (10)$ in theorem V.2.2.3, we obtain:

$L_0 =$

(00)	$A(10)$	$A^2(10)$	$A^3(10)$
$A(10)$	$A(10)\,A(10)$	$A(10)\,A^2(10)$	$A(10)\,A^3(10)$
$A^2(10)$	$A^2(10)A(10)$	$A^2(10)A^2(10)$	$A^2(10)A^3(10)$
$A^3(10)$	$A^3(10)A(10)$	$A^3(10)A^2(10)$	$A^3(10)A^3(10)$

$=$

(00)	(01)	(11)	(10)
(01)	(00)	(10)	(11)
(11)	(10)	(00)	(01)
(10)	(11)	(01)	(00)

.

The other two latin squares are obtained by a cyclic permutation of the last three rows of L_0. Thus,

$L_1 =$

(00)	(01)	(11)	(10)
(10)	(11)	(01)	(00)
(01)	(00)	(10)	(11)
(11)	(10)	(00)	(01)

(00)	(01)	(11)	(10)
(11)	(10)	(00)	(01)
(10)	(11)	(01)	(00)
(01)	(00)	(10)	(11)

$= L_2$.

To simplify the notation we set (00) = 1, (01) = 2, (11) = 3, (10) = 4 to obtain:

$L_0 =$

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

, $L_1 =$

1	2	3	4
4	3	2	1
2	1	4	3
3	4	1	2

, and $L_3 =$

1	2	3	4
3	4	1	2
4	3	2	1
2	1	4	3

We are now ready to exhibit a generating matrix of order $n-1 = q^m - 1$ with entries from GF(q) for those n promised before.

n	Generator	Order	m	Generator	Order
2^2	$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$	3	2^3	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$	7
2^4	$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$	15	2^5	$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \end{bmatrix}$	31
2^6	$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$	63	2^7	$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$	127
2^8	$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$	255	2^9	$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$	511

n	Generator	Order	n	Generator	Order
3^2	$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$	8	3^3	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$	26
3^4	$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{bmatrix}$	80	3^5	$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$	242
3^6	$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$	728	5^3	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 2 & 0 \end{bmatrix}$	124
5^2	$\begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}$	24	7^2	$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$	48
5^4	$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 3 & 3 & 0 & 3 \end{bmatrix}$	624	11^2	$\begin{bmatrix} 0 & 1 \\ 3 & 3 \end{bmatrix}$	120
7^3	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 2 & 0 \end{bmatrix}$	342	13^2	$\begin{bmatrix} 0 & 1 \\ 5 & 5 \end{bmatrix}$	168
17^2	$\begin{bmatrix} 0 & 1 \\ 5 & 5 \end{bmatrix}$	288	19^2	$\begin{bmatrix} 0 & 1 \\ 4 & 4 \end{bmatrix}$	360
23^2	$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$	528	27^2	$\begin{bmatrix} 0 & 1 \\ 3 & 3 \end{bmatrix}$	728
29^2	$\begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}$	840	31^2	$\begin{bmatrix} 0 & 1 \\ 2 & 2 \end{bmatrix}$	960

To shed more light on the given procedure we go through another example. Let $n = 2^3$. Then

$G^* = \{(000), (001), (010), (011), (100), (101), (110), (111)\}$

$g = \{(100), (010), (001)\}$ and $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

$$Ag = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} (100) \\ (010) \\ (001) \end{bmatrix} = \begin{bmatrix} (010) \\ (001) \\ (101) \end{bmatrix} .$$

Let x in theorem V.2.2.3 be (100). Then,

$A(100) = (010)$,

$A^2(100) = (001)$,

$A^3(100) = (101)$,

$A^4(100) = (111)$,

$A^5(100) = (110)$,

$A^6(100) = (011)$, and

$A^7(100) = (100)$.

Therefore, we obtain L_0 as follows:

$L_0 =$

(000)	(010)	(001)	(101)	(111)	(110)	(011)	(100)
(010)	(000)	(011)	(111)	(101)	(100)	(001)	(110)
(001)	(011)	(000)	(100)	(110)	(111)	(010)	(101)
(101)	(111)	(100)	(000)	(010)	(011)	(110)	(001)
(111)	(101)	(110)	(010)	(000)	(001)	(100)	(011)
(110)	(100)	(111)	(011)	(001)	(000)	(101)	(010)
(011)	(001)	(010)	(110)	(100)	(101)	(000)	(111)
(100)	(110)	(101)	(011)	(011)	(010)	(111)	(000)

Setting (000) = 1, (010) = 2, (001) = 3, (101) = 4, (111) = 5, (110) = 6 , (011) = 7, (100) = 8, then L_0 in a compact form will be:

$L_0 =$

1	2	3	4	5	6	7	8
2	1	7	5	4	8	3	6
3	7	1	8	6	5	2	4
4	5	8	1	2	7	6	3
5	4	6	2	1	3	8	7
6	8	5	7	3	1	4	2
7	3	2	6	8	4	1	5
8	6	4	3	7	2	5	1

.

Now, we can derive $L_1, L_2, \ldots, L_6$ from L_0 by a cyclic permutation of the last 7 rows of L_0, for example,

$L_1 =$

1	2	3	4	5	6	7	8
8	6	4	3	7	2	5	1
2	1	7	5	4	8	3	6
3	7	1	8	6	5	2	4
4	5	8	1	2	7	6	3
5	4	6	2	1	3	8	7
6	8	5	7	3	1	4	2
7	3	2	6	8	4	1	5

, $L_2 =$

1	2	3	4	5	6	7	8
7	3	2	6	8	4	1	5
8	6	4	3	7	2	5	1
2	1	7	5	4	8	3	6
3	7	1	8	6	5	2	4
4	5	8	1	2	7	6	3
6	8	5	7	3	1	4	2
6	8	5	7	3	1	4	2

,

and so on. Note the way L_1 is derived from L_0: except for the first row of L_0 and L_1, which are identical, the i^{th} row of L_0 becomes the $(i+1)^{th}$ row of L_1, and the last row of L_0 becomes the second row of L_1. In general L_j is derived from L_{j-1} in the same fashion as L_1 is derived from L_0.

V. 2. 3. $n = q_1^{m_1} q_2^{m_2} \ldots q_r^{m_r}$, <u>where q_i is a prime such that $q_i \neq q_j$ if $i \neq j$</u> and m_i is a positive integer, $i = 1, 2, \ldots, r$.

Theorem V.2.3.1. Let $n = q_1^{m_1}, q_2^{m_2}, \ldots, q_r^{m_r}$ be the prime power decomposition of n. Then, there exists an $O(n, \gamma)$ set based on a group, where $\gamma = \min(q_1^{m_1}, q_2^{m_2}, \ldots, q_r^{m_r}) - 1$.

Construction. Let $n_i = q_i^{m_i}$. Then, by the method of theorem V.2.2.3 construct an $O(n_i, n_i - 1)$ set $S_i = \{L_{i1}, L_{i2}, \ldots, L_{in_i-1}\}$, $i = 1, 2, \ldots, r$. Now, let $S_i^* = \{L_{i1}, L_{i2}, \ldots, L_{i\gamma}\}$, $i = 1, 2, \ldots, r$. Then, $H = \{A_1, A_2, \ldots, A_\gamma\}$ is an $O(n, \gamma)$ set where $A_j = L_{1j} \otimes L_{2j} \otimes \ldots \otimes L_{rj}$. $\otimes$ denotes the Kronecker product operation.

Demonstration V.2.3.1. Let $n = 12 = 2^2 \cdot 3$. Then, $\gamma = 2$,

$S_1 = L_{11} =$

1	2	3
2	3	1
3	1	2

, $L_{12} =$

1	2	3
3	1	2
2	3	1

,

$S_2 = L_{21} =$

1	2	3	4
2	1	4	3
3	4	1	2
4	3	2	1

, $L_{22} =$

1	2	3	4
4	3	2	1
2	1	4	3
3	4	1	2

, $L_{23} =$

1	2	3	4
3	4	1	2
4	3	2	1
2	1	4	3

$S_1^* = \{L_{11}, L_{12}\}$ and S_2^* say $\{L_{21}, L_{22}\}$. Then, the reader can easily verify that

$$H = \{A_1 = L_{11} \otimes L_{21},\ A_2 = L_{12} \otimes L_{22}\}$$

is an $O(12, 2)$ set.

Remark. Let n and γ be the same as in theorem V.2.3.1. Then it can be shown that automorphism method fails to produce more than γ mutually orthogonal latin

squares. We shortly show that this inherent defect is due to the mapping not to the group structure.

Definition V, 2, 1. Consider for each positive integer n an abstract group G of order n with binary operation $*$. Let Ω be the collection of all one-to-one mappings of G into itself. Then two maps σ and ψ in Ω are said to be orthogonal if for any g in G,

$$(\sigma Z) * (\psi Z)^{-1} = g$$

has a unique solution Z in G. In particular if σ is an identity map then ψ is said to be an orthomorphism map. A t-subset of Ω is said to be a mutually orthogonal set if every two maps in this t-subset are orthogonal.

Let $L(\cdot)$ be an $n \times n$ square. We make a one-to-one correspondence between the rows of $L(\cdot)$ and the elements of G. Thus, by row x we shall mean the row corresponding to the element x in G. Similarly we make a one-to-one correspondence between the columns of $L(\cdot)$ and the elements of G. The cell of $L(\cdot)$ which occurs in the intersection of row x and column y is called the cell (x,y).

Theorem V, 2, 3, 2. Let σ be in Ω. Put in the cell (x,y) of $L(\cdot)$ the element $(\sigma x) * y$ of G. Call the resulting square $L(\sigma)$. Then $L(\sigma)$ is a latin square of order n on G. Moreover if $\{\sigma_1, \sigma_2, \ldots, \sigma_t\}$ is a set of t mutually orthogonal maps then $\{L(\sigma_1), \ldots, L(\sigma_t)\}$ is an $O(n,t)$ set.

Demonstration V, 2, 3, 2. Let $G = \{0, 1, 2\}$ with the binary operation $x_1 + x_2 = x_3 \pmod 3$, x_i in G. Then the maps σ and ψ with the following definitions are orthogonal.

$$\sigma(0) = 0 \qquad \psi(0) = 0$$
$$\sigma(1) = 1 \qquad \psi(1) = 2$$
$$\sigma(2) = 2 \qquad \psi(2) = 1 \ .$$

The corresponding latin squares to σ and ψ are:

$$L(\sigma) = \begin{array}{|c|c|c|}\hline 0 & 1 & 2 \\ \hline 1 & 2 & 0 \\ \hline 2 & 0 & 1 \\ \hline \end{array} \ , \quad L(\psi) = \begin{array}{|c|c|c|}\hline 0 & 1 & 2 \\ \hline 2 & 0 & 1 \\ \hline 1 & 2 & 0 \\ \hline \end{array}$$

which are orthogonal.

V. 3. Construction of $O(n,t)$ sets based on t different groups of order n

Up to now we have been concerned with the construction of $O(n,t)$ sets using a group of order n which admits certain mappings. In this section we want to show that for some n's and t's one can construct $O(n,t)$ sets based on t different groups each of order n. This approach proved useful because it lead to the construction of an $O(15, 3)$ set. We should mention that our motivation to search along these lines has stemmed from the following theorem, with a negative flavor, proved by Mann [1944].

Theorem V. 3. 1. It is impossible to construct an $O(5,2)$ set based on two different permutation groups.

For a while we thought that this theorem might be true for other orders. However, it was found that, fortunately, this is not the case as the following two theorems show:

Theorem V. 3. 2. It is possible to construct $O(7,2)$ sets based on two different cyclic permutation groups of order 7.

Proof: By construction $\{L_1, L_2\}$ is an O(7,2) set where

$L_1 =$

1	2	3	4	5	6	7
3	7	6	1	4	2	5
6	5	2	3	1	7	4
2	4	7	6	3	5	1
7	1	5	2	6	4	3
5	3	4	7	2	1	6
4	6	1	5	7	3	2

, $L_2 =$

1	2	3	4	5	6	7
2	3	4	5	6	7	1
3	4	5	6	7	1	2
4	5	6	7	1	2	3
5	6	7	1	2	3	4
6	7	1	2	3	4	5
7	1	2	3	4	5	6

.

L_1 and L_2 are based on two different permutation groups as can easily be seen from the different structure of their rows. To be specific L_1 is based on the cyclic permutation group generated by $\begin{pmatrix}1&2&3&4&5&6&7\\3&7&6&1&4&2&5\end{pmatrix}$ and L_2 is based on the cyclic permutation group generated by $\begin{pmatrix}1&2&3&4&5&6&7\\2&3&4&5&6&7&1\end{pmatrix}$. Note that, since L_1 and L_2 are based on cyclic permutation groups, then by theorem V.2.1.1 $\{L_1\}$ and $\{L_2\}$ can be embedded in O(7,6) sets. However, whether or not $\{L_1, L_2\}$ can be embedded in a larger set is an open problem.

Theorem V.3.3. It is possible to construct O(15, 3) sets based on three different cyclic permutation groups of order 15.

We remind the reader that every group of order 15 is cyclic.

Proof: By construction $\{L_1, L_2, L_3\}$ is an O(15, 3) set where

$L_1 =$

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
2	3	4	5	6	7	8	9	10	11	12	13	14	0	1
4	5	6	7	8	9	10	11	12	13	14	0	1	2	3
6	7	8	9	10	11	12	13	14	0	1	2	3	4	5
8	9	10	11	12	13	14	0	1	2	3	4	5	6	7
10	11	12	13	14	0	1	2	3	4	5	6	7	8	9
12	13	14	0	1	2	3	4	5	6	7	8	9	10	11
14	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1	2	3	4	5	6	7	8	9	10	11	12	13	14	0
3	4	5	6	7	8	9	10	11	12	13	14	0	1	2
5	6	7	8	9	10	11	12	13	14	0	1	2	3	4
7	8	9	10	11	12	13	14	0	1	2	3	4	5	6
9	10	11	12	13	14	0	1	2	3	4	5	6	7	8
11	12	13	14	0	1	2	3	4	5	6	7	8	9	10
13	14	0	1	2	3	4	5	6	7	8	9	10	11	12

generated by $\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 0 & 1 \end{pmatrix}$,

$L_2 =$

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	11	10	7	9	14	13	6	0	3	4	2	8	5	12
11	2	4	6	3	12	5	13	1	7	9	10	0	14	8
2	10	9	13	7	8	14	5	11	6	3	4	1	12	0
10	4	3	5	6	0	12	14	2	13	7	9	11	8	1
4	9	7	14	13	1	8	12	10	5	6	3	2	0	11
9	3	6	12	5	11	0	8	4	14	13	7	10	1	2
3	7	13	8	14	2	1	0	9	12	5	6	4	11	10
7	6	5	0	12	10	11	1	3	8	14	13	9	2	4
6	13	14	1	8	4	2	11	7	0	12	5	3	10	9
13	5	12	11	0	9	10	2	6	1	8	14	7	4	3
5	14	8	2	1	3	4	10	13	11	0	12	6	9	7
14	12	0	10	11	7	9	4	5	2	1	8	13	3	6
12	8	1	4	2	6	3	9	14	10	11	0	5	7	13
8	0	11	9	10	13	7	3	12	4	2	1	14	6	5

generated by $\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 1 & 11 & 10 & 7 & 9 & 14 & 13 & 6 & 0 & 3 & 4 & 2 & 8 & 5 & 12 \end{pmatrix}$, and

L_3 =

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
7	4	8	9	11	2	5	12	13	10	0	14	1	3	6
12	11	13	10	14	8	2	1	3	0	7	6	4	9	5
1	14	3	0	6	13	8	4	9	7	12	5	11	10	2
4	6	9	7	5	3	13	11	10	12	1	2	14	0	8
11	5	10	12	2	9	3	14	0	1	4	8	6	7	13
14	2	0	1	8	10	9	6	7	4	11	13	5	12	3
6	8	7	4	13	0	10	5	12	11	14	3	2	1	9
5	13	12	11	3	7	0	2	1	14	6	9	8	4	10
2	3	1	14	9	12	7	8	4	6	5	10	13	11	0
8	9	4	6	10	1	12	13	11	5	2	0	3	14	7
13	10	11	5	0	4	1	3	14	2	8	7	9	6	12
3	0	14	2	7	11	4	9	6	8	13	12	10	5	1
9	7	6	8	12	14	11	10	5	13	3	1	0	2	4
10	12	5	13	1	6	14	0	2	3	9	4	7	8	11

generated by $\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 7 & 4 & 8 & 9 & 11 & 2 & 5 & 12 & 13 & 10 & 0 & 14 & 1 & 3 & 6 \end{pmatrix}$.

Whether or not (L_1, L_2, L_3) can be embedded in an $O(15, t)$, $t > 3$, set is an open problem.

V.4. Concluding Remark

Johnson et al. [1961] and Bose et al. [1960] independently found, by an electronic computer, five mutually orthogonal latin squares by first finding five mutually orthogonal maps for an abelian group of order 12. The $O(12, 5)$ set exhibited below is the set found by Johnson et al. [1961]. Note that the top square is obtained, after a proper renaming, as the direct product of a latin square of order 2 and a cyclic latin square of order 6 being both orthogonally mateless. Moreover, every other square is obtained by proper row permutations, determined by an orthomorphism, from the top square.

Final Remark. The group method fails to produce an $O(n,t)$ set, $t \geq 2$ for any n of the form $4t + 2$. This is so because the Cayley table of any group of order $n = 4t + 2$, which is a latin square of order n, is orthogonally mateless.

0	1	2	3	4	5	6	7	8	9	10	11
5	0	1	2	3	4	11	6	7	8	9	10
4	5	0	1	2	3	10	11	6	7	8	9
3	4	5	0	1	2	9	10	11	6	7	8
2	3	4	5	0	1	8	9	10	11	6	7
1	2	3	4	5	0	7	8	9	10	11	6
6	7	8	9	10	11	0	1	2	3	4	5
11	6	7	8	9	10	5	0	1	2	3	4
10	11	6	7	8	9	4	5	0	1	2	3
9	10	11	6	7	8	3	4	5	0	1	2
8	9	10	11	6	7	2	3	4	5	0	1
7	8	9	10	11	6	1	2	3	4	5	0

0	1	2	3	4	5	6	7	8	9	10	11
6	7	8	9	10	11	0	1	2	3	4	5
10	11	6	7	8	9	4	5	0	1	2	3
4	5	0	1	2	3	10	11	6	7	8	9
11	6	7	8	9	10	5	0	1	2	3	4
5	0	1	2	3	4	11	6	7	8	9	10
9	10	11	6	7	8	3	4	5	0	1	2
7	8	9	10	11	6	1	2	3	4	5	0
2	3	4	5	0	1	8	9	10	11	6	7
8	9	10	11	6	7	2	3	4	5	0	1
1	2	3	4	5	0	7	8	9	10	11	6
3	4	5	0	1	2	9	10	11	6	7	8

0	1	2	3	4	5	6	7	8	9	10	11
3	4	5	0	1	2	9	10	11	6	7	8
6	7	8	9	10	11	0	1	2	3	4	5
5	0	1	2	3	4	11	6	7	8	9	10
9	10	11	6	7	8	3	4	5	0	1	2
7	8	9	10	11	6	1	2	3	4	5	0
4	5	0	1	2	3	10	11	6	7	8	9
10	11	6	7	8	9	4	5	0	1	2	3
1	2	3	4	5	0	7	8	9	10	11	6
2	3	4	5	0	1	8	9	10	11	6	7
11	6	7	8	9	10	5	0	1	2	3	4
8	9	10	11	6	7	2	3	4	5	0	1

0	1	2	3	4	5	6	7	8	9	10	11
10	11	6	7	8	9	4	5	0	1	2	3
5	0	1	2	3	4	11	6	7	8	9	10
7	8	9	10	11	6	1	2	3	4	5	0
1	2	3	4	5	0	7	8	9	10	11	6
9	10	11	6	7	8	3	4	5	0	1	2
3	4	5	0	1	2	9	10	11	6	7	8
8	9	10	11	6	7	2	3	4	5	0	1
4	5	0	1	2	3	10	11	6	7	8	9
11	6	7	8	9	10	5	0	1	2	3	4
6	7	8	9	10	11	0	1	2	3	4	5
2	3	4	5	0	1	8	9	10	11	6	7

0	1	2	3	4	5	6	7	8	9	10	11
2	3	4	5	0	1	8	9	10	11	6	7
7	8	9	10	11	6	1	2	3	4	5	0
8	9	10	11	6	7	2	3	4	5	0	1
4	5	0	1	2	3	10	11	6	7	8	9
11	6	7	8	9	10	5	0	1	2	3	4
10	11	6	7	8	9	4	5	0	1	2	3
6	7	8	9	10	11	0	1	2	3	4	5
9	10	11	6	7	8	3	4	5	0	1	2
5	0	1	2	3	4	11	6	7	8	9	10
3	4	5	0	1	2	9	10	11	6	7	8
1	2	3	4	5	0	7	8	9	10	11	6

VI. Projecting Diagonals Construction of O(n,t) Sets

A very simple procedure (sort of the "man-on-the-street" approach) of constructing balanced incomplete block and partially balanced incomplete block designs for $v = k^2$ items in incomplete blocks of size k has been utilized since the late 1940's by the author and has its counterpart in constructing O(n,t) sets. First we shall illustrate its use in incomplete block experiment design construction, and then we show how it applies to the construction on O(n,t) set. The theoretical basis for this method may be derived directly from the preceding section.

The procedure becomes apparent through an example. Suppose that $v = 9$ and $k = 3$. After writing the first square as illustrated below, take successive diagonals of the preceding square and use them to form the incomplete blocks of a square, thus:

Square 1

1	2	3
4	5	6
7	8	9

Square 2

1	5	9
2	6	7
3	4	8

Square 3

1	6	8
2	4	9
3	5	7

Square 4

1	4	7
2	5	8
3	6	9

As we have noted this is a resolvable balanced incomplete design with the parameters $v = 9 = k^2$, $k = 3$, $r = 4 = k + 1$, $b = 12 = k(k+1)$, and $\lambda = 1$, where the rows of the above squares form the incomplete blocks.

To form a partially balanced incomplete block design for $v = k^2$ in incomplete blocks of size k one may use any 2, any 3, ..., any k arrangements (or squares). To illustrate the formation of a partially balanced incomplete block

design for $v = 6 = k(k-1)$, $r = 2$ or $3 = k$, and $k' = 2$ simply delete the numbers 7, 8, and 9 from the last $k = 3$ arrangements. The deletion of certain symbols from the set $1, 2, \ldots, v$ is known as "variety cutting". For $k^2 = 25$ and $k = 5$ partially balanced incomplete block designs may be constructed for $v = 10$ and $k^* = 2$, $v = 15$ and $k^+ = 3$, and $v = 20$ and $k' = 4$ by the above "variety cutting" procedure.

Also, the successive diagonals method is useful for $v = k^2$ in incomplete blocks of size k for any odd k. For example, for $v = 225$ and $k = 15$ four arrangements or squares may be quickly constructed by the above method. Likewise, the "variety cutting" procedure may be utilized to obtain 2 or 3 arrangements for $v = 15p$, $2 \leq p \leq 15$, varieties.

The above method has its counterpart in constructing mutually orthogonal latin squares and this possibility is briefly mentioned in Fisher and Yates [1957] in this context. Again the method becomes apparent through an example. First write the latin square in standard order and of the form given below for the first square, then project the main right diagonal of the preceding square into the first column of a square, and then write the symbol order in the same manner as in the first square. As a first example, let the order of the latin square be 3; the squares are:

first square

1	2	3
2	3	1
3	1	2

second square

1	2	3
3	1	2
2	3	1

third square

1	2	3
1	2	3
1	2	3

Thus, the main right diagonal of the first square is 1,3,2 which becomes the first column of the second square. Then, write the first row as 1,2,3, the second row as 3,1,2, and the third row as 2,3,1. For the third square, which is not a latin square, the right main diagonal of the second square is 1,1,1 and this becomes the first column of the third square; the rows are then completed. If we then take the right main diagonal of the third square, we obtain the first square.

As a second illustrative example, the five squares for order n = 5 which are constructed by successively projecting diagonals, are:

first square

1	2	3	4	5
2	3	4	5	1
3	4	5	1	2
4	5	1	2	3
5	1	2	3	4

second square

1	2	3	4	5
3	4	5	1	2
5	1	2	3	4
2	3	4	5	1
4	5	1	2	3

third square

1	2	3	4	5
4	5	1	2	3
2	3	4	5	1
5	1	2	3	4
3	4	5	1	2

fourth square

1	2	3	4	5
5	1	2	3	4
4	5	1	2	3
3	4	5	1	2
2	3	4	5	1

fifth square

1	2	3	4	5
1	2	3	4	5
1	2	3	4	5
1	2	3	4	5
1	2	3	4	5

The fifth square is not a latin square but may be utilized to construct the first square through use of the method of successive projections of the main diagonals.

The method may be utilized for any odd order n and will produce $q_1 - 1$ orthogonal latin squares for $n = q_1, q_2, \ldots, q_s$ where $q_i < q_{i+1}$ and $q_1, q_2, \ldots, q_s$

is the prime power decomposition of n. Thus, for $n = 15 = 3(5)$ a pair $(q_1 - 1 = 3 - 1 = 2)$ of orthogonal latin squares is easily produced. For $n = 35 = 5(7)$, a quartet of mutually orthogonal latin squares is readily produced by the projecting diagonals method.

VII. Relating Between Complete Confounding and Simple Orthomorphisms

We shall illustrate the ideas by going through a complete example taking $n = 12 = 2^2 \times 3$. For this purpose we take the ring of 12 elements (obtained by utilizing Raktoe's [1969] results) as follows:

$GF(2^2)$	I_3	$GF(3)$	I_4
0	0	0	0
1 ➡	3	1 ➡	4
x	3x	2	2
x+1	3x+3		

$R_{12} = I_3 \oplus I_4 = \{0,1,2,3,4,5,3x, 3x+1, 3x+2, 3x+3, 3x+4, 3x+5\}$

R_{12} is a commutative ring under addition and multiplication (moddd6, $3x^2 + 3x + 3$, $4x + 4$) in the following sense:

e.g.: (a). $(3x+3) + (3x+4) = 6x + 7 = 1$; here we have to reduce only mod 6 to get the answer.

(b). $(3x+1) \cdot (3x+4) = 9x^2 + 15x + 4$

$= 3x^2 + 3x + 4$

$= [3+0]x^2 + [3+0]x + [0+4] = (3x^2 + 3x) + 4$

$= 3 + 1 = 1$; here first we had to reduce mod 6, then mod $3x^2 + 3x + 3$ leaving us immediately 3 and 4, which is irreducible mod $4x + 4$, thus resulting in 1.

Explicitly, to facilitate arithmetic, the addition and mutliplication of these 12 elements are:

+	0	1	2	3	4	5	3x	3x+1	3x+2	3x+3	3x+4	3x+5
0	0	1	2	3	4	5	3x	3x+1	3x+2	3x+3	3x+4	3x+5
1		2	3	4	5	0	3x+1	3x+2	3x+3	3x+4	3x+5	3x
2			4	5	0	1	3x+2	3x+3	3x+4	3x+5	3x	3x+1
3				0	1	2	3x+3	3x+4	3x+5	3x	3x+1	3x+2
4					2	3	3x+4	3x+5	3x	3x+1	3x+2	3x+3
5						4	3x+5	3x	3x+1	3x+2	3x+3	3x+4
3x							0	1	2	3	4	5
3x+1								2	3	4	5	0
3x+2									4	5	0	1
3x+3										0	1	2
3x+4											2	3
3x+5												4

·	0	1	2	3	4	5	3x	3x+1	3x+2	3x+3	3x+4	3x+5
0	0	0	0	0	0	0	0	0	0	0	0	0
1		1	2	3	4	5	3x	3x+1	3x+2	3x+3	3x+4	3x+5
2			4	0	2	4	0	2	4	0	2	4
3				3	0	3	3x	3x+3	3x	3x+3	3x	3x+3
4					4	2	0	4	2	0	4	2
5						1	3x	3x+5	3x+4	3x+3	3x+2	3x+1
3x							3x+3	3	3x+3	3	3x+3	3
3x+1								3x+4	5	3x	1	3x+2
3x+2									3x+1	3	3x+5	1
3x+3										3x+3	3	3x
3x+4											3x+1	5
3x+5												3x+4

Now, associate with a latin square of order 12 the $3^2 \times 4^2 = [3 \times 4] \times [3 \times 4]$ $= 12 \times 12$ lattice square with the following breakdown of the 143 degrees of freedom:

A^4	2	C^3	3
B^4	2	D^3	3
A^4B^4	2	C^3D^3	3
A^4B^2	2	C^3D^{3x}	3
		C^3D^{3x+3}	3

A^4C^3	6	B^4C^3	6	$A^4B^4C^3$	6	$A^4B^2C^3$	6
A^4D^3	6	B^4D^3	6	$A^4B^4D^3$	6	$A^4B^2D^3$	6
$A^4C^3D^3$	6	$B^4C^3D^3$	6	$A^4B^4C^3D^3$	6	$A^4B^2C^3D^3$	6
$A^4C^3D^{3x}$	6	$B^4C^3D^{3x}$	6	$A^4B^4C^3D^{3x}$	6	$A^4B^2C^3D^{3x}$	6
$A^4C^3D^{3x+3}$	6	$B^4C^3D^{3x+3}$	6	$A^4B^4C^3D^{3x+3}$	6	$A^4B^2C^3D^{3x+3}$	6

For any row or column confounding we need to confound effects totaling up to 11 degrees of freedom. There are natural candidates available. In fact, we may choose for our first lattice square the confounding scheme in many ways. A scheme resulting in a pair of orthogonal latin squares is the following:

Confounding scheme of our 12 × 12 lattice square

	$(A^4C^3)_0$	$(A^4C^3)_1$	$(A^4C^3)_2$	$(A^4C^3)_3$	$(A^4C^3)_4$	$(A^4C^3)_5$	$(A^4C^3)_{3x}$	$(A^4C^3)_{3x+1}$	$(A^4C^3)_{3x+2}$	$(A^4C^3)_{3x+3}$	$(A^4C^3)_{3x+4}$	$(A^4C^3)_{3x+5}$
$(B^4D^3)_0$	0000	4030	2000	0030	4000	2030	00 3x 0	40 3x+3 0	20 3x 0	00 3x+3 0	40 3x 0	20 3x+3 0
$(B^4D^3)_1$	0403	4433	2403	0433	4403	2433	04 3x 3	44 3x+3 3	24 3x 3	04 3x+3 3	44 3x 3	24 3x+3 3
$(B^4D^3)_2$	0200	4230	2200	0230	4200	2230	02 3x 0	42 3x+3 0	22 3x 0	02 3x+3 0	42 3x 0	22 3x+3 0
$(B^4D^3)_3$	0003	4033	2003	0033	4003	2033	00 3x 3	40 3x+3 3	20 3x 3	00 3x+3 3	40 3x 3	20 3x+3 3
$(B^4D^3)_4$	0400	4430	2400	0430	4400	2430	04 3x 0	44 3x+3 0	24 3x 0	04 3x+3 0	44 3x 0	24 3x+3 0
$(B^4D^3)_5$	0203	4233	2203	0233	4203	2233	02 3x 3	42 3x+3 3	22 3x 3	02 3x+3 3	42 3x 3	22 3x+3 3
$(B^4D^3)_{3x}$	000 3x	403 3x	200 3x	003 3x	400 3x	203 3x	00 3x 3x	40 3x+3 3x	20 3x 3x	00 3x+3 3x	40 3x 3x	20 3x+3 3x
$(B^4D^3)_{3x+1}$	040 3x+3	443 3x+3	240 3x+3	043 3x+3	440 3x+3	243 3x+3	04 3x 3x+3	44 3x+3 3x+3	24 3x 3x+3	04 3x+3 3x+3	44 3x 3x+3	24 3x+3 3x+3
$(B^4D^3)_{3x+2}$	020 3x	423 3x	220 3x	023 3x	420 3x	223 3x	02 3x 3x	42 3x+3 3x	22 3x 3x	02 3x+3 3x	42 3x 3x	22 3x+3 3x
$(B^4D^3)_{3x+3}$	000 3x+3	403 3x+3	200 3x+3	003 3x+3	400 3x+3	203 3x+3	00 3x 3x+3	40 3x+3 3x+3	20 3x 3x+3	00 3x+3 3x+3	40 3x 3x+3	20 3x+3 3x+3
$(B^4D^3)_{3x+4}$	040 3x	443 3x	240 3x	043 3x	440 3x	243 3x	04 3x 3x	44 3x+3 3x	24 3x 3x	04 3x+3 3x	44 3x 3x	24 3x+3 3x
$(B^4D^3)_{3x+5}$	020 3x+3	423 3x+3	220 3x+3	023 3x+3	420 3x+3	223 3x+3	02 3x 3x+3	42 3x+3 3x+3	22 3x 3x+3	02 3x+3 3x+3	42 3x 3x+3	22 3x+3 3x+3

LATIN SQUARE 1: Treatments identified with $A^4 B^4 C^3 D^3$

0	1	2	3	4	5	3x	3x+1	3x+2	3x+3	3x+4	3x+5
1	2	3	4	5	0	3x+1	3x+2	3x+3	3x+4	3x+5	3x
2	3	4	5	0	1	3x+2	3x+3	3x+4	3x+5	3x	3x+1
3	4	5	0	1	2	3x+3	3x+4	3x+5	3x	3x+1	3x+2
4	5	0	1	2	3	3x+4	3x+5	3x	3x+1	3x+2	3x+3
5	0	1	2	3	4	3x+5	3x	3x+1	3x+2	3x+3	3x+4
3x	3x+1	3x+2	3x+3	3x+4	3x+5	0	1	2	3	4	5
3x+1	3x+2	3x+3	3x+4	3x+5	3x	1	2	3	4	5	0
3x+2	3x+3	3x+4	3x+5	3x	3x+1	2	3	4	5	0	1
3x+3	3x+4	3x+5	3x	3x+1	3x+2	3	4	5	0	1	2
3x+4	3x+5	3x	3x+1	3x+2	3x+3	4	5	0	1	2	3
3x+5	3x	3x+1	3x+2	3x+3	3x+4	5	0	1	2	3	4

LATIN SQUARE 2: Treatments identified with $A^{4}B^{2}C^{3}D^{3x}$

0	1	2	3	4	5	3x	3x+1	3x+2	3x+3	3x+4	3x+5
3x+2	3x+3	3x+4	3x+5	3x	3x+1	2	3	4	5	0	1
4	5	0	1	2	3	3x+4	3x+5	3x	3x+1	3x+2	3x+3
3x	3x-1	3x+2	3x+3	3x+4	3x+5	0	1	2	3	4	5
2	3	4	5	0	1	3x+2	3x+3	3x+4	3x+5	3x	3x+1
3x+4	3x+5	3x	3x-1	3x-2	3x+3	4	5	0	1	2	3
3x+3	3x+4	3x+5	3x	3x+1	3x-2	3	4	5	0	1	2
5	0	1	2	3	4	3x+5	3x	3x+1	3x+2	3x+3	3x+4
3x+1	3x+2	3x+3	3x+4	3x+5	3x	1	2	3	4	5	0
3	4	5	0	1	2	3x+3	3x+4	3x+5	3x	3x+1	3x+2
3x+5	3x	3x+1	3x+2	3x+3	3x+4	5	0	1	2	3	4
1	2	3	4	5	0	3x+1	3x+2	3x+3	3x+4	3x+5	3x

Using the complete confounding approach as outlined above, one can construct $\min[(2^2-1), (3-1)] = 2$ mutually orthogonal latin squares and no more as can easily be observed from the degrees of freedom table.

From the multiplication table of our ring R_{12}, we observe that 1, 5, 3x+1, 3x+2, 3x+4 and 3x+5 are the 6 non-zero divisors (i.e. elements with multiplicative inverses). Following Bose, Chakravarti and Knuth [1960], we consider simple automorphisms in R_{12} of the form:

$$\sigma(r) = r^* r$$

where r^* is a given fixed element having a multiplicative inverse (because only these elements are capable of producing automorphisms of R_{12}). Let now our aim be to produce two orthomorphisms which in turn will produce an O(12, 2) set. For this purpose consider the automorphisms:

$$I(r) = r$$

$$\sigma(r) = r^* \cdot r \ .$$

Now $I-\sigma$ implies the condition that r in the equation $[r^* \cdot r - r] = c$ has a unique solution for every c of R_{12}. In our setting this means that $r(r^* - 1) = c$ has a unique solution, i.e., $r(r^* + 5) = c$ has a unique solution which in turn implies that $(r^* + 5)^{-1}$ exists in R_{12}.

Substituting in the values of r^* we see that:

$[1 + 5]^{-1}$ does not exist in R_{12}

$[5 + 5]^{-1}$ " " " " "

$[(3x+1) + 5]^{-1}$ does not exist in R_{12}

$$[(3x+2)+5]^{-1} = [3x+1]^{-1} \text{ exists in } R_{12}$$

$$[(3x+4)+5]^{-1} \text{ does not exist in } R_{12}$$

$$[(3x+5)+5]^{-1} = [3x+4]^{-1} \text{ exists in } R_{12} .$$

Hence we have obtained two pairs of orthomorphisms namely:

$$I(r) = r \qquad I(r) = r$$

and

$$\alpha_1(r) = (3x+2)r \qquad \alpha_2(r) = (3x+5)r .$$

The O(12, 2) set presented above using complete confounding corresponds to the first pair of maps. It may be easily shown that simple maps of the type $\alpha(r) = r^* \cdot r$ lead to O(12, 2) sets or in general to an O(n, a) set, where $a = \min (p_1^{n_1}-1, p_2^{n_2}-1, \ldots, p_k^{n_k}-1)$ and $n = \prod_{i=1}^{k} p_i^{n_i}$ so that the complete confounding approach is equivalent to the construction of a set of a orthomorphisms.

VIII. Some Remarks on "Orthomorphism" Construction of O(n,t) Sets

In 1959 and 1960 E. T. Parker showed by a combination of classification of cases (with considerable elimination of isomorphic repetition accessible to cut down on computer time, followed by computer runs) that there are obtainable only 5 orthogonal latin squares of order 12, all restricted to be copies of the non-cyclic Abelian group with latin squares related by row permutations. (The researchers cited call this the method of orthomorphisms. Parker considers this a method, but only based on freaks of luck; further, Parker feels that "orthomorphism" admits no precise definition.)

Parker made another finding, also by hard classification of cases followed by computer runs, which Marshall Hall feels is more important than that cited above. No pair of order-12 orthogonal squares of the type mentioned can be extended to a complete set of any sort; i.e., further orthogonal latin squares are allowed to be completely general.

What might be obtainable for orthogonal squares of order 20 in like fashion, row-permuting the non-cyclic Abelian group of order 20 is an interesting matter for speculation — conceivably one might even produce a complete set (19 orthogonal squares equivalent to a plane). Knuth and Parker discussed the problem about 1963, and concluded that exhaustive search is out of the question; still a fortunate sample of cases might produce an attractive result.

In 1960 Parker looked at the row-permutation ("orthomorphism" of Bose and Mesner [1959]) approach for the group of order 15, and proved by Hasse-Minkowski invariants that a complete set could not be so obtained. He dropped

further work; but some persistence could quite possibly yield such as five orthogonal squares of order 15 .

A hybrid attack on order 15 or 20 might be undertaken by an ambitious investigator. (The facts for order 12 mentioned above rule out chances here.) One might produce sets of orthogonal latin squares of row-permuted group type, using automorphisms of the group latin squares to eliminate — or, that failing, reduce — isomorphic repetition. It would not be shrewd to program a computer to produce all transversals of a group latin square, for running time and output would be excessive; then for any hint of efficiency it would be necessary to turn about and do a reduction on the computer output. After a set of row-permuted latin squares (possibly exhaustive for order 15, but almost certainly only a sample for order 20 , large enough that computer searching would require realistic amounts of time, one might proceed with the next step. Produce all transversals of the set of orthogonal squares by computer, then fit these together in all possible ways (again by computer) to form orthogonal mates of the given set of orthogonal latin squares. Unlike Parker's assertion above about complete sets of order-10 squares, there is no known argument implying impossibility of producing 14 orthogonal latin squares of order 15 by this hybrid attack.

It is well-known that removing one line from the plane, usually called the line of infinity, the remaining $n^2 + n$ lines can be arranged into $2n$ lines passing through two points at infinity which are arbitrary up to notation and coordinatization of the plane, and n^2-n lines belonging to $n-1$ mutually orthogonal latin squares. If the line at infinity is chosen to be a secant and there are $2n$ lines, the lines pass through the two points of the oval such that each of the $n-1$ latin squares consists of $\frac{n}{2}$ secants and $\frac{n}{2}$ non-intersectors passing through each of the $n-1$ points at infinity other than the points of the oval.

Using the described method, it was assumed that a plane of order 10 exists. Under this assumtion 21 lines could be exhibited arbitrarily up to notation. Out of these lines one was taken to be the line of infinity and the remaining 20 used to coordinatize the plane. Then by trial and error twenty more lines were found which formed two orthogonal latin squares. The method used to construct these squares differs from the one described in literature.

Unfortunately no more squares could be found using this [illegible] computer established that the two squares did not yield an additional mutually orthogonal mate. Clearly it could happen that the choice of the first two was unfortunate. The same method was applied to the plane of order 12. Here the trial and error method failed to produce even two orthogonal squares. It may be worthwhile to remark that the construction of the plane and consequently the search for orthogonal latin squares does not require the assumption that the oval consists of the maximum number of points $n + 2$. However, if the plane does not include an oval consisting of $n + 2$ points the lines could not be classified into two categories only and this complicates the construction of the plane. Let us illustrate the method in the case n equals 10. It is easy to show that in this case the oval must consist of at least 6 points. However, the case of an oval of 6 points would be ignored since in this case every quadrangle would have to have collinear diagonals. On the other hand, a plane of order ten must be a non-Desarguessian and hence must contain

a non-degenerate quadrangle with noncollinear diagonals. Suppose that the plane contains a quadrangle with noncollinear diagonals and suppose that the plane contains an oval consisting of seven points then the 104 points of the plane which do not belong to the oval could be classified into three categories:

(i) points lying on 3 secants, 1 tangent 7 nonintersectors

(ii) " " " 2 " 3 " 6 "

(iii) " " " 1 " 5 " 5 "

Let us name the number of points in each category by x, y, z respectively. Clearly $x + y + z = 104$.

Counting the intersections of the secants and the tangents we get the further equations:

$$3x + y = 105$$

$$3y + 10z = 525$$

The unique solutions of this system of equations are $x = 20$, $y = 45$, $z = 39$. One could start the construction of the plane under the present assumption and investigate the possibilities of obtaining orthogonal latin squares in this way.

X. Code Construction of O(n,t) Sets

Given an n-symbol alphabet, e.g., 1,2,...,n, and a set of k-tuples of the n symbols, we denote the set of all k-tuples by $C_{k,n}$. This set may be thought of as a vector space or as a k-dimensional hypercube with edges of length n. Any subset of $C_{k,n}$ is denoted as a block code with a block length of k. The elements of the subset are denoted as code words. The number of symbols by which any two code words differ is called the Hamming distance. If any pair of code words in the subset differs by a Hamming difference of at least r, the block code is called a distance r code. A distance r code is called an (r-1)/2-error correcting code because fewer than (r-1)/2 changes leaves the word closer to its original form than to any other code word in the subset. For similar reasons, a distance r code has also been designated as an (r-1)-error-dectecting code.

In an interesting paper, Golomb and Posner [1964] discuss the relationships between a subset of n^2 code words and an O(n,t) set and relate these to ideas developed from a consideration of a set of n^2 super rooks of power t on the n^{t+2} chessboard such that no two super rooks attack each other. The new concepts of rook domains and rook packing were found to be very useful in providing a geometrical view of the results.

Any subset of n^2 words from $C_{3,n}$ which forms a single-error-detecting code may be used to construct a latin square of order n as any pair of the triples differs by at least two symbols. Likewise, any subset of n^2 words

from $C_{t+2,n}$ with a Hamming distance of $t+1$ may be utilized to construct an $O(n,t)$ set. These results are embodied in the following theorem (from Golomb and Posner [1964]):

Theorem X.8.1 The following three concepts are equivalent:

i) an $O(n,t)$ set

ii) A set of n^2 nonattacking super rooks of power t on the n^{t+2} board, for even t, also the following, a set of n^2 super rooks of power $t/2$ on the n^{t+2} board such that no cell is attacked twice; that is, such that the rook domains are nonoverlapping.

iii) A distance $t+1$ code of block length $t+2$ with n^2 words from an n-symbol alphabet.

For those interested in code construction, reference may be made to Mann [1968] and Peterson [1961] and the literature citations therein. We shall merely illustrate the method of construction of an $O(n,t)$ set from n^2 words of length $t+2$ and Hamming distance $t+1$ through an example. Let $n = 3$ and $t = 2$. Then the $n^2 = 9$ code words with length 4 and Hamming distance 3 and the corresponding latin squares are:

					latin squares of order 3								
						0	1	2			0	1	2
0000	0111	0222		0	0	1	2		0	0	1	2
1012	1120	1201	to produce	1	1	2	0	and	1	2	0	1
2021	2102	2210		2	2	0	1		2	1	2	0

where the first symbol corresponds to row number, the second to column number,

the third to symbols in the first latin square, and the fourth to symbols in the second latin square. The two latin squares form an $O(3,2)$ set. Note that any pair of the quadruples differs in at least three symbols.

The analogy of the above with many of the concepts from fractional replication and orthogonal arrays is immediately apparent. The equivalences of many of the results in these fields need to be systematically noted much in the same manner that Golomb and Posner [1964] note various equivalences among $O(n,t)$ sets, error-correcting codes, and n^2 nonattacking rooks on an n^{t+2} chessboard.

XI. Pairwise Balanced Design Construction of O(n,t) Sets

Central to the constructions of orthogonal latin squares of Bose and Shrikhande [1959] and of Parker [1959, 1960] is the following which might be called a "folk theorem," being credited to no specific investigator: From a set of t orthogonal latin squares of order n one may produce a set of n^2 ordered (t+2)-tuples on n^2 symbols such that each pair of distinct positions contains each ordered pair of symbols (exactly once); the converse construction can also be carried out. (Some, such as Bose, prefer to call the set of (t+2)-tuples an orthogonal array.) There is nothing difficult to prove in this construction. Two arbitrary positions in the (t+2)-tuples are identified with row and column indices in matrices, and each other position with entries in one of the matrices. The equivalence between orthogonality of latin squares and the conditions on the (t+2)-tuples is then fairly apparent.

Parker [1960] contributed the following to the construction of orthogonal latin squares. If there exists a pair of orthogonal latin squares of order m, then there exists a pair of orthogonal latin squares of order 3m + 1.

Let the 3m + 1 symbols be $X_1,\ldots,X_m$ and the residue classes (mod 2m + 1). Form the (latin square) array

$$\begin{array}{cccc} X_1 & 0 & 1 & -1 \\ 0 & X_1 & -1 & 1 \\ 1 & -1 & X_1 & 0 \\ -1 & 1 & 0 & X_1 \end{array} .$$

for each i, $1 \leq i \leq m$, each row is one of the ordered quadruples. In turn, the list of quadruples is built up by adding each integer (mod $2m + 1$) to all four positions at once, the X_i symbols being unchanged by the addition. The set of $4m(2m + 1)$ ordered quadruples just described contains in each pair of distinct positions exactly one occurrence of each ordered pair made up of an X_i and a residue class in either order, and of each ordered pair made up of two distinct residue classes. The required set of ordered quadruples is completed by adjoining: i) all ordered quadruples (j,j,j,j), $j = 0, \ldots, 2m$; ii) a set of ordered quadruples on the X_i symbols corresponding to a pair of orthogonal latin squares of order m guaranteed by the hypothesis to exist.

Bose and Shrikhande (1959, published 1959 and 1960 partly in a 3-author paper with Parker) developed a sequence of constructive theorems which led in steps to disproof of Euler's conjecture for all orders $4t + 2 > 6$. Their central theorem given here does not exhaust their methods, but virtually all their results rest on this theorem. We begin with a definition. A pairwise balanced design, $PB(n; k_1, \ldots, k_t)$ is a collection of subsets of a set of n elements, each subset having number of elements one of the k_i, and such that each pair of distinct elements in the set of n occurs in a unique subset of the PB. (Note: unlike in balanced incomplete block designs, the subsets of a PB are not restricted to have equal numbers of elements.) Now for the main theorem of Bose and Shrikhande. If a $PB(n; k_1, \ldots, k_t)$ exists, and for each i, $1 \leq i \leq t$, a set of m orthogonal latin squares of order k_i exists, then a set of $m - 1$ orthogonal latin squares of order n exists. Loosely speaking, the sets of ordered tuples for each subset

of the PB are constructed and these fit together to form a set of ordered tuples for the full set of n elements. The decrease from m to m-1 orthogonal latin squares occurs because in fitting the pieces together to form the large set of ordered tuples, it is necessary that each set of ordered tuples formed from a subset of the PB include each $(i,i,\ldots,i)$, where i ranges over the elements of that subset. (It is sufficient that this condition be fulfilled in the construction. Thus the theorem might be stated in slightly stronger form: "If ... $1 \leq i \leq t$, a set of m orthogonal latin squares of order k_i with a transversal exists, then a set of m orthogonal latin squares of order n exists.") Now for a more nearly formal version of the proof. If there exists a set of m orthogonal latin squares of order n, then there exists a set of the appropriate sort of n^2 ordered (m+1)-tuples with each symbol repeated in an (m+1)-tuple m+1 times. (The condition mentioned is satisfied with (m+2)-tuples if the set of orthogonal latin squares has a transversal.) One need simply put together the ordered tuples on each subset of the PB in turn, subject to the important condition that within each subset of the PB, each tuple of repetitions of each symbol be included. Carrying this out on the alphabet of the symbols in each subset of the PB, one has the construction for the set of orthogonal latin squares in the conclusion: each ordered tuple of a repeated symbol among the n is used only once.

A representative and very interesting example (Bose and Shrikhande informed Parker that this was the first case of disproof of Euler's conjecture produced in their joint work at a blackboard) yields 5 mutually orthogonal latin squares of order 50 via the PB construction. One forms the affine plane of

order 7, then adjoins exactly one ideal point on each line of one class of parallel lines. This yields a PB(50; 8,7). Since there exist 6 orthogonal latin squares of each order 8 and 7, there exist 6 - 1 = 5 orthogonal latin squares of order 50.

There is a limitation on the Bose-Shrikhande PB construction. Aside from trivial PB designs, having a single subset of all elements, any PB has a subset with at most one more element than the square root of the number of elements in the large set. Thus other techniques are requisite to produce more than $\sqrt{n}$ orthogonal latin squares of order not a prime-power.

~~755~~
751

XII. Product Composition of O(n,t) Sets

About 70 years ago, for the first time, Tarry [1899] in his half-page note asserted that if there exists an O(a, 2) set and if there exists an O(b, 2) set then there exists an O(ab, 2) set. He exhibited the following O(12, 2) set, by composing two O(3, 2) and O(4, 2) sets, to demonstrate the truth of his assertion. Note that in the following square the set of first integers belong to one latin square and the set of second integers belong to the second latin square. No more description is given by Tarry.

2-3	1-1	3-2	8-12	7-10	4-11	11-6	10-4	12-5	5-9	4-7	6-8
3-1	2-2	1-3	9-10	8-11	7-12	12-4	11-5	10-6	6-7	5-8	4-9
1-2	3-3	2-1	7-11	0-12	8-10	10-5	12-6	11-4	4-8	6-9	5-7
11-9	10-7	12-8	5-6	4-4	6-5	2-12	1-10	3-11	8-3	7-1	9-2
12-7	11-8	10-9	6-4	5-5	4-6	3-10	2-11	1-12	9-1	8-2	7-3
10-8	12-9	11-7	4-5	6-6	5-4	1-11	3-12	2-10	7-2	9-3	8-1
5-12	4-10	6-11	11-3	10-1	12-2	8-9	7-7	9-8	2-6	1-4	3-5
6-10	5-11	4-12	12-1	11-2	10-3	9-7	8-8	7-9	3-4	2-5	1-6
4-11	6-12	5-10	10-2	12-3	11-1	7-8	9-9	8-7	1-5	3-6	2-4
8-6	7-4	9-5	2-9	1-7	3-8	5-3	4-1	6-2	11-12	10-10	12-11
9-4	8-5	7-6	3-7	2-8	1-9	6-1	5-2	4-3	12-10	11-11	10-12
7-5	9-6	8-4	1-8	3-9	2-7	4-2	6-3	5-1	10-11	12-12	11-10

Tarry did not observe any generalization of his method. Perhaps this was due to the fact that he, like so many other researchers, was only concerned with sets of type O(n, 2). Probably he was not aware of the existence of a larger set.

About 23 years later MacNeish [1922] demonstrated:

1) The existence and a construction of an $O(n, n-1)$ set for n a prime or prime power integer.

2) A generalization of Tarry's procedure viz., if there exists an $O(a, r)$ set and if there exists an $O(b, r)$ set then there exists an $O(ab, r)$ set.

3) By a successive application of 1) and 2) he showed that if $n = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_t^{\alpha_t}$ is the prime-power decomposition of n then there exists an $O(n, r)$ set where $r = \min\{p_i^{\alpha_i} - 1,\ i = 1, 2, \ldots, t\}$.

MacNeish could not embed his $O(n, r)$ set generated in 3) in a larger set. This unsuccessful attempt, reinforced by Euler's conjecture, led MacNeish to prove (erroneously) geometrically that $O(n, t)$ sets do not exist for $t > r$, and therefore, as a confirmation of Euler's conjecture. The preceding argument of MacNeish is known as MacNeish's conjecture in the literature. By constructing an $O(21, 3)$ set Parker [1959] gave a counter example to MacNeish's conjecture. Later Bose, Shrikhande, and Parker [1960] completely demolished Euler's conjecture except for $n = 6$. It should be mentioned that MacNeish's conjecture has not been totally disproved yet. For instance, no one as yet as far as we know, has constructed an $O(15, 4)$ set (an $O(15, 3)$ set is given in section V for the first time) or an $O(20, 3)$ set. We believe that MacNeish should be given substantial credit for his non-erroneous contributions. It is to be regretted that MacNeish is often cited in the literature only for his false conjecture.

Even though Tarry and MacNeish did not attach any name to their procedure, it is not difficult to see that it is the method of Kronecker product of matrices. Therefore, we can state, more formally, their results as follows:

Theorem (Tarry-MacNeish). If $\{A_1, A_2, \ldots, A_r\}$ is an $O(n,r)$ set and if $\{B_1, B_2, \ldots, B_r\}$ is an $O(m,r)$ set, then $\{A_1 \otimes B_1, A_2 \otimes B_2, \ldots, A_r \otimes B_r\}$, where $\otimes$ denotes the Kronecker product operation of matrices, is an $O(nm, r)$ set.

The preceding arguments clearly support the choice of the title for this section and is in contrast to the choice of the name for the procedure given in section XIII.

XIII. Sum Composition Construction of O(n,t) Sets

XIII.1. Introduction. Perhaps one of the most useful techniques for the construction of combinatorial systems is the method of composition. To mention some, here are few well-known examples: 1) If there exists a set of t orthogonal latin squares of order n_1 and if there exists a set of t orthogonal latin squares of order n_2, then there exists a set of t orthogonal latin squares of order n_1n_2. 2) If there are Steiner triple systems of order v_1 and v_2, there is a Steiner triple system of order $v = v_1v_2$. 3) If H_1 and H_2 are two Hadamard matrices of order n_1 and n_2 respectively, then the Kronecker product of H_1 and H_2 is a Hadamard matrix of order n_1n_2. 4) If Room squares of order n_1 and n_2 exist, then a Room square of order n_1n_2 exists. 5) If BIB (v_1,k,λ_1) and BIB (v_2,k,λ_2) exist and if $f(\lambda_2v_2^2) \geq k$, then BIB $(v_1v_2,k,\lambda_1\lambda_2)$ exists where $f(\lambda_2v_2^2)$ denotes the maximum number of constraints which are possible in an orthogonal array of size $\lambda_2v_2^2$, with v_2 levels, strength 2, and index λ_2. 6) As a final example, the existence of orthogonal arrays $(\lambda_iv_i^t, q_i, v_i, t)$, $i = 1,2,\ldots,r$ implies the existence of the orthogonal array $(\lambda v^t, q, v, t)$, where $\lambda = \lambda_1\lambda_2\ldots\lambda_r$, $v = v_1v_2\ldots v_r$ and $q = \min(q_1, q_2, \ldots, q_r)$.

The reader will note that each of the above examples involved a product type composition. The method that we will describe utilizes a sum type composition, by means of which one can possibly construct sets of orthogonal latin squares for all $n \geq 10$.

XIII.2. Definitions. In the sequel by an $O(n,t)$ set we mean a set of t mutually orthogonal latin squares of order n.

a) A transversal (directrix) of a latin square L of order n on an n-set Σ is a collection of n cells such that the entries of these cells exhaust the set Σ and every row and column of L is represented in this collection. Two transversals are said to be parallel if they have no cell in common.

b) A collection of n cells is said to form a common transversal for an $O(n, t)$ set if the collection is a transversal for each of these t latin squares. Similarly, two common transversals are said to be parallel if they have no cell in common.

Example. The underlined and parenthesized cells form two parallel common transversals for the following $O(4, 2)$ set.

$$\left\{\begin{array}{cccc} 1 & 2 & (3) & \underline{4} \\ (2) & \underline{1} & 4 & 3 \\ \underline{3} & (4) & 1 & 2 \\ 4 & 3 & \underline{2} & (1) \end{array}\right. \qquad \left.\begin{array}{cccc} 1 & 2 & (3) & \underline{4} \\ (4) & \underline{3} & 2 & 1 \\ \underline{2} & (1) & 4 & 3 \\ 3 & 4 & \underline{1} & (2) \end{array}\right\}$$

XIII.3. Composing Two Latin Squares of Order n_1 and n_2

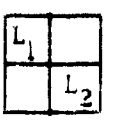
A very natural question in the theory of latin squares is the following: Given two latin squares L_1 and L_2 of order n_1 and n_2 $(n_1 \geq n_2)$ respectively. In how many ways can one compose L_1 and L_2 in order to obtain a latin square L_3 of order m, where m is a function of n_1 and n_2 only? This question can be partially answered as follows. First, it is well-known that the Kronecker product $L_3 = L_1 \times L_2$ is a latin square of order $m = n_1 n_2$ irrespective of the combinatorial structure of L_1 and L_2. Secondly, we show that if L_1 has a certain combinatorial structure, then one can construct a latin square L of

order $n = n_1 + n_2$. Naturally enough we call this procedure a "method of sum composition".

Even though our method of sum composition does not work for all pairs of latin squares, it has an immediate application in the construction of orthogonal latin squares including those of order $4t + 2$, $t \geq 2$. We emphasize that the combinatorial structure of orthogonal latin squares constructed by the method of sum composition is completely different from those of known orthogonal latin squares in the literature. Therefore, it is worthwhile to study these squares for the purpose of constructing new finite projective planes.

We shall now describe the method of "sum composition". Let L_1 and L_2 be two latin squares of order n_1 and n_2, $n_1 \geq n_2$, on two non-intersecting sets $\Sigma_1 = \{a_1, a_2, \ldots, a_{n_1}\}$ and $\Sigma_2 = \{b_1, b_2, \ldots, b_{n_2}\}$ respectively. If L_1 has n_2 parallel transversals then we can compose L_1 with L_2 to obtain a latin square L of order $n = n_1 + n_2$. Note that for any pair (n_1, n_2), there exists L_1 and L_2 with the above requirement, except for $(2,1)$, $(2,2)$, $(6,5)$ and $(6,6)$.

To produce L put L_1 and L_2 in the upper left and lower right corner respectively. Call the resulting square C_1, which looks as follows:

$C_1 =$ 

Name the n_2 transversals of L_1 in any manner from 1 to n_2. Now fill the cell $(i, n_1 + k)$, $k = 1,2,\ldots,n_2$, with that element of transversal k which appears in row i, $i = 1,2,\ldots,n_1$. Fill also the cell $(n_1 + k, j)$, $k = 1,2,\ldots,n_2$,

with that element of transversal k which appears in column j, $j = 1,2,\ldots,n_1$. Call the resulting square C_2. Now every entry of C_2 is occupied with an element either from Σ_1 or Σ_2, but C_2 is obviously not a latin square on $\Sigma_1 \cup \Sigma_2$. However, if we replace each of the n_1 entries of transversal k with b_k, it is easily verified that the resulting square which we call L is a latin square of order n on $\Sigma_1 \cup \Sigma_2$.

The procedure described for filling the first n_1 entries of the row (column) $n_1 + k$ with the corresponding entries of transversal k is, naturally enough, called <u>the projection of transversal k on the first n_1 entries of row (column) $n_1 + k$</u>.

We shall now elucidate the above procedure via an example. Let $\Sigma_1 = \{1,2,3,4,5\}$, $\Sigma_2 = \{6,7,8\}$,

$$L_1 = \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 3 & 4 \\ 4 & 5 & 1 & 2 & 3 \\ 3 & 4 & 5 & 1 & 2 \\ 2 & 3 & 4 & 5 & 1 \end{array} \quad \text{and} \quad L_2 = \begin{array}{ccc} 6 & 7 & 8 \\ 7 & 8 & 6 \\ 8 & 6 & 7 \end{array}.$$

Note that the cells on the same curve in L_1 form a transversal.

$$C_1 = \begin{array}{|ccccc|c|} \hline 1 & 2 & 3 & 4 & 5 & \\ 5 & 1 & 2 & 3 & 4 & \\ 4 & 5 & 1 & 2 & 3 & \\ 3 & 4 & 5 & 1 & 2 & \\ 2 & 3 & 4 & 5 & 1 & \\ \hline & & & & & 6\ 7\ 8 \\ & & & & & 7\ 8\ 6 \\ & & & & & 8\ 6\ 7 \\ \hline \end{array} \quad \text{and} \quad C_2 \begin{array}{|ccccc|ccc|} \hline 1 & 2 & 3 & 4 & 5 & 1 & 2 & 3 \\ 5 & 1 & 2 & 3 & 4 & 4 & 5 & 1 \\ 4 & 5 & 1 & 2 & 3 & 2 & 3 & 4 \\ 3 & 4 & 5 & 1 & 2 & 5 & 1 & 2 \\ 2 & 3 & 4 & 5 & 1 & 3 & 4 & 5 \\ \hline 1 & 3 & 5 & 2 & 4 & 6 & 7 & 8 \\ 5 & 2 & 4 & 1 & 3 & 7 & 8 & 6 \\ 4 & 1 & 3 & 5 & 2 & 8 & 6 & 7 \\ \hline \end{array}$$

And finally

$L =$

6	7	8	4	5	1	2	3
7	8	2	3	6	4	5	1
8	5	1	6	7	2	3	4
3	4	6	7	8	5	1	2
2	6	7	8	1	3	4	5
1	3	5	2	4	6	7	8
5	2	4	1	3	7	8	6
4	1	3	5	2	8	6	7

which is a latin square of order 8 on $\Sigma_1 \cup \Sigma_2 = \{1,2,\ldots,8\}$.

Remark. Note that it is by no means required that the projection of transversals on the rows and columns should have the same ordering. Indeed, for the fixed set of ordered n_2 transversals, we have $n_2!$ choices of projections on columns and $n_2!$ choices of projections on the rows. Hence we can generate at least $(n_2!)^2$ different latin squares of order $n = n_1 + n_2$ composing L_1 and L_2.

XIII.4. Construction of $O(n,2)$ Sets by Method of Sum Composition. In order to construct an $O(n,2)$ set for $n = n_1 + n_2$, we require that $n_1 \geq 2n_2$ and there should exist an $O(n_2,2)$ set, and an $O(n_1,2)$ set with $2n_2$ parallel transversals. It is easy to show that any $n \geq 10$ can be decomposed in at least one way into $n_1 + n_2$ which fulfill the above requirements. We now present two theorems which state that for certain n one can construct an $O(n,2)$ set by the method of sum composition.

Theorem XIII.4.1. Let $n_1 = p^\alpha \geq 7$ for any odd prime p and positive integer α, excluding $n_1 = 13$. Then there exists an $O(n,2)$ set which can be constructed by composition of two $O(n_1,2)$ and $O(n_2,2)$ sets for $n_2 = (n_1 - 1)/2$ and $n = n_1 + n_2$.

We shall first give the method of construction and then a proof that the constructed set is an $O(n,2)$ set.

Construction. Let $B(r)$ be the $n_1 \times n_1$ square with element $r\alpha + \alpha_j$ in its (i,j) cell, α_i, α_j, $0 \neq r$ in $GF(n_1)$, $i,j = 1,2,\ldots,n_1$. Then it is easy to see that $\{B(1), B(x), B(y)\}$, $y = x^{-1}$, $x \neq 1$, is an $O(n_1, 3)$ set. Consider the n_1 cells in $B(1)$ with $\alpha_i + \alpha_j = k$ a fixed element in $GF(n_1)$. Then the corresponding cells in $B(x)$ and $B(y)$ form a common transversal for the set $\{B(x), B(y)\}$. Name this common transversal by k. It is then obvious that two common transversals k_1 and k_2, $k_1 \neq k_2$ are parallel and hence $\{B(x), B(y)\}$ has n_1 common parallel transversals. Now let (A_1, A_2) be any $O(n_2, 2)$ set, which always exists, on a set Ω non-intersecting with $GF(n_1)$. For any λ in $GF(n_1)$ we can find $(n_1 - 1)/2$ pairs of distinct elements belonging to $GF(n_1)$ such that the sum of the two elements of each pair is equal to λ. Let $\{S\}$ and $\{T\}$ denote the collection of the first and the second elements of these $(n_1 - 1)/2$ pairs respectively. Note that for a fixed λ the set $\{S\}$ can be constructed in $(n_1 - 1)(n_1 - 3)\ldots 1$ distinct ways. Now fix λ and let L_1 denote any of the $(n_2!)^2$ latin squares that can be generated by the sum composition of $L(x)$ and A_1 using transversals determined by the n_2 elements of $\{S\}$. Let L_2 be the latin square derived from the composition of $L(y)$ and A_2 using the n_2 transversals determined by the elements of $\{T\}$ and the following projection rule: Project transversals t_i, $i = 1,2,\ldots,n_2$ on the row (column) which upon superposition of L_2 on L_1 this row (column) should coincide with the row (column) stemmed from the transversal $\lambda - t_i$. Shortly we shall prove that $\{L_1, L_2\}$ forms on $O(n,2)$ set.

The preceding arguments shows that $\{L_1, L_2\}$ can be constructed non-isomorphically in at least $(n_1-3)(n_2!)^2[n_1(n_1-1)(n_1-3)\ldots 1]$ ways. For instance in the case of $n_1 = 7$, there is at least 12096 non-isomorphic pairs of orthogonal latin squares of order 10. Therefore, Euler has been wrong in his conjecture by a very wide margin.

Note that we can construct infinitely many pairs of orthogonal latin squares of order $4t + 2$ by the method of theorem XIII.4.1. For $p \equiv 7 \bmod 8$ and α odd $p^{\alpha} = (8t + 5)/3)$. Hence $n_1 + n_2 = 4t + 2$.

Proof: The constructional procedure clearly reveals that:

A. L_1 and L_2 are latin squares of order n on $GF(n_1) \cup \Omega$.

B. Upon superposition of L_1 on L_2 the following are true:

b_1. Every element of Ω appears with every other element of Ω.

b_2. Every element of Ω appears with every element of $GF(n_1)$.

b_3. Every element of $GF(n_1)$ appears with every element of Ω.

Therefore, all we have to prove is that every element of $GF(n_1)$ appears with every other element of $GF(n_1)$. To prove this recall that $B(x)$ is orthogonal to $B(y)$. However, since we removed the n_2 transversals from $B(x)$ determined by the n_2 elements of $\{S\}$ and n_2 transversals from $B(y)$ determined by the n_2 elements of $\{T\}$ therefore the following $2n_2n_1$ pairs have been lost.

$(xa_i + a_j, ya_i + a_j)$ with $a_i + a_j = \gamma$ for any $\gamma \in GF(n_1)$, $\gamma \neq \lambda$.

We claim that the given projection rules guarantee the capture of these lost pairs by the $2n_2n_1$ bordered cells. To show this note that the superposition of the

projected transversal s from $B(x)$ on the projected transversal $t = \lambda - s$ from $B(y)$ will capture the n_1 pairs.

$$(x\alpha_i + \alpha_j,\ y\alpha_i + \alpha_j) \text{ with } \alpha_i + \alpha_j = k = [y(\lambda - s) + s]/(1+y)$$

If these transversals have been projected on row border and n_1 pairs

$$(x\alpha_i + \alpha_j,\ y\alpha_i + \alpha_j) \text{ with } \alpha_i + \alpha_j = k = [s(y-1) + (s-\lambda)(x-1)]/(y-x)$$

if these transversals have been projected on column border. Now because $k + k' = \lambda$ and if $s_1 \neq s_2$ then $k_1 \neq k_2$ and $k_1' \neq k_2'$ hence the $2n_2n_1$ pairs which have been resulted from the projection of transversals determined by $\{S\}$ and $\{T\}$ will jointly capture the $2n_2n_1$ lost pairs and thus a proof.

We shall now clarify the above constructional procedure by an example.

Example. Let $n_1 = 7$, $GF(7) \quad \{0,1,2,\ldots,6\}$. Then for $x = 2$, $y = x^{-1} = 4$ we have

$\{B(1), B(2), B(4)\} =$

0 1 2 3 4 5 6	0 1 2 3 4 5 6	0 1 2 3 4 5 6
1 2 3 4 5 6 0	2 3 4 5 6 0 1	4 5 6 0 1 2 3
2 3 4 5 6 0 1	4 5 6 0 1 2 3	1 2 3 4 5 6 0
3 4 5 6 0 1 2	6 0 1 2 3 4 5	5 6 0 1 2 3 4
4 5 6 0 1 2 3	1 2 3 4 5 6 0	2 3 4 5 6 0 1
5 6 0 1 2 3 4	3 4 5 6 0 1 2	6 0 1 2 3 4 5
6 0 1 2 3 4 5	5 6 0 1 2 3 4	3 4 5 6 0 1 2

For $n_2 = (n_1-1)/2 = 3$ let $\Omega_2 = \{7,8,9\}$ and

$$\{A_1, A_2\} = \begin{matrix} 7 & 8 & 9 \\ 8 & 9 & 7 \\ 9 & 7 & 8 \end{matrix}, \quad \begin{matrix} 7 & 8 & 9 \\ 9 & 7 & 8 \\ 8 & 9 & 7 \end{matrix}.$$ Finally for $\lambda = 0$, $\{S\} = \{1,2,3\}$ and $\{T\} = \{6,5,4\}$ we have $\{L_1, L_2\} =$

0	7	8	9	4	5	6	1	2	3	0	1	2	3	7	8	9	6	5	4
7	8	9	5	6	0	1	2	3	4	4	5	6	7	8	9	3	2	1	0
8	9	6	0	1	2	7	3	4	5	1	2	7	8	9	6	0	5	4	3
9	0	1	2	3	7	8	4	5	6	5	7	8	9	2	3	4	1	0	6
1	2	3	4	7	8	9	5	6	0	7	8	9	5	6	0	1	4	3	2
3	4	5	7	8	9	2	6	0	1	8	9	1	2	3	4	7	0	6	5
5	6	7	8	9	3	4	0	1	2	9	4	5	6	0	7	8	3	2	1
2	1	0	6	5	4	3	7	8	9	3	0	4	1	5	2	6	7	8	9
4	3	2	1	0	6	5	8	9	7	6	3	0	4	1	5	2	9	7	8
6	5	4	3	2	1	0	9	7	8	2	6	3	0	4	1	5	8	9	7

the reader can easily verify that $\{L_1, L_2\}$ is an $O(10,2)$ set.

Remarks.

1) The method of theorem XIII.4.1 fails for $n_1 = 13$ only because there is no $O(6,2)$ set. Otherwise, there will be no orthogonality contradiction on the other parts of L_1 and L_2 with their 6×6 lower right square missing.

2) In the case of $n_1 = 7$, if we let $\{S\} = \{0,1,3\}$ and $\{T\} = \{2,4,5\}$ then the requirement $y = x^{-1}$ is not necessary. However then we do not have a unified projection rule for the formation of L_2 as was provided for the case $y = x^{-1}$ by theorem XIII.4.1. To give the complete list of solutions let (a_1,a_2,a_3) and (b_1,b_2,b_3) be any two permutations of the set $\{8,9,10\}$. If we project transversals $(0,1,3)$ on the rows (a_1,a_2,a_3) and columns (b_1,b_2,b_3) in the formation of L_1, then the following table indicates what permutation of transversals $\{2,4,5\}$ should be projected on the rows (a_1,a_2,a_3) and columns (b_1,b_2,b_3) in the formation of L_2. Obviously these permutations will be a function of the pair (x,y).

Pair (x,y)	Rows a_1, a_2, a_3	Columns b_1, b_2, b_3
(2, 3)	4, 2, 5	4, 2, 5
(2, 3)	2, 5, 4	2, 5, 4
(2, 4)	2, 5, 4	4, 2, 5
(2, 5)	4, 2, 5	4, 2, 5
(2, 6)	2, 5, 4	2, 5, 4
(3, 4)	2, 5, 4	2, 5, 4
(3, 5)	2, 5, 4	4, 2, 5
(3, 5)	4, 2, 5	5, 4, 2
(3, 5)	4, 2, 5	2, 5, 4
(3, 5)	5, 4, 2	2, 5, 4
(3, 6)	4, 2, 5	2, 5, 4
(3, 6)	5, 4, 2	4, 2, 5
(4, 5)	2, 5, 4	2, 5, 4
(4, 6)	5 4, 2	4, 2, 5
(4, 6)	2, 5, 4	2, 5, 4
(4, 6)	5, 4, 2	5, 4, 2

(This table is by no means exhaustive.)

The reader may note that whenever $y = x^{-1}$ in the above table the given solution(s) are different from the one provided by the method of theorem XIII. 4. 1.

Thus we can conclude that any pair of orthogonal latin squares of order 7 based on the GF(7) can be composed with a pair of orthogonal latin squares of

order 3 and make a pair of orthogonal latin squares of order 10. In addition, since we have six choices for (a_1,a_2,a_3) and (b_1,b_2,b_3) hence from every line in the above table we can produce 36 non-isomorphic O(10,2) sets or $16 \times 36 = 576$ sets for the entire table. Since all these pairs are non-isomorphic with all previous pairs, produced by theorem XIII.4.1, thus by the method of sum composition one can at least produce 12,672 non-isomorphic O(10,2) sets.

We believe that for other values of n_1 there are sets of {S} and {T} together with proper projections which makes the restriction $y = x^{-1}$ unnecessary.

Theorem XIII.4.2. Let $n_1 = 2^\alpha \geq 8$ for any positive integer α. Then there exists an $O(n,2)$ set which can be constructed by composition of two $O(n_1,2)$ and $O(n_2,2)$ sets for $n_2 = n_1/2$ and $n = n_1 + n_2$.

We shall here give only the method of construction. A similar argument as in theorem XIII.4.1 will show that the constructed set is an O(n,2) set.

Construction. In a similar fashion as in theorem XIII.4.1 construct the set $\{B(1), B(x), B(y)\}$ over $GF(2^\alpha)$. Let also $\{A_1,A_2\}$ be any $O(n_2,2)$ set, which always exists, on a set Ω non-intersecting with $GF(2^\alpha)$. For any $\lambda \neq 0$ in $GF(2^\alpha)$ we can find $n_1/2$ pairs of distinct elements belonging to $GF(2^\alpha)$ such that the sum of the two elements of each pair is equal to λ. Let {S} and {T} denote the collection of the first and the second elements of these $n_1/2$ pairs respectively. Note that for a fixed λ the set {S} can be constructed in $n_1(n_1-2)(n_1-4)\ldots1$ distinct ways. Now form L_1 from the sum composition of B(x) and A_1 and L_2 from the sum composition of B(y) and A_2 using the same projection rule as given in theorem XIII.4.1. Now $\{L_1,L_2\}$ is an O(n,2) set.

Example. Let $n = 8$, $GF(8) = \{0,1,2,\ldots,7\}$ with the following addition (+) and multiplication (×) tables:

+	0	1	2	3	4	5	6	7
0	0	1	2	3	4	5	6	7
1	1	0	6	4	3	7	2	5
2	2	6	0	7	5	4	1	3
3	3	4	7	0	1	6	5	2
4	4	3	5	1	0	2	7	6
5	5	7	4	6	2	0	3	1
6	6	2	1	5	7	3	0	4
7	7	5	3	2	6	1	4	0

×	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7
2	0	2	3	4	5	6	7	1
3	0	3	4	5	6	7	1	2
4	0	4	5	6	7	1	2	3
5	0	5	6	7	1	2	3	4
6	0	6	7	1	2	3	4	5
7	0	7	1	2	3	4	5	6

Then for $x = 2$, $y = x^{-1} = 7$ we have

$\{B(1), B(2), B(7)\} =$

0	1	2	3	4	5	6	7
1	0	6	4	3	7	2	5
2	6	0	7	5	4	1	3
3	4	7	0	1	6	5	2
4	3	5	1	0	2	7	6
5	7	4	6	2	0	3	1
6	2	1	5	7	3	0	4
7	5	3	2	6	1	4	0

0	1	2	3	4	5	6	7
2	6	0	7	5	4	1	3
3	4	7	0	1	6	5	2
4	3	5	1	0	2	7	6
5	7	4	6	2	0	3	1
6	2	1	5	7	3	0	4
7	5	3	2	6	1	4	0
1	0	6	4	3	7	2	5

0	1	2	3	4	5	6	7
7	5	3	2	6	1	4	0
1	0	6	4	3	7	2	5
2	6	0	7	5	4	1	3
3	4	7	0	1	6	5	2
4	3	5	1	0	2	7	6
5	7	4	6	2	0	3	1
6	2	1	5	7	3	0	4

For $n_2 = n_1/2 = 4$ let $\Omega = \{A, B, C, D\}$ and

$$\{A_1, A_2\} = \begin{array}{cccc} A & B & C & D \\ B & A & D & C \\ C & D & A & B \\ D & C & B & A \end{array} \qquad \begin{array}{cccc} A & B & C & D \\ D & C & B & A \\ B & A & D & C \\ C & D & A & B \end{array}$$

For $S = \{0,1,\ldots,4\}$ and $\{2,3\}$ we have $\{L_1, L_2\} =$

A	B	2	C	D	5	6	7	0	1	3	4
B	A	0	D	C	4	1	3	6	2	5	7
3	4	A	0	1	D	B	C	7	5	2	6
C	D	5	A	B	2	7	6	1	0	4	3
D	C	4	B	A	0	3	1	2	6	7	5
6	2	D	5	7	A	C	B	3	4	0	1
7	5	B	2	6	C	A	D	4	3	1	0
1	0	C	4	3	B	D	A	5	7	6	2
0	6	7	1	2	3	4	5	A	B	C	D
2	1	3	6	0	7	5	4	B	A	D	C
4	7	6	3	5	1	0	2	C	D	A	B
5	3	1	7	4	6	2	0	D	C	B	A

0	1	D	3	4	A	C	B	5	7	6	2
7	5	C	2	6	B	D	A	0	1	3	4
D	C	6	B	A	7	2	5	3	4	0	1
2	6	B	7	5	C	A	D	1	0	4	3
3	4	A	0	1	D	B	C	7	5	2	6
A	B	5	C	D	2	7	6	4	3	1	0
C	D	4	A	B	0	3	1	6	2	5	7
B	A	1	D	C	3	0	4	2	6	7	5
4	2	7	6	3	5	1	0	A	B	C	D
6	3	0	4	2	1	5	7	D	C	B	A
5	0	3	1	7	4	6	2	B	A	D	C
1	7	2	5	0	6	4	3	C	D	A	B

which is an $O(12,2)$ set.

Discussion. The necessary requirements for the construction of an $O(n,t)$ set, $n = n_1 + n_2$, $t < n_2$, by the method of sum composition are: The existence of an $O(n_1,t)$ set, $n_1 \geq tn_2$, with at least tn_2 common parallel transversals, and an $O(n_2,t)$ set. These conditions are obviously satisfied whenever n_1 and n_2 are prime powers.

While for some values of n there exists only a unique decomposition fulfilling the above requirements, for infinitely many other values of n there are abundant such decompositions.

It seems that if there exists an $O(n_2,2)$ set and if $n = n_1 + n_2$, $n_1 \geq 2n_2$ then one can construct an $O(n,2)$ set by the method of sum composition if n_1 is a prime power. To support this observation and shed some more light on the method of sum composition we present in subsequent pages some highlights of the results which we hope to complete and submit for publication shortly.

In the following for each given decomposition of n we exhibit an $O(n,2)$ set which has been derived by the method of sum composition. We shall represent the pairs in a form that the curious reader can easily reconstruct the original sets. Hereafter the notation $L_1 \perp L_2$ means that L_1 is orthogonal to L_2.

1) $12 = 9 + 3$

A	B	C	4	5	6	7	8	9	1	2	3
B	C	A	1	2	3	4	5	6	9	7	8
C	A	B	7	8	9	1	2	3	5	6	4
2	3	1	5	6	4	A	B	C	8	9	7
8	9	7	2	3	1	B	C	A	4	5	6
5	6	4	8	9	7	C	A	B	3	1	2
3	1	2	A	B	C	9	7	8	6	4	5
9	7	8	B	C	A	6	4	5	2	3	1
6	4	5	C	A	B	3	1	2	7	8	9
1	5	9	6	7	2	8	3	4	A	B	C
7	2	6	3	4	8	5	9	1	B	C	A
4	8	3	9	1	5	2	6	7	C	A	B

1	2	3	4	5	6	A	B	C	8	9	7
9	7	8	3	1	2	B	C	A	6	4	5
5	6	4	8	9	7	C	A	B	1	2	3
6	4	5	A	B	C	3	1	2	7	8	9
2	3	1	B	C	A	8	9	7	5	6	4
7	8	9	C	A	B	4	5	6	3	1	2
A	B	C	2	3	1	5	6	4	9	7	8
B	C	A	7	8	9	1	2	3	4	5	6
C	A	B	6	4	5	9	7	8	2	3	1
4	9	2	5	7	3	6	8	1	A	B	C
3	5	7	1	6	8	2	4	9	C	A	B
8	1	6	9	2	4	7	3	5	B	C	A

2) $14 = 11 + 3$, the only decomposition which fulfills the necessary requirements.

A	B	C	3	4	5	6	7	8	9	10	0	1	2
B	C	9	10	0	1	2	3	4	5	A	6	7	8
C	4	5	6	7	8	9	10	0	A	B	1	2	3
10	0	1	2	3	4	5	6	A	B	C	7	8	9
6	7	8	9	10	0	1	A	B	C	5	2	3	4
2	3	4	5	6	7	A	B	C	0	1	8	9	10
9	10	0	1	2	A	B	C	6	7	8	3	4	5
5	6	7	8	A	B	C	1	2	3	4	9	10	0
1	2	3	A	B	C	7	8	9	10	0	4	5	6
8	9	A	B	C	2	3	4	5	6	7	10	0	1
4	A	B	C	8	9	10	0	1	2	3	5	6	7
0	5	10	4	9	3	8	2	7	1	6	A	B	C
7	1	6	0	5	10	4	9	3	8	2	B	C	A
3	8	2	7	1	6	0	5	10	4	9	C	A	B

0	1	2	3	4	5	6	7	A	B	C	9	10	8
8	9	10	0	1	2	5	A	B	C	7	5	6	4
5	6	7	8	1	10	A	B	C	3	4	1	2	0
2	3	4	5	6	A	B	C	10	0	1	8	9	7
10	0	1	2	A	B	C	6	7	8	9	4	5	3
7	8	9	A	B	C	2	3	4	5	6	0	1	10
4	5	A	B	C	9	10	0	1	2	3	7	8	6
1	A	B	C	5	6	7	8	9	10	0	3	4	2
A	B	C	1	2	3	4	5	6	7	8	10	0	9
B	C	8	9	10	0	1	2	3	4	A	6	7	5
C	4	5	6	7	8	9	10	0	A	B	2	3	1
6	10	3	7	0	4	8	1	5	9	2	A	B	C
3	7	0	4	8	1	5	9	2	6	10	C	A	B
9	2	6	10	3	7	0	4	8	1	5	B	C	A

3) $15 = 12 + 3$, $15 = 11 + 4$ are the only decompositions which fulfill the necessary requirements. However, we consider here the latter decomposition since we can utilize the properties of Galois field GF(11).

A	B	C	D	4	5	6	7	8	9	10	0	1	2	3
B	C	D	5	6	7	8	9	10	0	A	1	2	3	4
C	D	6	7	8	9	10	0	1	A	B	2	3	4	5
D	7	8	9	10	0	1	2	A	B	C	3	4	5	6
8	9	10	0	1	2	3	A	B	C	D	4	5	6	0
10	0	1	2	3	4	A	B	C	D	9	5	6	0	1
1	2	3	4	5	A	B	C	D	10	0	6	7	1	2
3	4	5	6	A	B	C	D	0	1	2	7	8	2	3
5	6	7	A	B	C	D	1	2	3	4	8	9	3	4
7	8	A	B	C	D	2	3	4	5	6	9	10	4	5
9	A	B	C	D	3	4	5	6	7	9	10	0	5	6
0	10	9	8	7	6	5	4	3	2	1	A	B	C	D
2	1	0	10	9	8	7	6	5	4	3	B	A	D	C
4	3	2	1	0	10	9	8	7	6	5	C	D	A	B
6	5	4	3	2	1	0	10	9	8	7	D	C	B	A

0	1	2	3	4	5	A	B	C	D	10	8	6	9	7
6	7	8	9	10	A	B	C	D	4	5	2	0	3	1
1	2	3	4	A	B	C	D	9	10	0	7	5	8	6
7	8	9	A	B	C	D	3	4	5	6	1	10	2	0
2	3	A	B	C	D	8	9	10	0	1	6	4	7	5
8	A	B	C	D	2	3	4	5	6	7	0	9	1	10
A	B	C	D	7	8	9	10	0	1	2	5	3	6	4
B	C	D	1	2	3	4	5	6	7	A	10	8	0	9
C	D	6	7	8	9	10	0	1	A	B	4	2	5	3
D	0	1	2	3	4	5	6	A	B	C	9	7	10	8
5	6	7	8	9	10	0	A	B	C	D	3	1	4	2
9	4	10	5	0	6	1	7	2	8	3	A	B	C	D
10	5	0	7	1	7	2	8	3	9	4	D	C	B	A
3	9	4	10	5	0	6	1	7	2	8	B	A	D	C
4	10	5	0	6	1	7	2	8	3	9	C	D	A	B

4) $17 = 13 + 4$ and $17 = 12 + 5$ are the only decompositions which fulfill the necessary requirements.

The following pair is derived through the first decomposition.

A	B	C	D	4	5	6	7	8	9	10	11	12	0	1	2	3	0	1	2	3	4	5	6	7	A	B	C	D	12	9	8	11	10
B	C	D	8	9	10	11	12	0	1	2	3	A	4	5	6	7	8	9	10	11	12	0	1	A	B	C	D	6	7	3	2	5	4
C	D	12	0	1	2	3	4	5	6	7	A	B	8	9	10	11	2	4	5	6	7	8	A	B	C	D	0	1	2	10	9	12	7
D	3	4	5	6	7	8	9	10	11	A	B	C	12	0	1	2	11	12	0	1	2	A	B	C	D	7	8	9	10	4	3	6	5
7	8	9	10	11	12	0	1	2	A	B	C	D	3	4	5	6	6	7	8	9	A	B	C	D	1	2	3	4	5	11	10	0	12
12	0	1	2	3	4	5	6	A	B	C	D	11	7	8	9	10	1	2	3	A	B	C	D	8	9	10	11	12	0	5	4	7	6
4	5	6	7	8	9	10	A	B	C	D	2	3	11	12	0	1	9	10	A	B	C	D	2	3	4	5	6	7	8	12	11	1	0
9	10	11	12	0	1	A	B	C	D	6	7	8	2	3	4	5	4	A	B	C	D	9	10	11	12	0	1	2	3	6	5	8	7
1	2	3	4	5	A	B	C	D	10	11	12	0	6	7	8	9	A	B	C	D	3	4	5	6	7	8	9	10	11	0	12	2	1
6	7	8	9	A	B	C	D	1	2	3	4	5	10	11	12	0	B	C	D	10	11	12	0	1	2	3	4	5	A	7	6	9	8
11	12	0	A	B	C	D	5	6	7	8	9	10	1	2	3	4	C	D	4	5	6	7	8	9	10	11	12	A	B	1	0	3	2
3	4	A	B	C	D	9	10	11	12	0	1	2	5	6	7	8	D	11	12	0	1	2	3	4	5	6	A	B	C	8	7	10	9
8	A	B	C	D	0	1	2	3	4	5	6	7	9	10	11	12	5	6	7	8	9	10	11	12	0	A	B	C	D	2	1	4	3
0	9	5	1	10	6	2	11	7	3	12	8	4	A	B	C	D	7	0	6	12	5	11	4	10	3	9	2	8	1	A	B	C	D
5	1	10	6	2	11	7	3	12	8	4	0	9	B	A	D	C	12	5	11	4	10	3	9	2	8	1	7	0	6	D	C	B	A
10	6	2	11	7	3	12	8	4	0	9	5	1	C	D	A	B	10	3	9	2	8	1	7	0	6	12	5	11	4	B	A	D	C
2	11	7	3	12	8	4	0	9	5	1	10	6	D	C	B	A	3	8	1	7	0	6	12	5	11	4	10	3	9	C	D	A	B

5) We do not know whether there exists either an $O(14,2)$ set with 8 common parallel transversals or an $O(15,2)$ set with 6 common parallel transversals. Therefore the only decomposition of 18 which fulfills the necessary requirements is $18 = 13 + 5$. The following pair is constructed through this decomposition.

A	B	C	D	E	5	6	7	8	9	10	11	12	0	1	2	3	4	0	1	2	3	4	5	6	A	B	C	D	E	12	7	9	10	11	8
B	C	D	E	6	7	8	9	10	11	12	0	A	1	2	3	4	5	7	8	9	10	11	12	A	B	C	D	E	5	6	0	2	3	4	1
C	D	E	7	8	9	10	11	12	0	1	A	B	2	3	4	5	6	1	2	3	4	5	A	B	C	D	E	11	12	0	6	8	9	10	7
D	E	8	9	10	11	12	0	1	2	A	B	C	3	4	5	6	7	8	9	10	11	A	B	C	D	E	4	5	6	7	12	1	2	3	0
E	9	10	11	12	0	1	2	3	A	B	C	D	4	5	6	7	8	2	3	4	A	B	C	D	E	10	11	12	0	1	5	7	8	9	6
10	11	12	0	1	2	3	4	A	B	C	D	E	5	6	7	8	9	9	10	A	B	C	D	E	3	4	5	6	7	8	11	0	1	2	12
12	0	1	2	3	4	5	A	B	C	D	E	11	6	7	8	9	10	3	A	B	C	D	E	9	10	11	12	0	1	2	4	6	7	8	5
1	2	3	4	5	6	A	B	C	D	E	12	0	7	8	9	10	11	A	B	C	D	E	2	3	4	5	6	7	8	9	10	12	0	1	11
3	4	5	6	7	A	B	C	D	E	0	1	2	8	9	10	11	12	B	C	D	E	8	9	10	11	12	0	1	2	A	3	5	6	7	4
5	6	7	8	A	B	C	D	E	1	2	3	4	9	10	11	12	0	C	D	E	1	2	3	4	5	6	7	8	A	B	9	11	12	0	10
7	8	9	A	B	C	D	E	2	3	4	5	6	10	11	12	0	1	D	E	7	8	9	10	11	12	0	1	A	B	C	2	4	5	6	3
9	10	A	B	C	D	E	3	4	5	6	7	8	11	12	0	1	2	E	0	1	2	3	4	5	6	7	A	B	C	D	8	10	11	12	9
11	A	B	C	D	E	4	5	6	7	8	9	10	12	0	1	2	3	6	7	8	9	10	11	12	0	A	B	C	D	E	1	3	4	5	2
0	12	11	10	9	8	7	6	5	4	3	2	1	A	B	C	D	E	4	11	5	12	6	0	7	1	8	2	9	3	10	A	B	C	D	E
2	1	0	12	11	10	9	8	7	6	5	4	3	B	C	D	E	A	12	6	0	7	1	8	2	9	3	10	4	11	5	E	A	B	C	D
4	3	2	1	0	12	11	10	9	8	7	6	5	C	D	E	A	B	10	4	11	5	12	6	0	7	1	8	2	9	3	D	E	A	B	C
6	5	4	3	2	1	0	12	11	10	9	8	7	D	E	A	B	C	5	12	6	0	7	1	8	2	9	3	10	4	11	C	D	E	A	B
8	7	6	5	4	3	2	1	0	12	11	10	9	E	A	B	C	D	11	5	12	6	0	7	1	8	2	9	3	10	4	B	C	D	E	A

6) We do not know whether there exists either an O(18,2) set with 8 common parallel transversals or an O(15,2) set with 14 common parallel transversals. Therefore the only decomposition of 22 which fulfill the necessary requirements are $22 = 19 + 3$ and $22 = 17 + 5$.

a: $22 = 19 + 3$,

A	B	C	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	0	1	2
B	C	5	6	7	8	9	10	11	12	13	14	15	16	17	18	0	1	A	2	3	4
C	7	8	9	10	11	12	13	14	15	16	17	18	0	1	2	3	A	B	4	5	6
9	10	11	12	13	14	15	16	17	18	0	1	2	3	4	5	A	B	C	6	7	8
12	13	14	15	16	17	18	0	1	2	3	4	5	6	7	A	B	C	11	8	9	10
15	16	17	18	0	1	2	3	4	5	6	7	8	9	A	B	C	13	14	10	11	12
18	0	1	2	3	4	5	6	7	8	9	10	11	A	B	C	15	16	17	12	13	14
2	3	4	5	6	7	8	9	10	11	12	13	A	B	C	17	18	0	1	14	15	16
5	6	7	8	9	10	11	12	13	14	15	A	B	C	0	1	2	3	4	16	17	18
8	9	10	11	12	13	14	15	16	17	A	B	C	2	3	4	5	6	7	18	0	1
11	12	13	14	15	16	17	18	0	A	B	C	4	5	6	7	8	9	10	1	2	3
14	15	16	17	18	0	1	2	A	B	C	6	7	8	9	10	11	12	13	3	4	5
17	18	0	1	2	3	4	A	B	C	8	9	10	11	12	13	14	15	16	5	6	7
1	2	3	4	5	6	A	B	C	10	11	12	13	14	15	16	17	18	0	7	8	9
4	5	6	7	8	A	B	C	12	13	14	15	16	17	18	0	1	2	3	9	10	11
7	8	9	10	A	B	C	14	15	16	17	18	0	1	2	3	4	5	6	11	12	13
10	11	12	A	B	C	16	17	18	0	1	2	3	4	5	6	7	8	9	13	14	15
13	14	A	B	C	18	0	1	2	3	4	5	6	7	8	9	10	11	12	15	16	17
16	A	B	C	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	17	18	0
0	17	15	13	11	9	7	5	3	1	18	16	14	12	10	8	6	4	2	A	B	C
3	1	18	16	14	12	10	8	6	4	2	0	17	15	13	11	9	7	5	B	C	A
B	4	2	0	17	15	13	11	9	7	5	3	1	18	16	14	12	10	8	C	A	B

0	1	2	3	4	5	6	7	8	9	10	11	12	A	B	C	16	17	18	14	15	13
13	14	15	16	17	18	0	1	2	3	4	5	A	B	C	9	10	11	12	7	8	6
7	8	9	10	11	12	13	14	15	16	17	A	B	C	2	3	4	5	6	0	1	18
1	2	3	4	5	6	7	8	9	10	A	B	C	14	15	16	17	18	0	12	13	11
14	15	16	17	18	0	1	2	3	A	B	C	7	8	9	10	11	12	13	5	6	4
8	9	10	11	12	13	14	15	A	B	C	0	1	2	3	4	5	6	7	17	18	16
2	3	4	5	6	7	8	A	B	C	12	13	14	15	16	17	18	0	1	10	11	9
15	16	17	18	0	1	A	B	C	5	6	7	8	9	10	11	12	13	14	3	4	2
9	10	11	12	13	A	B	C	17	18	0	1	2	3	4	5	6	7	8	15	16	14
3	4	5	6	A	B	C	10	11	12	13	14	15	16	17	18	0	1	2	8	9	7
16	17	18	A	B	C	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	0
10	11	A	B	C	15	16	17	18	0	1	2	3	4	5	6	7	8	9	13	14	12
4	A	B	C	8	9	10	11	12	13	14	15	16	17	18	0	1	2	3	6	7	5
A	B	C	1	2	3	4	5	6	7	B	9	10	11	12	13	14	15	16	18	0	17
B	C	13	14	15	16	17	18	0	1	2	3	4	5	6	7	8	9	A	11	12	10
C	6	7	8	9	10	11	12	13	14	15	16	17	18	0	1	2	A	B	4	5	3
18	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	A	B	C	16	17	15
12	13	14	15	16	17	18	0	1	2	3	4	5	6	7	A	B	C	11	9	10	8
6	7	8	9	10	11	12	13	14	15	16	17	18	0	A	B	C	4	5	2	3	1
11	18	6	13	1	8	15	3	10	17	5	12	0	7	14	2	9	16	4	A	B	C
5	12	0	7	14	2	9	16	4	11	18	6	13	1	8	15	3	10	17	C	A	B
17	5	12	0	7	14	2	9	16	4	11	18	6	13	1	8	15	3	10	B	C	A

b: $22 = 17 + 5$,

A	B	C	D	E	5	6	7	8	9	10	11	12	13	14	15	16	0	1	2	3	4
B	C	D	E	6	7	8	9	10	11	12	13	14	15	16	0	A	1	2	3	4	5
C	D	E	7	8	9	10	11	12	13	14	15	16	0	1	A	B	2	3	4	5	6
D	E	8	9	10	11	12	13	14	15	16	0	1	2	A	B	C	3	4	5	6	7
E	9	10	11	12	13	14	15	16	0	1	2	3	A	B	C	D	4	5	6	7	8
10	11	12	13	14	15	16	0	1	2	3	4	A	B	C	D	E	5	6	7	8	9
12	13	14	15	16	0	1	2	3	4	5	A	B	C	D	E	11	6	7	8	9	10
14	15	16	0	1	2	3	4	5	6	A	B	C	D	E	12	13	7	8	9	10	11
16	0	1	2	3	4	5	6	7	A	B	C	D	E	13	14	15	8	9	10	11	12
1	2	3	4	5	6	7	8	A	B	C	D	E	14	15	16	0	9	10	11	12	13
3	4	5	6	7	8	9	A	B	C	D	E	15	16	0	1	2	10	11	12	13	14
5	6	7	8	9	10	A	B	C	D	E	16	0	1	2	3	4	11	12	13	14	15
7	8	9	10	11	A	B	C	D	E	0	1	2	3	4	5	6	12	13	14	15	16
9	10	11	12	A	B	C	D	E	1	2	3	4	5	6	7	8	13	14	15	16	0
11	12	13	A	B	C	D	E	2	3	4	5	6	7	8	9	10	14	15	16	0	1
13	14	A	B	C	D	E	3	4	5	6	7	8	9	10	11	12	15	16	0	1	2
15	A	B	C	D	E	4	5	6	7	8	9	10	11	12	13	14	16	0	1	2	3
0	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	A	B	C	D	E
2	1	0	16	15	14	13	12	11	10	9	8	7	6	5	4	3	B	C	D	E	A
4	3	2	1	0	16	15	14	13	12	11	10	9	8	7	6	5	C	D	E	A	B
6	5	4	3	2	1	0	16	15	14	13	12	11	10	9	8	7	D	E	A	B	C
8	7	6	5	4	3	2	1	0	16	15	14	13	12	11	10	9	E	A	B	C	D

0	1	2	3	4	5	6	7	8	A	B	C	D	E	14	15	16	10	11	12	13	9
9	10	11	12	13	14	15	16	A	B	C	D	E	5	6	7	8	1	2	3	4	0
1	2	3	4	5	6	7	A	B	C	D	E	13	14	15	16	0	9	10	11	12	8
10	11	12	13	14	15	A	B	C	D	E	4	5	6	7	8	9	0	1	2	3	16
2	3	4	5	6	A	B	C	D	E	12	13	14	15	16	0	1	8	9	10	11	7
11	12	13	14	A	B	C	D	E	3	4	5	6	7	8	9	10	16	0	1	2	15
3	4	5	A	B	C	D	E	11	12	13	14	15	16	0	1	2	7	8	9	10	6
12	13	A	B	C	D	E	2	3	4	5	6	7	8	9	10	11	15	16	0	1	14
4	A	B	C	D	E	10	11	12	13	14	15	16	0	1	2	3	6	7	8	9	5
A	B	C	D	E	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	0	13
B	C	D	E	9	10	11	12	13	14	15	16	0	1	2	3	A	5	6	7	8	4
C	D	E	0	1	2	3	4	5	6	7	8	9	10	11	A	B	13	14	15	16	12
D	E	8	9	10	11	12	13	14	15	16	0	1	2	A	B	C	4	5	6	7	3
E	16	0	1	2	3	4	5	6	7	8	9	10	A	B	C	D	12	13	14	15	11
7	8	9	10	11	12	13	14	15	16	0	1	A	B	C	D	E	3	4	5	6	2
16	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	15	11	12	13	14	10
8	9	10	11	12	13	14	15	16	0	A	B	C	D	E	6	7	2	3	4	5	1
5	14	6	15	7	16	8	0	9	1	10	2	11	3	12	4	13	A	B	C	D	E
14	6	15	7	16	8	0	9	1	10	2	11	3	12	4	13	5	E	A	B	C	D
6	15	7	16	8	0	9	1	10	2	11	3	12	4	13	5	14	D	E	A	B	C
15	7	16	8	0	9	1	10	2	11	3	12	4	13	5	14	6	C	D	E	A	B
13	5	14	6	15	7	16	8	0	9	1	10	2	11	3	12	4	B	C	D	E	A

XIV. Computer Construction of O(10, t) Sets

In about fifteen years the effectiveness of computers in searching for orthogonal sets of latin squares of order ten has increased strikingly. Still the problem is so large that there seems to be little reason for optimism that the order ten problem can be completed by computers. More precisely, if (as most conversant with the problem consider quite plausible) no orthogonal triple of orthogonal latin squares of order ten exists, then the number of cases to consider seems too large for an exhaustive proof by computer to be achievable. The number of latin squares of order ten is astronomical.

About 1953 Paige and Tompkins [195≡] programmed SWAC to search for squares orthogonal to a fixed latin square of order ten. A few hours of running produced no orthogonal square, and was regarded as a bit of experimental evidence for the truth of Euler's conjecture. Calculations based on the progress made in the search led to the extrapolation that over fifty million years of computer time would be required to search for all squares orthogonal to a latin square of order ten put into SWAC intially. (At about the same time a similar program was written and similar results obtained with MANIAC at Los Alamos; this attempt has not been reported in print.)

In 1959, after Euler's conjecture had been disproved for all orders $4t+2 > 6$, Parker programmed UNIVAC 1206 to search for squares orthogonal to a latin square of order ten. The running time was sharply less than for SWAC or MANIAC, about thirty minutes for the majority of latin squares. This was accomplished by generating and storing all transversals of the input latin square,

then searching for all ways to form latin squares from the list of transversals. (A transversal, or directrix, is a set of cells of a latin square, one in each row, one in each column, and one containing each digit.) The striking gain in speed over the earlier efforts occurred largely because the number of transversals of a typical latin square of order ten is roughly 850, much less than 10! ; and, of course, the search was several levels deep. (SWAC and MANIAC were programmed to build up starts of latin squares toward orthogonal mates by filling in cells to form rows.)

There were two main outcomes from considerable running of Parker's 1206 program: 1) Orthogonal triples of order ten latin squares are not numerous; more precisely, only a small fraction, if any, order ten squares extend to triples. Some 400 latin squares were run. Some were random, some were computer output fed back as input and hence known to have an orthogonal mate, and some were considered interesting candidates for intuitive reasons by Parker and others. Not once did an exhaustive search for orthogonal mates of an input latin square include a pair orthogonal to one another. Mild evidence may be claimed supporting the opinion that no order-ten orthogonal triple exists. 2) Of a computer-generated sample of 100 random latin squares of order ten (program by R. T. Ostrowski), 62 have orthogonal mates. Thus, unlike triples, order ten orthogonal pairs are quite common. Euler's intuition for order ten was not only wrong, but in this sense wrong by a large margin. It was this finding which tempted Parker for a time to believe that repeated runs of the program should have a good chance of producing at least a triple, but many failures dimmed this optimism.

In 1967 John W. Brown programmed IBM 7094 to decide whether an input latin square of order ten can be extended to an orthogonal triple. The running time was one half minute. Almost needless to say, transversals again were generated. Searching for patterns of transversals toward extension to a triple produced a speed gain over the previous program for orthogonal pairs. Brown endeavored to get every drop of speed from the machine. As before, hundreds of input order-ten latin squares produced no orthogonal triple.

XV. On the Equivalence of O(n,t) Sets With Other Combinatorial Systems

XV.0. Summary

In this section we have densely summarized some of the results obtained by author and at least fourteen others in order to demonstrate the importance of the theory of mutually orthogonal latin squares. We have shown that fourteen well-known and important combinatorial systems with certain parameters are actually equivalent to a set of mutually orthogonal latin squares. A schematic representation of these equivalences has been demonstrated in four wheels which we have called "Fundamental Wheels of Combinatorial Mathematics".

XV.1. Introduction

The theory of mutually orthogonal latin squares owes its importance to the fact that many well-known combinatorial systems are actually equivalent to a set of mutually orthogonal latin squares; viz., finite projective plane, finite Euclidean plane, net, BIB, PBIB, orthogonal arrays, a set of mutually orthogonal matrices, error correcting codes, strongly regular graphs, complete graphs, a balanced set of t-restrictional lattice designs, difference sets, Hadamard matrices, and an arrangement of non attacking rooks on hyperdimensional chess board. These combinatorial systems are unquestionably potent and effective in all branches of combinatorial mathematics, and in particular, in the construction of experimental designs. Therefore, a statement that the theory of mutually orthogonal latin squares is perhaps the most important theory in the field of experiment designs is not in the least exaggerated as far as this author is concerned.

Our purpose in this section is to demonstrate the relation of a set of mutually orthogonal latin squares with the above mentioned combinatorial systems. We shall present the essence of the known results available only in scattered literature in one theorem which we consider to be a "fundamental theorem of combinatorial mathematics". For the definitions of these combinatorial systems and the proof of the forthcoming theorem see the list of references given at the end of this paper.

XV. 2. Notation

For the sake of conciseness we introduce the following notations:

0) $O(n,t)$ denotes a set of t mutually orthogonal latin squares of order n.

1) $MOM(n,t)$ denotes a set of t mutually orthogonal $n \times n$ matrices.

2) $OA(n,t)$ denotes a set of orthogonal arrays of size n^2, depth t, n levels, and strength 2.

3) $Net(n,t)$ denotes a net of order n and degree t.

4) $Code(n,r,t;m)$ denotes a set of n code words each of length r such that any two code words are at least at Hamming distance $\geq t$ on an m-set Σ with m distinct elements. We remind the reader that such a code is also called $(t-1)$-error detecting code or $(t-1)/2$-error correcting code because such a code is capable of detecting up to $t-1$ errors and correct up to $(t-1)/2$ errors in each transmitted code word.

5) $PBIB(b,v,r,k,\lambda_1,\lambda_2)$ denotes a partially balanced incomplete block design with b blocks each of size k, v treatments with r replication of each, and association indices λ_1 and λ_2.

6) SR-Graph (A) denotes the strongly regular graph with incidence matrix A .

7) Non #(n,t) denotes an arrangement of n mutually non attacking rooks on the t-dimensional $n \times n$ chess board.

8) PG(2,s) denotes a finite projective plane of order s (not necessarily Desarguessian).

9) $\mathcal{E}(2,s)$ denotes a finite Euclidean plane of order s .

10) BIB(b,v,r,k,λ) denotes a balanced incomplete block design with b blocks each of size k, v treatments with r replications of each, and association index λ .

11) K-Graph (A) denotes the complete graph with incidence matrix A .

12) DIF(v,k,λ) denotes a difference set with parameters v, k, and λ .

13) BLRL(s) denotes a balanced set of 1-restrictional lattice design for s treatments. Note that a 1-restrictional balanced lattice designs for simply a BIB design.

14) HAD(n) denotes a symmetric normalized Hadamard matrix of order n .

Hereafter we also adopt the following two notations:

i) A $\Longleftrightarrow$ B means A implies B and B implies A .

ii) A $\Longrightarrow$ B means A implies B . Whether or not B implies A is undecided.

2. The Result:

Theorem

(a) For any pair of positive integers n and t we have:

1) $O(n,t) \Longleftrightarrow MOM(n,t+2)$

2) $O(n,t) \Longleftrightarrow OA(n,t+2)$

3) $O(n,t) \Longleftrightarrow Net(n,t+2)$

4) $O(n,t) \Longleftrightarrow Code(n^2, t+2, t+1; n)$

5) $O(n,t) \Longleftrightarrow PBIB(n^2, n(t+2), n, t+2, 0, 1)$

6) $O(n,t) \Longleftrightarrow$ SR-Graph (A) where A is the incidence matrix associated with PBIB in 5).

7) $O(n,t) \Longleftrightarrow$ Non $\#(n^2, n^{t+2})$.

(b) If $t = n-1$ then also:

8) $O(n,n-1) \Longleftrightarrow PG(2,n)$

9) $O(n,n-1) \Longleftrightarrow C(2,n)$

10) $O(n,n-1) \Longleftrightarrow BIB(n^2+n+1, n^2+n+1, n+1, n+1, 1)$

11) $O(n,n-1) \Longleftrightarrow Code(n^2+n+1, n^2+n+1, 2n; 2)$

12) $O(n,n-1) \Longleftrightarrow$ K-Graph (A) where A is the incidence matrix associated with BIB in 10)

13) $O(n,n-1) \Longleftrightarrow DIF(n^2+n+1, n+1, 1)$.

(c) If $n = p^m$ where p is a prime and m is a positive integer then also the following:

14) $O(p^m, p^m-1) \Longleftrightarrow BLRL(p^m)$.

(d) If $n = 2r$ and $t = r-2$, $r \geq 3$ then the following are also true:

15) $O(2r,r-2) \Longrightarrow HAD(4r^2)$

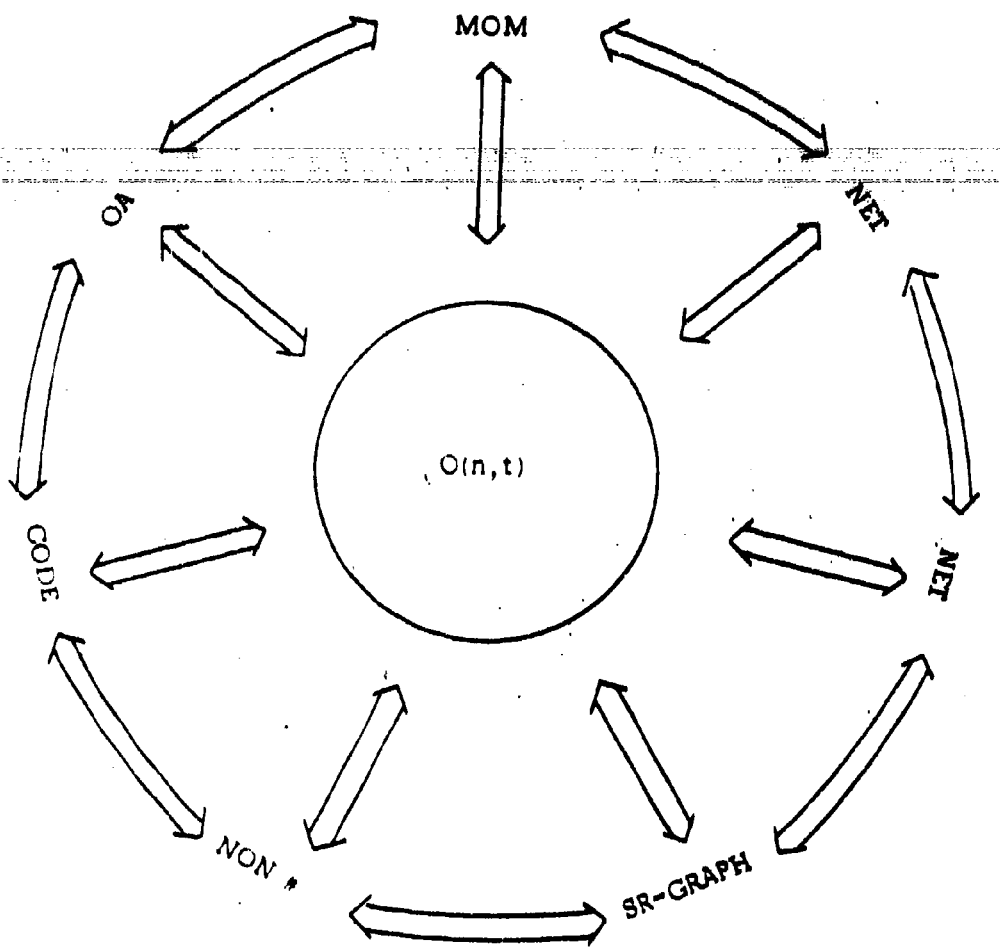
16) $O(2r,r-2) \Longrightarrow BIB(4r^2-1, 4r^2-1, 2r^2-1, 2r^2-1, r^2-1)$

17) $O(2r,r-2) \Longrightarrow Code(4r^2-1, 4r^2-1, 2r^2; 2)$

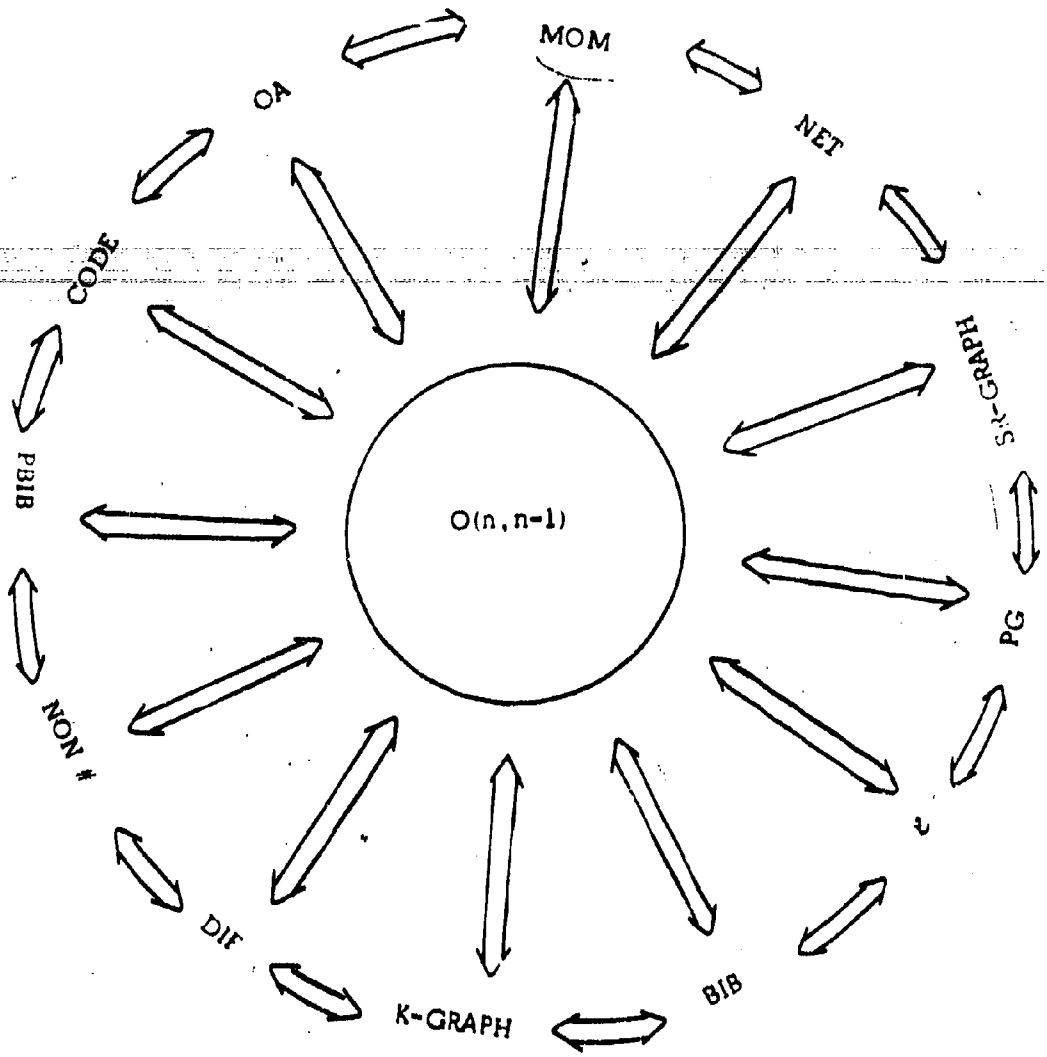
18) $O(2r,r-2) \Longrightarrow Code(8r^2, 4r^2, 2r^2; 2)$

19) $O(2r,r-2) \Longrightarrow DIF(4r^2-1, 2r^2-1, r^2-1)$.

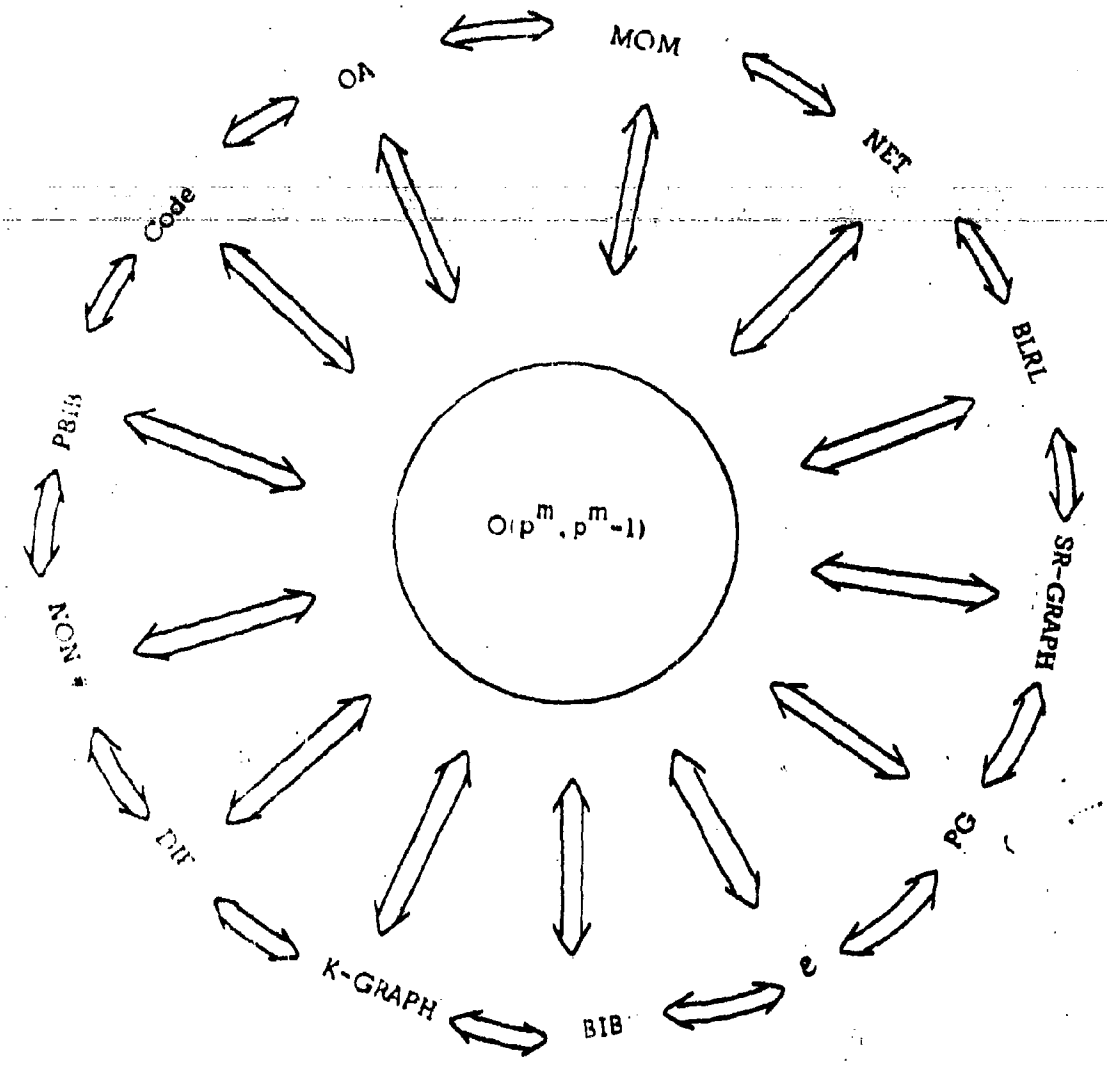
A complete schematic representation of this theorem can be demonstrated in four wheels which will be called "fundamental wheels of combinatorial mathematics". For the sake of compactness we shall omit the associated parameters with each system in these wheels except for $O(n,t)$. By knowing the values of n and t in the given $O(n,t)$ sets, then the reader can easily find the associated parameters with other systems in the wheels from the proper part of above theorem.



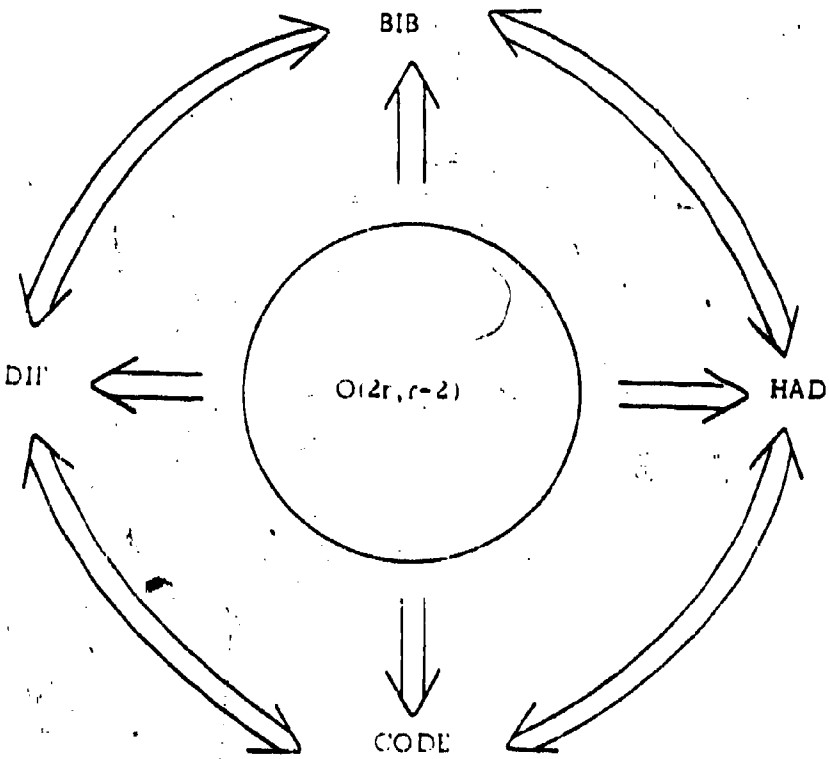
Wheel 1. For any positive integer n and t .



Wheel 2. For any positive integer n .



Wheel 4. For any prime p and positive integer m .



Wheel 4. For any positive integer $r \geq 3$.
(see also wheel 1)

XVI. Acknowledgements

This work was partially supported under the following research grants or contracts:

(i) Public Health Research Grant GM-05900, Cornell University

(ii) United States Army Contract Number DA-31-124-ARO-D-462, University of Wisconsin

(iii) National Science Foundation Grant GP-12881, Michigan State University

(iv) Office of Naval Research Contract N0014-A67-0305-0008, University of Illinois.

XVII. Literature Cited

1. Beckenbach, E. F. (ed.) [1964], Applied Combinatorial Mathematics, John Wiley and Sons, Inc., New York, London, Sydney.
2. Bose, R. C. [1939], On the application of the properties of Galois Fields to the problem of construction of Hyper-Graeco-Latin squares, Sankhyā 3: 323-338.
3. Bose, R. C. [1939], On the construction of balanced incomplete block designs, Annals of Eugenics 9: 343-399.
4. Bose, R. C. [1950], A note on orthogonal arrays (abstract), Annals of Mathematical Statistics 21: 304-305.
5. Bose, R. C. [1963], Strongly regular graphs, partial geometrics and partially balanced designs, Pacific Journal of Mathematics 13: 389-419.
6. Bose, R. C. Chakravarti, I. M., and Knuth, D. E. [1960], On methods of constructing sets of mutually orthogonal Latin squares using a computer, I. Technometrics 2: 507-516.
7. Bose, R. C. and Clatworthy, W. H. [1955], Some classes of partially balanced designs, Annals of Mathematical Statistics 26: 212-232.
8. Bose, R. C. and Shrikhande, S. S. [1959], On the falsity of Euler's conjecture about the non-existence of two orthogonal Latin squares of order 4t + 2, Proceedings of the National Academy of Sciences, U.S.A. 45: 734-737.

9. Bose, R. C. and Shrikhande, S. S. [1960], On the construction of pairwise orthogonal Latin squares and the falsity of a conjecture of Euler, Transactions of the American Mathematical Society 95: 191-209.

10. Bose, R. C., Shrikhande, S. S. and Parker, E. T. [1960], Further results on the construction of mutually orthogonal Latin squares and the falsity of Euler's conjecture, Canadian Journal of Mathematics 12: 189-203.

11. Bruck, R. H. [1951], Finite nets, I. Numerical invariants, Canadian Journal of Mathematics 3: 94-107.

12. Bruck, R. H. [1963], Finite nets, II. Uniqueness and embedding, Pacific Journal of Mathematics 13: 421-457.

13. Bruck, R. H. [1963], What is a loop? in "Studies in Modern Algebra" (editor, A. A. Albert), Math. Assoc. Amer. and Prentice Hall, Englewood Cliffs, New Jersey, pp. 59-99.

14. Dembowski, P. [1968], Finite Geometries, Springer-Verlag, Berlin, Heidelberg, New York.

15. Euler, L. [1782], Recherches sur une nouvelle espèce des quarrés magiques, Verh. Zeeuwsch Genoot. Wetenschappen, Vlissigen 9: 85-239.

16. Fisher, R. A. and Yates, F. [1957], Statistical Tables for Biological, Agricultural and Medical Research, 5th edition (first edition in 1938), Oliver and Boyd, Edinburgh and London.

17. Golomb, S. W. and Posner, E. C. [1964], Rook domains, Latin squares, affine planes, and error-distributing codes, IEEE Transactions on Information Theory IT-10: 196-208.

18. Hall, M. [1943], Projective planes, Transactions of the American Mathematical Society 54: 229-277.

19. Hall, M. Jr. [1967], Combinatorial Theory, Blaisdell Publishing Co., Waltham, Massachusetts, Toronto and London.

20. Hall, M. and Paige, L. J. [1955], Complete mappings of finite groups, Pacific Journal of Mathematics 5: 541-549.

21. Hedayat, A. [1968], On Singer 1-permutation, Unpublished paper, Biometrics Unit, Cornell University, Ithaca, New York.

22. Hedayat, A. [1969], On the theory of the existence, non-existence, and the construction of mutually orthogonal F-squares and latin squares, Ph. D. Thesis, Cornell University, Ithaca, New York.

23. Hedayat, A. [1969], On the equivalence of a set of mutually orthogonal latin squares with other combinatorial systems, Paper RM-237, Dept. of Statistics and Probability, Michigan State University, East Lansing, Michigan.

24. Hedayat, A. and Federer, W. T. [1969], An application of group theory to the existence and non-existence of orthogonal Latin squares, Biometrika 56(3):

25. Hedayat, A. and Federer, W. T. [1969], On the equivalence of Mann's group automorphism method of constructing an O(n, n-1) set and Raktoe's collineation method of constructing a balanced set of ℓ-restrictional prime-powered lattice designs, (unpublished paper), Biometrics Unit, Cornell University, Ithaca, New York.

26. Hedayat, A. and Seiden, E. [1969], On a method of sum composition of orthogonal latin squares. Paper No. RM-238, Dept. of Statistics and Probability, Michigan State University, East Lansing, Michigan.

27. Hedayat, A. And Seiden, E. [1969], Some contributions to the theory of F-squares. Unpublished paper, Cornell University and Michigan State University.

28. Johnson, D. M., Dulmadge, A. L., and Mendelsohn, N. S. [1961], Orthomorphisms of groups and orthogonal Latin Squares I., Canadian Journal of Mathematics 13: 356-372.

29. Kempthorne, O. [1952], The Design and Analysis of Experiments, John Wiley and Sons, Inc., New York.

30. Levi, F. W. [1942], Finite Geometrical Systems, University of Calcutta, Calcutta, India.

31. MacNeish, H. F. [1922], Euler's squares, Annals of Mathematics 23: 221-227.

32. Mann, H. B. [1942], The construction of orthogonal Latin squares, Annals of Mathematical Statistics 13: 418-423.

33. Mann, H. B. [1943], On the construction of sets of orthogonal Latin squares. The Annals of Mathematical Statistics 14: 401-414.

34. Mann, H. B. [1944], On orthogonal Latin squares, Bulletin of the American Mathematical Society 50: 249-257.

35. Mann, H. B. (ed.) [1968], Error Correcting Codes, John Wiley and Sons, Inc., New York, London, Sydney, Toronto, pp. xii + 231.

36. Paige, L. J. [1951], Complete mappings of finite groups, Pacific Journal of Mathematics 1: 111-116.

37. Paige, L. J. and Tompkins, C. B. [1960], The size of the 10 × 10 orthogonal latin square problem, Proceedings of the Tenth Symposium in Applied Mathematics of the American Mathematics Society 10: 71-83.

38. Parker, E. T. [1959], Construction of some sets of mutually orthogonal latin squares, Proceedings of the American Mathematics Society 10: 946-949.

39. Parker, E. T. [1962], On orthogonal latin squares, Proceedings of Symposia in Pure Mathematics, American Mathematics Society 6: 43-46.

40. Peterson, W. W. [1961], Error-Correcting Codes, John Wiley and Sons, Inc., New York.

41. Raktoe, B. L. [1967], Application of cyclic collineations to the construction of balanced f-restrictional prime powered lattice designs, Annals of Mathematical Statistics 38: 1127-1141.

42. Raktoe, B. L. [1969], Combining elements from finite fields in mixed factorials, The Annals of Mathematical Statistics 40: 498-504.

43. Raktoe, B. L. and Federer, W. T. [1969], On irregular fractions of an s^m factorial, Unpublished paper, Biometrics Unit, Cornell University, Ithaca, New York.

44. Ryser, H. J. [1963], Combinatorial Mathematics, The Carus Mathematical Monographs No. 14, Mathematical Association of America and John Wiley and Sons, Inc., New York.

45. Sade, A. [1958], Groupïdes orthogonaux, Publicationes Mathematicae, 5: 229-240.

46. Silverman, R. [1960], A metrization for power sets with applications to combinatorial analysis, Canadian Journal of Mathematics 12: 158-176.

47. Singer, J. [1960], A class of groups associated with Latin squares, The American Mathematical Monthly 67: 235-240.

48. Stevens, W. L. [1939], The completely orthogonalized Latin squares, Annals of Eugenics 9: 83-93.

49. Tarry, G. [1899], Sur le probleme d'Euler des n^2 officiers, L'Intermédiaire des Mathématiciens 6: 251-252.

ON CONFIDENCE LIMITS FOR THE PERFORMANCE OF A SYSTEM WHEN FEW FAILURES ARE ENCOUNTERED

Sam C. Saunders
Boeing Scientific Research Laboratories
Seattle, Washington

SUMMARY. In some situations encountered today components, or assemblies, are so reliable that no failures are observed within the time available for testing. This can pose a problem in both interpretation and analysis. We consider here the problem of determining lower confidence bounds on the reliability of a complex system, such as the Saturn 1-C, when each component is assumed to have an exponential life and different components have different multiplicities within the system. We discuss the assumptions necessary to obtain confidence limits using the likelihood of the data when only a few failures are encountered. The bounds resulting from two models are compared. The first model is Bayesian with uniform prior distribution of the failure rates. The second model regards the failure rates virtually as unknown constants. Here the argument is made that models of the first type are deficient in several regards in comparison with the second.

0. INTRODUCTION. The problem of determining the probability of successful operation of a large complex system when one has data only on the reliability of the components has, over the past decade, been the subject of many investigations. However, much of the literature was of a proprietary nature and was never published, for example, see [1], [2] and the references there.

Some of the studies, see [7] and [9], were based on an asymptotic theory for which the precision of the approximation is unknown. Currently, much of the analysis is based on Bayesian methods utilizing subjective prior assumptions, see [12] and [13].

Because estimating the probability of failure under some models requires that at least one failure be observed, the statistician may be placed in the uncomfortable position of having less confidence in his estimates of reliability when fewer failures obtain. Ultimately, when the system becomes near perfect and no failures are observed the statistician has no confidence if his procedures are necessarily based on failure analysis. In this unsatisfactory situation it is an understandable reaction of persons with good engineering judgment and statistical intuition, to form a distrust of statistical inference and its "numerologists," see [3]. Some recent surveys have been made to determine the most useful and applicable procedures for current needs. One of the most comprehensive is [8].

In this note we examine some statistical techniques which do not depend upon the sampling method yet are applicable when there is a paucity of observed failures among the components which have been tested. The archtypical situation for this study will be the Saturn 1-C and the data which was available prior to the first launch, as given in Table I.

The remainder of this article has been reproduced photographically from the author's manuscript.

1. THE BASIC MODEL

Consider a complex system designed to perform in a specified manner when all of its components, grouped in assemblies, are operating. At question is the confidence we can have that it will perform adequately for a specified time.

We shall consider the reliability of the system for a specified time t as being given by

$$h(\underline{\lambda},t) = \exp\left\{-\sum_{i=1}^{m} \lambda_i \omega_i t\right\} \tag{1.1}$$

where the ω_i are known weighting factors for the unknown parameters $\underline{\lambda} = (\lambda_1,\ldots,\lambda_m)$. The problem is obtaining a lower confidence bound on (1.1) from the limited amount of data about the λ_i for $i=1,\ldots,m$.

The particular form (1.1) can arise in several ways. The first we mention is that it is itself a lower bound on system reliability.

If the time until failure of the i^{th} component is exponential with unknown hazard rate λ_i and these m components are in a coherent system, then there exists a set of integers $q_1,\ldots,q_m$, where q_i is the multiplicity of the i^{th} component within the system of order $\sum_1^m q_i$, which can be used to obtain a lower bound on the reliability of the system at any time $t > 0$. This lower bound is $h(\underline{\lambda},t)$ as given in (1.1) with ω_i replaced by q_i.

For a proof of this result, see [4].

An assembly containing several components can malfunction by having different components fail, e.g., a pressure system can either rupture or leak. These separate ways of failure are called "failure modes" in the current terminology, however, they correspond to malfunctions within the subsystem and need not necessarily cause a system failure.

We assume

1° The time until malfunction in each mode of a given assembly has constant hazard rate and all are independent.

Suppose we separate the possible modes of malfunction for each assembly into mutually exclusive, functionally independent, classes labeling the time until malfunction T_{ij} for the i^{th} mode of the j^{th} assembly. The time until malfunction of the assembly by any mode is $T_j = \min_i(T_{ij})$ and the hazard rate of T_j is $\lambda_j = \sum_j \lambda_{ij}$, with the obvious interpretation of λ_{ij} as the hazard rate of T_{ij}. Unfortunately, the system may have different vulnerability to such malfunctions depending upon the period within the mission phase.

Thus the second situation in which the form (1.1) can arise is a series system with such malfunctioning assemblies. Assume

2° Given a malfunction in the j^{th} mode of the i^{th} assembly during the k^{th} time interval $(t^{(k-1)}, t^{(k)})$ of a mission, system failure will result with known conditional probability β_{ijk}.

For a mission of fixed length t we define the beta factor for the i^{th} assembly, which is a known constant, by the equation

$$\beta_i = \sum_{j,k} \beta_{ijk}[t^{(k)} - t^{(k-1)}]$$

where $t^{(0)} = 0$, $\lim t^{(k)} = t$. Thus from 1°, 2° the probability of system safety over the time interval $(0,t)$ is $h(\lambda,t)$ as given in (1.1) with $\omega_i t$ replaced by β_i. Without loss of generality we shall henceforth assume that all time intervals are expressed in fractions of this fixed mission length t, to wit, assume $t = 1$.

Needless to say, in the practical example given, all of these complications occur simultaneously. Moreover, the determination of the q_i through minimal state reliability analysis is itself a non-trivial task, not to speak of the analysis of the malfunction modes and their effect upon mission success (or vehicle safety). All of this is a necessary prelude for the determination of the beta factors. But we shall assume this work has been completed so that $\omega_i = q_i\beta_i$ is known, consequently equation (1.1) holds with $t = 1$.

The problem that we wish to discuss is of another genre, namely the methods for utilization of the data so as to determine a lower confidence bound for the reliability after the ω_i have been obtained. Since it is this aspect which is important we shall assume that $\beta_i \equiv 1$ for our data so as to further emphasize the differences between the two models under discussion. Of course this fictitiously makes the reliability estimate low.

2. THE DATA

In many cases the first (and sometimes the only) data one has concerning the reliability of the components comes from environmental tests. This test data must be reduced by engineering evaluation into the equivalent operational time during the given mission phase. These transformation factors, called K-factors, are used in industrial practice, see [1], to accelerate the testing and/or reduce its expense. Specifically, during the first phase of a mission a component may experience several types of vibration as well as several temperature and humidity changes. Consequently, testing the components in these separate environments must yield results requiring a transformation into the appropriate mission phase equivalent time. (The dangers of such a procedure are apparent but are taken in view of the exhorbitant cost of the alternative.)

We do not discuss this further but we merely point out that in such instances the data on the operational behavior of the components are not God-given, but rather are the construct of engineering knowledge and judgment.

Thus, the statistician is ultimately provided with data on all components in the form

$$x_i = (t_i, n_i) \qquad i=1,\ldots,m \tag{2.1}$$

where t_i is the total time, expressed in equivalent fractions of the mission length that the i^{th} component has been operated, and n_i is the total number of malfunctions of the i^{th} component during time t_i.

In the life qualification of components, it is usually the case that testing proceeds as long as there are funds available and this is usually neither until a fixed number of hours nor a fixed number of failures occur. Moreover, there are always extraneous circumstances which might terminate the testing program at any time. In view of this indeterminacy in the sampling procedure the treatment of the data that we adopt should not depend heavily upon a particular sampling scheme which might not obtain.

Assume that a number of identical components, say m, are put on test. What is observed at each trial is the random length of life, call it Y, when the component fails or the random time, say Z, at which the test is terminated for any reason other than failure of the component.

Thus we observe the event

$$[Y = y] \cap [Y < Z] \quad \text{or} \quad [Z = z] \cap [Y \geq Z]$$

which we note is not the minimum of Y and Z since we know whether Y or Z is observed. It is well known that the likelihood is independent of the sampling method for type I or type II censoring, i.e. stopping at either a fixed number of failures or after a fixed time".

We make the

Remark: If $(Z_1,\ldots,Z_m)$ is a vector of non-negative random variables, independent of all Y_i such that $[Y_i > Z_i]$, where $(Y_1,\ldots,Y_m)$ are themselves independently and identically distributed with common density function f and distribution F, then the likelihood of the event

$$\bigcap_{i=1}^{n} [Y_i \le Z_i][Y_i = y_i] \bigcap_{i=n+1}^{m} [Y_i > Z_i][Z_i = z_i],$$

where n ($\le m$) is the random number of failures observed, is of the form

$$C \prod_{i=1}^{n} f(y_i) \prod_{j=n+1}^{m} [1-F(z_j)] \tag{2.2}$$

and the constant C depends upon the outcomes $(y_1,\ldots,y_n)$ but not upon their distribution.

The proof is immediate. Let g be conditional density of $(Z_1,\ldots,Z_m)$ given $(Y_1,\ldots,Y_n)$ assuming $Z_1,\ldots,Z_m$ independent of $Y_{n+1},\ldots,Y_m$ on $\bigcap_{i=n+1}^{m} [Y_i > Z_i]$. Then the probability of the event specified above is

$$\int_{\{z_i>y_i,\ i=1,\ldots,n\}} \cdots \int \prod_{i=1}^{n} f(y_i) \prod_{j=n+1}^{m} [1-F(z_j)]\, g(z_1,\ldots,z_m|y_1,\ldots,y_n)\,dz_1,\ldots,dz_n$$

which upon simplification shows that

$$C = \int_{\{z_i>y_i,\ i=1,\ldots,n\}} \cdots \int g(z_1,\ldots,z_m|y_1,\ldots,y_n)\,dz_1,\ldots,dz_n$$

as claimed. ||

Taking the data $x_i = (t_i,n_i)$ for the i^{th} of m assemblies, the i^{th} assembly having exponential life with hazard rate λ_i, and substituting into (2.2) we obtain the likelihood

$$p(x_i|\lambda_i) = C(x_i)\lambda_i^{n_i} e^{-\lambda_i t_i} \tag{2.3}$$

where $C(x_i)$ is independent of λ_i.

Because the likelihood is the same for this very general sampling situation and we feel that the data by its nature requires such independence

we favor methods of statistical analysis which depend upon the likelihood.

From Bayesian Principles, Lindley [8], pp. 1,2, the joint posteriori density of $\underset{\sim}{\lambda} = (\lambda_1,\ldots,\lambda_m)$, based on the evidence $\underset{\sim}{x} = (x_1,\ldots,x_m)$ with likelihood given in (2.3), is

$$f(\underset{\sim}{\lambda}|\underset{\sim}{x}) \propto \prod_{i=1}^{m} p(x_i|\lambda_i)\pi(\underset{\sim}{\lambda}) \tag{2.4}$$

where π is the joint prior density of $\underset{\sim}{\lambda}$.

Two difficulties remain. One is to formulate a reasonable joint prior π and the second is then to calculate, other than symbolically, the posterior distribution of

$$V = \Sigma\beta_i\lambda_i, \text{ say } G(v|\underset{\sim}{x}) \text{ for } v > 0. \tag{2.5}$$

Following the usual method, p. 15, Lindley, loc. cit., the value v_0, depending upon $\underset{\sim}{x}$ and ε, $0 < \varepsilon < 1$ such that $G(v_0|\underset{\sim}{x}) = \varepsilon$, provides a lower $100\varepsilon\%$ Bayesian confidence bound for the system reliability e^{-V} given $\underset{\sim}{x}$ of the form e^{-v_0} and

$$P[\exp\{-\Sigma\beta_i\lambda_i\} > e^{-v_0}] = \varepsilon. \tag{2.6}$$

Essentially this method has been utilized to obtain confidence bounds on the reliability of certain systems and is presently the subject of much discussion, see [8]. In what follows we shall discuss two such methods and their reasonableness in dealing with the situation at hand.

3. A UNIFORM PRIOR

The first approach is to assume the special prior density

$$\pi(\underset{\sim}{\lambda}) \equiv 1 \quad \text{for all} \quad \lambda_i > 0. \tag{3.1}$$

This assumption is justified by the so-called principle of insufficient reason: since we know nothing specific about π we have insufficient reason to take π anything but uniform. Strictly speaking π as defined in (3.1) is a non-probabilistic prior. But, of course, one could consider it proportional to an approximation to a prior density.

Substituting (3.1) into (2.4) we find

$$f(\underset{\sim}{\lambda}|\underset{\sim}{x}) = \prod_{i=1}^{m} \frac{t_i(\lambda_i t_i)^{n_i} e^{-\lambda_i t_i}}{n_i!} \quad \text{for} \quad \lambda_i > 0. \tag{3.2}$$

The mathematical problem becomes that of finding the distribution of $V = \Sigma\beta_i\lambda_i$ where β_i are known constants and λ_i are gamma variates with known scale and shape parameters. To wit, each λ_j is $\Gamma(t_j, n_j+1)$ where $\Gamma(t,\nu)$ denotes the law with density, given $\nu > 0$

$$\frac{1}{\Gamma(\nu)} t^{\nu} x^{\nu-1} e^{-tx} \quad \text{for} \quad x > 0. \tag{3.3}$$

We also quote two related results, see p. 46ff, Feller [7].

If λ_j is $\Gamma(t_j,\nu_j)$, then $t_j\lambda_j$ is $\Gamma(1,\nu_j)$.

If $\nu_j > 1$, then $\lambda_j = \lambda_j' + \lambda_j''$ in distribution where λ_j' is $\Gamma(t_j,1)$ independent of λ_j'' which is $\Gamma(t_j,\nu_j-1)$.

Thus by the first remark we see that in distribution $V = \sum_{j=1}^{m} b_j\lambda_j$

where $b_j = \beta_j/t_j = 1/\tau_j$, $j=1,\ldots,m$ and each λ_j is now $\Gamma(1,n_j+1)$. By the second remark, for the data given in Table I where we have at most two failures, we see that in distribution

$$V = \sum_1^m b_i Z_i + \sum_1^r b_j Y_j + \sum_1^s b_k X_k \tag{3.4}$$

where s is the number of components with two failures during testing

$r-s$ is the number of components with one failure during testing

$m-r-s$ is the number of components with no failures during testing

and X_k, V_j, Z_i are all independent $\Gamma(1,1)$, i.e., exponential with unit mean, variates.

We now quote a result proved, for example, in [11] as a

Lemma 1: If $Z_1,\ldots,Z_k$ are independent exponential random variables with unit mean, then for $b_i > 0$, all distinct, we have

$$P[\sum_1^k b_i Z_i > u] = \sum_{j=1}^k B_j^{(k)} e^{-u/b_j} \tag{3.5}$$

where $B_1^{(1)} = 1$ and for $k \geq 2$

$$B_j^{(k)} = \prod_{\substack{i=1 \\ i\neq j}}^k \frac{b_i}{b_j - b_i} \quad \text{for} \quad j=1,\ldots,k. \tag{3.6}$$

Also these recursion relations hold

$$B_j^{(k)} = B_j^{(k-1)} b_i/(b_j-b_k), \quad j=1,\ldots,k-1 \quad \text{and} \quad B_k^{(k)} = 1 - \sum_{j=1}^{k-1} B_j^{(k)}.$$

Also we have

Lemma 2: The distribution of $\sum_1^m b_i Z_i + \sum_1^r b_j Y_j$ is

$$\sum_{i=1}^{m}\sum_{j=1}^{r} B_i^{(m)} B_j^{(r)} \{1-e^{-\tau_j t} {}_j\psi(i,j,t)\} \quad \text{for} \quad t > 0,$$

where $\tau_j = 1/b_j$ for $j=1,\ldots,\max(m,r)$

$$\psi(i,j,t) = \int_0^t \exp\{-\tau_i(t-y)-\tau_j y\}dy$$

$$= \begin{cases} te^{-\tau_i t} & \text{if} \quad i = j \\ \dfrac{e^{-\tau_i t} - e^{-\tau_j t}}{\tau_j - \tau_i} & \text{if} \quad i \neq j. \end{cases}$$

The proof is accomplished by the convolution of two distributions each of the form given in Lemma 1.

Consider the more general definition

$$V_k = \sum_{j=1}^{k}\sum_{i=1}^{m_j} b_i X_{ij} \tag{3.7}$$

where the X_{ij} are all independent exponential variates with unit mean. Let V_k have distribution F_k, then

$$V_k = V_{k-1} + \sum_{i=1}^{m_k} b_i X_{ik}.$$

Defining $\bar{F} = 1-F$ with any affixes, and taking τ_i as given in Lemma 2,

$$\bar{F}_k(u) = \int_0^\infty P\left[\sum_{i=1}^{m_k} b_i X_{ik} > u-v\right] dF_{k-1}(v)$$

$$= \bar{F}_{k-1}(u) + \sum_{i=1}^{m_k} B_i^{(m_k)} \left\{\int_0^u e^{-(u-v)\tau_i} dF_{k-1}(v)\right\}.$$

But the quantity in braces in the equation above becomes

$$\{\cdots\} = \bar{F}_{k-1}(u) - e^{-u\tau_i} - \tau_i \int_0^u \bar{F}_{k-1}(v) e^{-(u-v)\tau_i} dv.$$

Hence we have shown the following:

Lemma 3: The survival probability of V_k as defined in equation (3.7) is given in terms of the survival probability of V_{k-1} as

$$(3.8)\quad \bar{F}_k(u) = \sum_{i=1}^{m_k} B_i^{(m_k)} \left[e^{-u\tau_i} + \tau_i \int_0^u \bar{F}_{k-1}(v) e^{-(u-v)\tau_i} dv \right].$$

Note that (3.8) can be used to prove Lemma 2 and used recursively to find the distribution of V_k for small k. Thus we now have a computationally feasible method for the calculation of the distribution of V, called G. Using the lemmas above a machine program was written for the IBM 360, using double precision for the computation of the $B_j^{(k)}$, which tabulates the distribution of V in the region of interest. Using the data presented in Table I, this distribution is graphed in Figure 1.

For example, we find that if $v_0 = 12.8$, $G(v_0) = .95$. Thus a lower 95% Bayesian confidence limit for the system reliability is $e^{-12.8} \doteq 10^{-5}$.

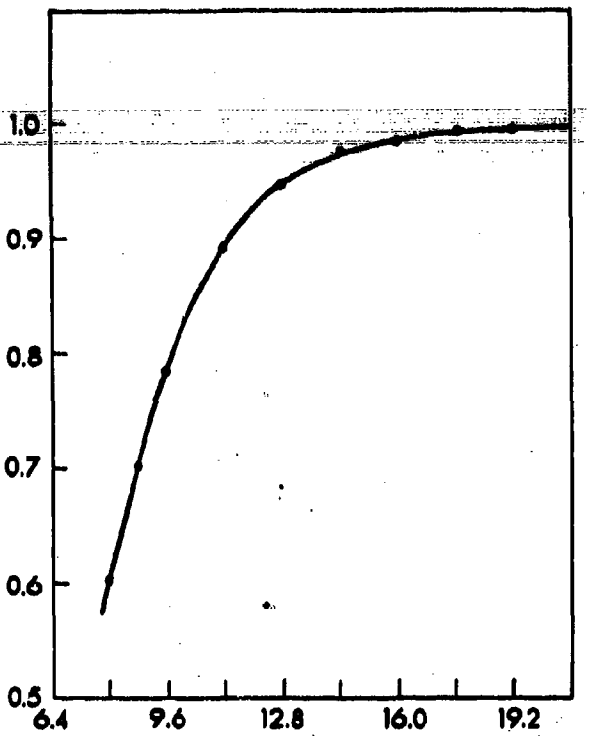


Figure 1

Graph of the distribution F of the random variable $V = \Sigma\beta_i\lambda_i$ in the region near unity. If v is the abscissa value, the value F(v) of the ordinate is the confidence the system reliability exceeds e^{-v}.

4. A CRITICAL DISCUSSION

A word about the computation necessitated by this method. It is clear from Table I that the differences of the τ_j are neither small nor uniform. A glance at the formula for the $B_j^{(k)}$ in equation (3.6) shows that in absolute value they can become very large for such cases. (In fact, for such data as we have for 67 components, values as high as 10^{20} are not impossible.) Since all $B_j^{(k)}$ summed over j must add to unity, some must be positive and some negative. However, because of the nature of machine decimal arithmetic, the summands will be rounded off and the machine cumulate the error. We should, by definition, have $G(0) = 0$ but computationally we do not. For example, referring to Figure 1, the machine value at $v = 3.2$ was $G(3.2) = .699 \times 10^{-3}$ but at $v = 2.4$, $G(2.4) = -.123 \times 10^{-1}$ with wider fluctuations for smaller values of v. Fortunately, we are interested in those values of the argument for which $G(v)$ is near one and the values of v necessarily become large enough to eliminate the errors due to this circumstance.

However, this is merely a limitation due to the accuracy of the method of computation which was adopted. We feel there is a much more primary objection.

Without any real loss of clarity to the fundamental ideas, let us fix our attention on an assembly with two separate modes of malfunction with hazard rates λ_1 and λ_2, say. Suppose this assembly was operated for a time t and no failure of either type was observed. By using the uniform prior of (3.1) the posterior distribution of the component hazard rate $\lambda(= \lambda_1+\lambda_2)$ considering the component as a unit, is

(4.1) $P[\lambda < a] = P[e^{-\lambda} > e^{-a}] = 1-e^{-at}$ for $a > 0$

so that e^{-a} is a lower bound on the reliability for a mission of unit length and the confidence level is $P[\lambda < a]$.

But on the other hand, by considering the posterior distribution of $\lambda_1 + \lambda_2$, using data $x_1 = x_2 = (t,0)$ and the principle of insufficient reason to apply the uniform prior for each mode, we have

(4.2) $P[\lambda < a] = 1-e^{-at}-ate^{-at}$ for $a > 0$

as the posterior distribution of the hazard rate λ of the assembly. But notice that (4.2) is less than (4.1) which was the distribution from the same data for the same assembly.

The point we are making is simply this: For a series system, with no component failing during test time $t > 0$, the confidence in the reliability of the system should be the same as that for each component, since the system and the components both experience the same operational time t without failure.

Our criticism of the former method is that the confidence in the reliability does not depend only upon the data, it also depends upon the arbitrary designation of component or assembly. If we arrive at different answers when using the same data, then something must be wrong.

To continue this point further, let us suppose that we have a series system with separate malfunction modes with hazard rates $\lambda_1,\ldots,\lambda_m$ each of which has acquired the same operational experience,

namely $x_i = (t, 0)$ for $i=1,\ldots,m$, i.e., no failures during operation for a length of time t. Again by using the uniform prior density we have the distribution of $\lambda = \lambda_1+\cdots+\lambda_m$ as

$$1 - \sum_{j=1}^{m} \frac{e^{-at}(at)^j}{j!} \tag{4.3}$$

which approaches zero as m approaches infinity regardless of the value of $ta > 0$.

Of course, (4.2) results from a different specification of the Bayesian model. The point is that each specification of another independent component will always result in a different posterior distribution. (Needless to say, a different prior will lead to a different posterior density for λ as well.)

Moreover, it is clear that almost any choice of prior density of λ which is the product of independent prior densities for each λ_i will result in a confidence level which is essentially the same as that given in (4.3), to wit so low as to be nonsensical for m large.

One modification suggested is to assume functional dependence with statistical independence, among the prior densities of λ_i. One such is to take the (conjugate) prior density of λ_i as $\Gamma(u_i,v_i)$ for $i=1,\ldots,m$, using here the notation of (3.5), subject to the constraint

$$v_i > 0, \qquad \sum_{1}^{m} v_i = 1. \tag{4.4}$$

Combining this with the likelihood of the form (2.3) shows

$$f(\underline{\lambda}|\underline{x}) \propto \prod_{i=1}^{m} [\lambda_i^{n_i+v_i-1} e^{-\lambda_i(u_i+t_i)}]$$

subject to (4.4) above. Thus the posterior density of λ_i is $\Gamma(u_i+t_i, n_i+v_i)$ and thus the mathematical problem becomes that of finding the distribution of $U = \Sigma c_i \lambda_i$ where $c_j = \beta_j/(u_j+t_j)$ and each λ_j is $\Gamma(1, n_j+v_j)$ for $i=1,\ldots,m$ subject to (4.4).

Because $n_j + v_j$ is not an integer we are faced with an analytic and computational problem beyond that of the preceding section. However, it is clear that this artifice does introduce enough degrees of freedom that proper choice of u_j, maintaining the restriction (4.4), can yield reliabilities of not unreasonable size. We do not pursue it further. The difficulty, making such an assumption untenable, is that ones prior knowledge about the reliability of a component should neither depend upon the prior densities of the other components in any way nor upon how many of them there are. These prior densities should be independent in every sense.

5. THE DEGENERATE PRIOR

The basis of Bayesian methodology is to regard the λ_i as random variables having a distribution which is to be constructed from prior knowledge. The more commonly accepted point of view is that the λ_i are unknown constants about which inference must be made. This is now the point of view which we essentially adopt.

If the λ_i were unknown positive real numbers, then there would exist a constant of proportionality between any two λ_i's which would be fixed, even though it was unknown.

Thus we make the assumption

3° There exists a constant of proportionality, say α_{ij}, between any two λ_i and λ_j.

If we have m different modes of malfunction, we define α_i for $i=1,\ldots,m$ as the *probability of malfunction in the* i^{th} *mode given that a malfunction in the system has occurred.* One sees that

$$\alpha_i = P[T_i < t \mid \sum_{j=1}^{m} [T_j < t]],$$

where we made the convention that the summation of events denotes the disjoint union. It follows that $\alpha_{ij} = \alpha_i/\alpha_j$ where

$$\alpha_i = \frac{\lambda_i}{\sum_{j=1}^{m} \lambda_j} \qquad i=1,\ldots,m. \tag{5.0.1}$$

Thus 3° is equivalent to taking the prior distribution, say $\Pi(\underset{\sim}{\lambda})$, to be singular with all measure concentrated along a ray out from the origin with the direction of the ray determined by the constants of proportionality. Specifically, we assume

$$(5.1)\quad d\Pi(\lambda_1,\ldots,\lambda_m)\begin{cases} > 0 & \text{if some } \lambda_i = \dfrac{\alpha_i\lambda_1}{\alpha_1} \text{ for } i \neq j \\ = 0 & \text{otherwise.}\end{cases}$$

In the case $m = 2$, $\Pi(\lambda_1,\lambda_2)$ is zero everywhere but along the ray $\lambda_2 = \dfrac{\alpha_2}{\alpha_1}\lambda_1$ out from the origin in the (λ_1,λ_2) plane. We wish to find the posterior density of $\sum_1^m \beta_i\lambda_i$. We make the change of variables $\rho_i = \beta_i\lambda_i$ and by (2.3) and $\tau_i = t_i/\beta_i$ we have

$$f(\underset{\sim}{\rho}|x) \propto \sum_{i=1}^{m} (\tau_i\rho_i)^{n_i} e^{-\tau_i\rho_i} d\Pi^*(\rho_1,\ldots,\rho_m)$$

where

$$\Pi^*(\rho_1,\ldots,\rho_m) = \Pi(\frac{\rho_1}{\beta_1},\ldots,\frac{\rho_m}{\rho_m}).$$

Thus

$$(5.2)\quad d\Pi^*(\rho_1,\ldots,\rho_m)\begin{cases} > 0 & \text{if some } \rho_i = \dfrac{\alpha_i\beta_i\rho_i}{\alpha_j\beta_j} \text{ for } i \neq j \\ = 0 & \text{otherwise.}\end{cases}$$

The density we seek is proportional to

$$(5.3)\quad \int_{\{\underset{\sim}{\rho}:\Sigma\rho_i=a\}} \cdots \int \prod_{i=1}^{m} (\tau_i\rho_i)^{n_i} e^{-\tau_i\rho_i} d\Pi^*(\rho_1,\ldots,\rho_m).$$

Consider the line in m-space

$$\ell(\rho_1) = (\rho_1, \frac{\alpha_2\beta_2}{\alpha_1\beta_1}\rho_1,\ldots, \frac{\alpha_m\beta_m}{\alpha_1\beta_1}\rho_1).$$

By equation (5.2) all the mass of Π^* is concentrated along the ray $\ell(\gamma_1)$ for $\rho_1 > 0$. In effect the only quantity that has a distribution is ρ_1 and we shall later see it makes no difference what this distribution is as long as it has support on $(0,\infty)$. This line intersects the plane $\sum_1^m \rho_i = a$ at a single point, namely ρ_1 such that

$$\rho_1 + \sum_{i=2}^{m} \frac{\alpha_i\beta_i}{\alpha_1\beta_1}\rho_1 = a$$

and solving for ρ_1 we find $\rho_1 = \gamma_1$ where we define

$$\gamma_i = \frac{\alpha_i\beta_i}{\sum_{j=1}^{m}\alpha_j\beta_j} \qquad i=1,\ldots,m. \tag{5.3.1}$$

Then the value of ρ_i at the point of intersection of the line ℓ with the plane $\Sigma\rho_i = a$ is $\rho_i = a\gamma_i$ for $i=1,\ldots,m$. Since all the measure of Π^* is concentrated along the line ℓ, the integration over the plane in (5.3) yields a single value at the singularity of the measure Π^*. It follows that the density we seek is proportional to the value of the integrand at that point, namely

$$\prod_{i=1}^{m} (\tau_i a\gamma_i)^{n_i} \exp\{-\sum_{i=1}^{m} \tau_i a\gamma_i\}.$$

If we define

$$\theta = \sum_{i=1}^{m} \tau_i\gamma_i = \frac{\Sigma t_i\alpha_i}{\Sigma\beta_i\alpha_i} \quad \text{and} \quad k = \sum_1^m n_i, \tag{5.4}$$

which are, respectively, a weighted mean of the t_i, and the total number of failures, we can write the posterior density of $V = \sum_1^m \rho_i$ as

$$(5.5) \qquad \frac{\theta(\theta a)^k e^{-a\theta}}{k!} \qquad \text{for} \qquad a > 0.$$

The distribution then is

$$(5.6) \qquad P[V < u] = \int_0^{\theta u} \frac{s^k e^{-s}}{k!}\, ds = 1 - \sum_{j=0}^{k} \frac{e^{-\theta u}(\theta u)^j}{j!}$$

which we recognize as a Chi-square distribution. If we set

$$(5.7) \qquad u = \frac{1}{2\theta}\, \chi_\varepsilon^2(2k+2),$$

where $\chi_\varepsilon^2(m)$ is the $100\varepsilon^{th}$ percentile of the Chi-square distribution with m degrees of freedom, we have e^{-u} providing a lower confidence bound of level ε.

We note that the computation for this method is trivial. We compute only the two quantities k and θ and then from a table of the Chi-square distribution calculate u and e^{-u}.

Unfortunately, equation (5.7) gives the confidence bound e^{-u} in terms of the alpha factors which are still unknown. Nonetheless, based upon the objective model that the failure rates are virtually unknown constants, we do arrive at (5.7) and knowledge concerning the a_i is what is needed to determine the confidence limit. However, this does not necessarily mean that the values of λ_i

need be known, for example, it is sufficient that their ratios be known. Perhaps in some instances engineering experience might be able to classify all the failure rates as multiples of fixed one, say the lowest, at least in a conservative manner.

Disregarding for the present the computation of θ, this method does obviate some of the conceptual difficulties which the preceding method possessed.

Firstly, the confidence bound is the same regardless of how the components are apportioned to subsystems within the system. In particular, if $\tau_1 = \cdots = \tau_m$, we obtain the same density of $\Sigma\beta_i\lambda_i$ as we would by considering the system as a single unit.

The addition of components to the system none of which have failed, i.e., data of the type $(t_i,0)$, do not necessarily cause the confidence to go rapidly to zero. (Of course, the confidence does depend upon t_i through θ.) It is clear from (5.5) that it is not the number of components but the number of failures which rapidly decrease the confidence.

In the special case when $\beta_{ijk} = \beta_i$ for all j,k we can make an intuitive interpretation of γ_i as the conditional probability of failure of the i^{th} assembly given that an assembly has failed. To see this, label the events "the i^{th} assembly fails" by F_i and "the i^{th} assembly malfunctions" by M_i. By definition

$$\beta_i = P[F_i|M_i], \qquad \alpha_i = P[M_i|\Sigma M_j]$$

and from the calculus of probabilities, since $F_i \subset M_i$

$$\alpha_i \beta_i = \frac{P(F_i)}{P[\Sigma M_j]}$$

and hence from (5.3.1) follows $\gamma_i = P[F_i | \Sigma F_i]$.

We now make two calculations to indicate the reliability values obtained by this method.

Example 1:

Let us suppose that $\gamma_i = \frac{1}{m}$ for $i=1,\ldots,m$. We recall that under certain conditions this would mean the event any one particular component had failed, knowing that exactly one component was in a failed state, was equally likely with the event any other component had failed.

From Table I we find $k = 8$ and compute from (9.4), $\theta = \Sigma\tau_i/m = 25.35$ and hence for $\epsilon = .95$, using the Chi-square value for 16 degrees of freedom, we have $u = (28.87)/50.7 = .569$. Thus $e^{-u} = .566$ is a lower 95% confidence limit for the system reliability.

Example 2:

Let us suppose $\alpha_i = \frac{1}{m}$ for $i=1,\ldots,m$ and from Table I, we again use (9.4) to compute $\theta = (\Sigma t_i)/(\Sigma\beta_i) = 16.27$. For $k = 8$, $\epsilon = .95$ we find $u = (28.87)/(32.54) = .887$ and $e^{-u} = .412$ is the lower 95% confidence limit for the reliability of the system.

6. BOUNDS ON θ

In this section we make the argument that what prior information one has about λ_i for $i=1,\ldots,m$ should be applied so as to determine bounds on θ rather than in the production of prior distributions of the component failure rates.

It is clear that if $\underline{\alpha} = (\alpha_1,\ldots,\alpha_m)$ is constrained and a lower bound $\theta_1 \leq \theta(\underline{\alpha})$ can be determined, then correspondingly from (5.7) $u_1 \geq u$, from which it follows that e^{-u_1} provides a lower confidence bound of level not less than ϵ. For example, the trivial inequality

$$\min_{i=1}^{m} \tau_i \leq \theta(\underline{\alpha})$$

will provide such a bound. However, unless the τ_i are nearly all equal, a state devoutly to be wished and planned for, it is not certain this bound would be a useful result. However, if testing were continued until $\tau_i = \tau_0$ for $i=1,\ldots,m$, we would then be in the favorable position that α_i need not be known.

If τ_j were the minimum of τ_i for $i=1,\ldots,m$, then $\tau_j = \theta(\underline{\alpha})$ implies $\gamma_j = 1$, $\gamma_i = 0$ for $i \neq j$ which in turn by (5.3.1) implies that $\alpha_j = 1$, $\alpha_i = 0$ for $i \neq j$ which requires by (5.0.1) that λ_j be infinitely large with respect to all other λ_i. This would seem to be an unlikely state of nature, one which might be reasonably excluded from consideration.

We now give some examples of information which in various degrees exclude the state mentioned above and are of a type which may provide a non-trivial bound.

Let us suppose it is known that

$$\sum_{i=1}^{m} \alpha_i \beta_i = p. \tag{6.1}$$

(This would mean in certain situations that the probability of a failure given a malfunction was known to be p.)

Thus the problem becomes that of minimizing the function $\theta(\underline{\alpha})$ as defined in (5.4) where (t_i, β_i) are known positive numbers subject to the restriction (6.1) and

$$\sum_{1}^{m} \alpha_i = 1, \quad \alpha_i \geq 0 \quad \text{for} \quad i=1,\ldots,m. \tag{6.2}$$

Call the set of $\underline{\alpha}$ satisfying (6.1) and (6.2) the set $\mathcal{A}_p$.

Clearly with the denominator fixed in (5.4) we have a linear programming problem with two constraints, for which the theory is well known.

Of course the restriction (6.1) is a mathematical convenience. What we desire are bounds on $\min_p \psi(p)$ for p taken over some subset of the range

$$\min \beta_i < p < \max \beta_i,$$

where

$$\psi(p) = \min\{\theta(\underline{\alpha}): \underline{\alpha} \in \mathcal{A}_p\}. \tag{6.3}$$

This can be obtained from a graph of $\psi(p)$, which is here accomplished with a linear program using p as a parameter. A plot using the data of Table I is given in Figure 2 as on illustration.

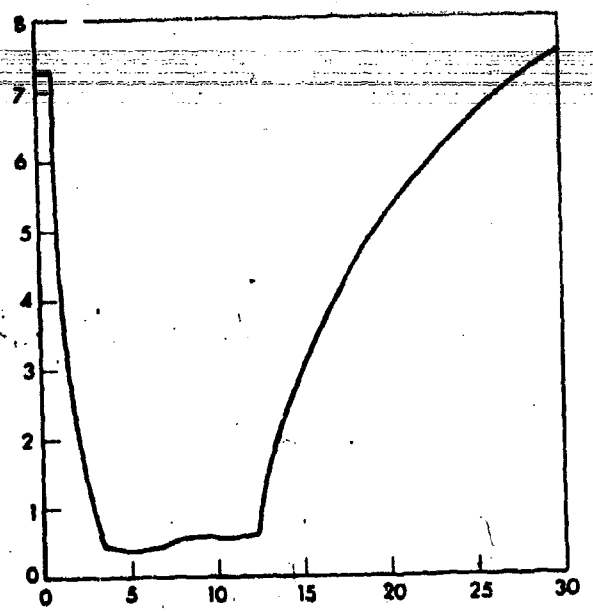


Figure 2

Graph of $\psi(p)$ for $1 \leq p \leq 30$.

Consider θ as a function of $\underline{\gamma} = (\gamma_1,\ldots,\gamma_m)$ as defined in (5.4). It is linear (hence convex) and we wish to minimize it subject to some restrictions on its domain which are measures of the variability of the λ_i for $i=1,\ldots,m$. First we shall consider the region $\mathcal{R}_x$ defined for $0 \le x \le \sqrt{m-1}$ by

$$(6.4) \quad 1 \le m\Sigma\gamma_i^2 \le x^2 + 1, \quad \Sigma\gamma_i = 1, \quad \gamma_i \ge 0 \quad \text{for} \quad i=1,\ldots,m.$$

The region $\mathcal{R}_x$ is convex and thus there exists a unique minimum for θ over the region. The method we shall use is Lagrange multipliers.

Let

$$\phi(\underline{\gamma}) = \theta(\underline{\gamma}) - \frac{m\lambda_1}{2}\Sigma\gamma_i^2 - \lambda_2\Sigma\gamma_i.$$

We wish to minimize ϕ subject to the conditions

$$(6.5) \quad m\Sigma\gamma_i^2 = x^2 + 1, \quad \Sigma\gamma_i = 1, \quad \gamma_i \ge 0 \quad \text{for} \quad i=1,\ldots,m.$$

Thus

$$\frac{\partial\phi}{\partial\gamma_j} = \tau_j - m\lambda_1\gamma_j - \lambda_2 \quad j=1,\ldots,m.$$

We now consider the three equations

$$(6.6) \quad \Sigma\gamma_i \frac{\partial\phi}{\partial\gamma_i} = 0, \quad \frac{1}{m}\Sigma\tau_i \frac{\partial\phi}{\partial\gamma_i} = 0, \quad \frac{1}{m}\Sigma\frac{\partial\phi}{\partial\gamma_i} = 0$$

which upon simplication, and imposing the restrictions of (6.5), yield three equations which are to be solved for θ by eliminating λ_1 and λ_2

First eliminating λ_2 we obtain

$$\lambda_1 = \frac{\sigma^2}{\theta-\bar{\tau}}, \qquad (\theta-\bar{\tau})^2 = x^2\sigma^2 \tag{6.7}$$

where $\bar{\tau}$ and σ are the mean and standard deviation of τ_j's, respectively.

Thus for a given value of x we can obtain the minimum value of θ from (6.7) as the smaller root

$$\theta = \bar{\tau} - x\sigma. \tag{6.8}$$

Also from $\frac{\partial \phi}{\partial \gamma_j} = 0$ we must have

$$\gamma_j = \lambda_1^{-1} m(\tau_j - \lambda_2) \geq 0 \quad \text{for} \quad j=1,\ldots,m \tag{6.9}$$

in order to satisfy the restrictions. Since by (6.7) and (6.8) $\lambda_1 = -\sigma/x$ we see that a sufficient condition to satisfy (6.9) is

$$\max(\tau_j) \leq \lambda_2 = \bar{\tau} + \sigma x^{-1}. \tag{6.10}$$

This is satisfied for all τ_j which are reasonably close together. It is clear from (6.8) that the minimum value of θ is a decreasing function of x.

Remark: If $\tau_1,\ldots,\tau_m$ satisfy (6.10) for a given x between $0 \leq x \leq \sqrt{m-1}$, then

$$\min\{\theta(\chi): \chi \in \mathcal{B}_x\} = \bar{\tau} - x\sigma$$

where $\bar{\tau},\sigma$ are defined in (6.7).

From this remark we see the restriction (6.4) yields a lower $100\epsilon\%$ confidence bound for the reliability, namely

(6.10.1) $\exp\{-x_\varepsilon^2(2k+2)/2(\bar{\tau}-x\sigma)\}$ for $0 \le x \le \sigma/(\max \tau_j - \bar{\tau})$.

Using the data in Table I we find $\bar{\tau} = 25.35$, $\sigma = 21.25$ and $\max \tau_j = 98.1$ and thus the range of x is $0 \le x \le .31$. A graph of (6.10.1) for this case with $k = 7$, $\varepsilon = .95$ is given in Figure 3.

Next we consider the problem of minimizing $\theta(\gamma)$ subject to

(6.11) $\Sigma\alpha_i^2 = \kappa$, $\Sigma\alpha_i = 1$, $\alpha_i \ge 0$ for $i=1,\ldots,m$.

Since we can write

$$\alpha_j = \gamma_j\beta_j^{-1}/(\Sigma\gamma_i\beta_i^{-1}) \quad \text{for} \quad j=1,\ldots,m,$$

the first restriction can be put in the alternate form $\Sigma\gamma_j^2\beta_j^{-2} = \kappa(\Sigma\gamma_j\beta_j^{-1})^2$. Again we use Lagrange multipliers to take advantage of the symmetry of the problem. Write

$$\phi = \theta - \frac{m\lambda_1}{2}\left[\Sigma\gamma_i^2\beta_i^{-2} - \kappa(\Sigma\gamma_i\beta_i^{-1})^2\right] - \lambda_2(\Sigma\gamma_i - 1)$$

$$\frac{\partial\phi}{\partial\gamma_j} = \tau_j - m\lambda_1[\gamma_j\beta_j^{-2} - \kappa(\Sigma\gamma_i\beta_i^{-1})\beta_j^{-1}] - \lambda_2.$$

We now consider the four equations

$$\frac{1}{m}\sum_1^m \beta_j\frac{\partial\phi}{\partial\gamma_j} = 0, \quad \frac{1}{m}\sum_1^m \beta_j^2\frac{\partial\phi}{\partial\gamma_j} = 0, \quad \sum_1^m \gamma_j\frac{\partial\phi}{\partial\gamma_j} = 0, \quad \sum_1^m \tau_j\beta_j^2\frac{\partial\phi}{\partial\gamma_j} = 0.$$

By setting $\delta = \Sigma\gamma_j\beta_j^{-1} = (\Sigma\alpha_i\beta_i)^{-1}$ for notational convenience and imposing the restraints as encountered we obtain four equations which contain the four variables $\lambda_1,\lambda_2,\theta,\delta$. Eliminating λ_1,λ_2 and δ gives a quadratic equation in θ, namely

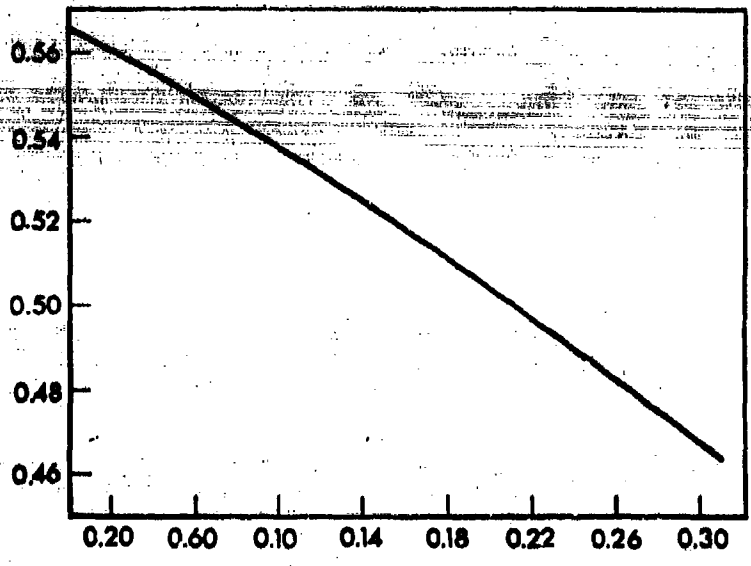


Figure 3

Values of $\exp\{-28.87/2(25.35-21.25x)\}$ for $0 \le x \le .31$: a lower 95% confidence bound on the reliability given $1/m \le \Sigma\gamma_1^2 \le \frac{x^2+1}{m}$.

$$\theta^2 A - 2\theta B + C = 0$$

where, recalling $t_i = \beta_i \tau_i$ for $i = 1, \ldots, m$,

$$A = (\bar{\beta})^2 - x\overline{\beta^2}, \qquad B = \bar{t}\bar{\beta} - x\overline{t\beta}, \qquad C = (\bar{t})^2 - x\overline{t^2}$$

and all coefficients depend upon the value of $x = 1 - (m\kappa)^{-1}$. Thus the minimum value is the smaller root, call it

$$\theta_x = (B_x - S_x)/A_x \tag{6.14}$$

where $S^2 = B^2 - AC$ or equivalently $S_x^2 = xa^2 - x^2 b^2$

and

$$a^2 = \frac{1}{m}\Sigma(t_i\bar{\beta} - \beta_i\bar{t})^2, \qquad b^2 = \overline{t^2}\overline{\beta^2} - (\overline{t\beta})^2.$$

It is clear that if x is constrained by

$$0 \le x \le \min\left\{\frac{(\bar{\beta})^2}{\overline{\beta^2}}, \frac{\bar{t}\bar{\beta}}{\overline{t\beta}}, \frac{(\bar{t})^2}{\overline{t^2}}, \frac{a^2}{2b^2}, 1 - \frac{1}{m}\right\}, \tag{6.14.1}$$

then the values of θ_x are meaningful.

We now argue that there is an interval of positive values of x for which θ_x is decreasing as a function of x. Note $\theta_x' \le 0$ iff

$$\zeta(x) = -S_x' A_x - \overline{\beta^2} S_x \le \overline{t\beta}\,(\bar{\beta})^2 - \bar{t}\bar{\beta}\overline{\beta^2} \tag{6.15}$$

To see that ζ is an increasing function, note that $\zeta'(x) = -A_x S''_x$. Now $S'(x) = (a^2-2xb^2)/2S(x)$ so that by (6.14.1), $S'(x) > 0$ and likewise we check $S''(x) < 0$ and hence $\zeta'(x) \geq 0$. Notice that $\zeta(0) = -\infty$ so there is an interval of values in x, for which (6.15) is true, and θ_x is decreasing. This region can be determined from (6.15) in each case of interest.

However, we must also satisfy the condition $\gamma_j \geq 0$ for $j=1,\ldots,m$. From $\frac{\partial \phi}{\partial \gamma_j} = 0$ follows, by using (6.13)

$$(6.16) \qquad \gamma_j \geq 0 \quad \text{iff} \quad 1 \geq \frac{\beta_1\theta - t_1}{\kappa m \lambda_1^{\delta}} = \frac{x(t_1-\beta_1(\theta))}{\bar{t}-\bar{\beta}\theta} .$$

But by the argument above $\theta_x < \theta_0 = \bar{t}/\bar{\beta}$ for $x > 0$. Hence the denominator of the right-hand side of (6.16) is positive. Thus $\gamma_i \geq 0$ for $j=1,\ldots,m$ iff

$$(6.17) \qquad \bar{t} - \theta_x\bar{\beta} \geq x \max_{j=1}^{m}\{t_j - \theta_x\beta_j\}.$$

To be presuaded that there is indeed a neighborhood of zero in which (6.17) is true we introduce the power series expansion of θ_x:

$$\theta_x = \frac{\bar{t}}{\bar{\beta}} - \sqrt{x}\,\frac{a}{(\bar{\beta})^2} + O(x).$$

Thus (6.17) is equivalent with

$$\frac{a}{\bar{\beta}} - \sqrt{x}\,\frac{\bar{t}\bar{\beta}^2 - \bar{\beta}\bar{t}\bar{\beta}}{(\bar{\beta})^2} + O(x) \geq \sqrt{x}\,\max_j[t_j - \theta_x\beta_j]$$

which is clearly true for x sufficiently small. Thus we can make the

Remark: If $(t_i, \hat{\beta}_i)$ $i=1,\ldots,m$ are given and x satisfies (6.14.1), then θ_x defined by (6.14) satisfies

$$\theta_x = \min\{\theta(\underline{\gamma}) : \underline{\gamma} \rightarrow \frac{1}{m} \leq \Sigma\alpha_i^2 \leq \frac{1}{m(1-x)}\}$$

whenever (6.17) is true.

Using the data from Table I we find that the upper bound for (6.14.1) is $a^2/2b^2 = .26056$ while the largest value of x satisfying (6.17) is .145. A graph of $\exp\{-\chi_\epsilon^2(2k+2)/2\theta_x\}$ for $k = 7$, $\epsilon = .95$ with θ_x as defined in (6.14) is given in Figure 4.

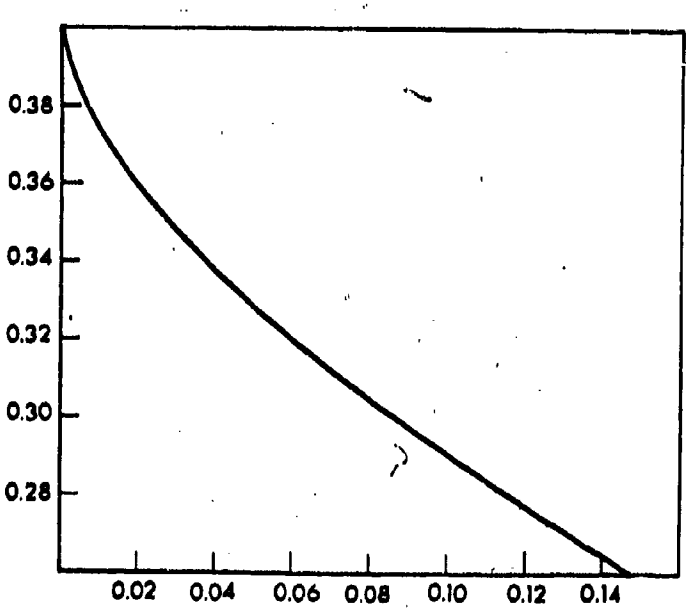


Figure 4

Graph of $\exp\{-\frac{28.87}{2\theta_x}\}$ for $0 < x < .15$: a lower 95% confidence bound on the reliability given $\frac{1}{m} \leq \Sigma\alpha_i^2 \leq \frac{1}{m(1-x)}$.

CONCLUDING REMARKS. By considering a series system with a given number of components, each component having been tested for the same time and each experiencing no failures, one intuitively feels since the entire system could have been operated, conceptually at least, without failure for that same period of time that the confidence one has in the system's reliability should be exactly the same as the confidence in each component's reliability. Moreover, this confidence should be the same irrespective of the number of components in the system.

By thinking how the confidence should behave for such a series system, as components are added with different test times and different numbers of failures, we see that the Bayesian approach with independent prior distributions of the failure rate fails to fulfill our expectation as to this incremental behavior. At the same time the second model chosen, with failure rates as virtual constants, does seem to behave in conformity with our intuition and moreover it has the added appeal of computational simplicity.

Lastly, for the practical case chosen namely early data for the Saturn 1-C which at this juncture we know is a highly reliable system, the second model gave reasonable interval estimates of the reliability while the first did not.

ACKNOWLEDGMENTS. The author would like to express his appreciation to the reliability group at the Launch Systems Branch of The Boeing Company for providing the data for this discussion and assistance in its organization. In particular, I want to thank Francis Bari who first called this question to my attention by mentioning some deficiencies of the extant methods for determining confidence bounds.

TABLE I

Summary of test data for Saturn I-C

t_i = test time in mission lengths, n_i = number of failures observed

$\omega_i = q_i$ = component multiplicity, $\tau_i = t_i/\beta_i$

t_i	n_i	ω_i	τ_i	t_i	n_i	ω_i	τ_i
318.5	0	16	19.9	48.7	0	1	48.7
138.8	0	16	8.7	33.9	0	1	33.9
69.4	0	16	4.3	30.2	0	1	30.2
159.2	0	4	39.8	45.7	0	1	45.7
187.9	0	8	23.5	36.5	0	1	36.5
144.9	0	4	36.2	50.6	0	2	28.3
69.7	0	4	17.4	45.2	0	2	22.6
148.7	0	4	37.2	22.6	0	2	11.3
146.8	0	4	36.7	37.7	0	2	18.8
15.1	2	2	7.5	49.9	0	1	49.9
7.5	0	12	.63	34.3	0	1	34.3
120.7	0	4	30.2	37.7	0	1	37.7
113.1	0	3	37.7	11.3	0	1	11.3
98.1	0	1	98.1	15.1	0	2	7.5
92.8	0	1	92.8	226.1	0	30	7.5
9.0	0	1	9.0	30.3	0	10	3.0
97.9	0	13	7.6	32.0	0	2	16
87.3	0	4	21.8	179.0	0	8	22.4
26.4	0	1	26.4	32.5	2	1	32.5
83.7	0	4	20.9	75.4	0	10	7.5
14.1	0	1	14.1	191.9	0	8	24
41.5	0	1	41.5	34.6	0	2	17.3
7.5	0	2	3.8	17.7	1	1	17.7
15.1	0	3	5.0	73.6	0	4	18.4
11.3	0	2	5.7	88.6	0	4	22.2
29.3	0	5	5.9	8.0	1	1	8.0
82.9	0	1	82.9	1.6	0	3	.5
7.5	1	1	7.5	2.4	0	6	.4
20.8	0	2	10.4	18.9	0	1	18.9
52.8	0	1	52.8	11.1	0	1	11.1
51.8	1	1	51.8	14.9	0	1	14.9
65.3	0	1	65.3	13	0	1	13
66.2	0	1	66.2	7.3	0	1	7.3
65.2	0	2	32.6				

BIBLIOGRAPHY

1. Allen, D. C., Carlson, C. H., Hubach, C. E. Procedure for Reliability Assessment with Confidence. Federal Electric Corporation Document No. AS-A-86-67, June 1967.

2. Dalton, R. E. An Evaluation of Methods for Construction of Confidence Limits for System Reliability. TRW Systems, Florida Operations, Contract AF 04(694)-806, June 1966.

3. Donnelly, Patrick P. Is Reliability Necessary? Ninth Annual West Coast Reliability Symposium, Western Periodicals Company, North Hollywood, California, pp. 61-65, February 1968.

4. Esary, J.D., Proschan, Frank, and Marshall, A. W. Some Reliability Applications of the Hazard Transform. BSRL Mathematical Note No. 605, May 1969; to appear in SIAM J. Appl. Math.

5. Feller, William. An Introduction to Probability Theory and Its Applications, Volume II, John Wiley and Sons, 1966.

6. Lindley, D. V. Probability and Statistics, Volume II, Inference Cambridge University Press, 1965.

7. Madansky, Albert. Approximate Confidence Limits for the Reliability of Series and Parallel Systems, Technometrics 4, 495-503, 1965.

8. Mann, Nancy R. Confidence Bounds for System Reliability from Subsystem Data: A Survey of the State of the Art, presented at the Fifteenth Conference on the Design of Experiments in Army Research, Development and Testing. (To be published.)

9. Myhre, J. M., Saunders, S. C. On Confidence Limits for the Reliability of Systems, Ann. Math. Statist. 39, 1463-1472, 1968.

10. Myhre, J. M., and Saunders, S. C. Comparison of Two Methods of Obtaining Approximate Confidence Intervals for System Reliability, Technometrics 10, 37-50, 1968.

11. Saunders, S. C. On the Determination of a Safe Life for Distributions Classified by Failure Rate, Technometrics 10, 361-377, 1968.

12. Springer, M. D. and Thompson, W. E. Bayesian Confidence Limits for Reliability of Redundant Systems When Tests are Terminated at First Failure, Technometrics 10, 29-36, 1968.

13. Zimmer, W. J., Breipohl, A. M., and Prairie, R. R. A Consideration of the Bayesian Approach to Reliability Evaluation, IEEE Transactions on Reliability, R-14, No. 2.

A PROBABILITY MODEL FOR THE ASSESSMENT OF HUMAN INCAPACITATION FROM PENETRATING MISSILE WOUNDS

William P. Johnson
William J. Bruchey, Jr.
Ballistic Research Laboratories
Aberdeen Proving Ground, Maryland

ABSTRACT. A mathematical model is proposed for the evaluation of the altered performance of one of the most complex systems known to man, himself. Probabilities are associated with a random fragment penetrating varying distances within the human body, striking a critical anatomical component and inflicting damage to the extent that the wound recipient would be unable to perform his assigned duties.

These probabilities are combined to determine the conditional probability of all events occurring, simultaneously, to a tactical soldier under battlefield conditions.

The remainder of this article has been reproduced photographically from the authors' manuscript.

Preceding page blank

I. INTRODUCTION

To evaluate the effectiveness of antipersonnel weapons it has become necessary to employ a quantitative casualty criterion. A suitable criterion for many purposes is the probability that the weapon will wound, fatally or severely, its intended target.

An old criterion of wounding power is the 58 foot pound rule. In its crudest form it states that missiles with less than 58 foot pounds of kinetic energy do not kill, that those with more than 58 foot pounds of kinetic energy do kill. This criterion was never intended to be more than a rough rule of thumb. Burns and Zuckerman[1*] made a more refined analysis in 1941 of the quantitative requirements for wounding, while Guerney[2] suggested that MV^3 was a more suitable criterion than the kinetic energy standard for wounding human targets.

Wound ballistics work carried out in this country during and since World War II has included studies of the relationship between the volumes of cavities formed in tissues and tissue simulants by missiles and the physical parameters describing the projectile on impact with the target. Such work has not yet provided numerical relations between the physical parameters of fragments and the probability that the wounded man will suffer any specified degree of impairment of his ability to function. For the evaluation of the antipersonnel weapons, there is needed a knowledge of the numerical probability that a man, struck by a projectile of specified characteristics, will thereafter be unable to perform the functions of his tactical role. For evaluation or design, the size of a wound is not directly employable as a useful criterion. So far, cavity studies have unfortunately led to little or no information suitable for the evaluation or design of fragmenting weapons.

The Burns and Zuckerman studies led to numerical probabilities of the type needed for evaluation. A later numerical study of a similar type was reported by McMillen and Gregg of the Princeton Department of Biology, in Missile Casualties Report No. 12, 6 Nov 1945, National Research Council, Division of Medical Sciences. McMillen and Gregg were concerned with wounds that they considered to be either fatal or severe.

*Superscript numerals refer to references.

These could be caused, they assumed, by projectiles reaching certain vulnerable regions inside the body after traversing protective layers of skin, soft tissue, and bone provided that the projectiles reach the vulnerable regions with velocities in excess of 750 m/sec. The thicknesses of the anatomical structures that had first to be traversed were determined from anatomical charts depicting cross sections of the body at approximately one inch intervals from head to feet. The velocities necessary to penetrate the various thicknesses of skin, tissue and bone were determined through in vivo experimentation. The selection of vulnerable regions was somewhat arbitrary, since no experiments were conducted to determine which regions were in fact of primary importance. The final results, which were therefore semi-theoretical in nature, were in the form of probabilities that random hits with steel balls on a human target would cause fatal or serious wounds. These probabilities were plotted against the striking energies of the balls, and were on the whole consistent with the earlier conclusions of Burns and Zuckerman.

A generalization of quite a different type was published by T. E. Sterne[3] in May 1951. In that study the experimental data employed by McMillen and Gregg were re-examined and re-interpreted. In addition, the calculations of McMillen and Gregg were repeated for randomly shaped fragments similar to bomb fragments instead of spheres. The temporary cavities caused by the fragments penetrating the target were taken into account by requiring that a fragment, in order to cause important injury to a vital region, must not only penetrate the intervening skin, tissue, and bone, but must also reach the vulnerable regions with sufficient remaining energy to produce a temporary cavity of 2 cubic centimeters.

In October of 1956, "A New Casualty Criterion for Wounding by Fragments" was published by Allen and Sperrazza.[4] This study revised the combination of mass and velocity combinations which appeared to be related to incapacitation. Instead of employing the MV/A parameter used by Sterne, MV^{β}, where $1 < \beta < 2$, was introduced for the first time. This criterion has been used extensively to determine the relationship between trauma and fragment characteristics. However, close examination of the techniques used within the current model reveals that the resultant

quantitative data is based upon numerous subjective assumptions which vary from one evaluation to the next and that the resultant data are not probabilistic numbers in the mathematical sense.

With the advantage of the knowledge available in the area from earlier investigations, it is the purpose of the current study to establish a model that will supersede the currently used fragment criteria. The specific objectives of the proposed study are:

(1) To establish a truly probabilistic model for the assessment of human incapacitation from penetrating missile wounds.

(2) To eliminate the necessity for tissue retardation firings and in vivo experimentation as a prerequisite for the evaluation of the wounding potential of future fragments.

(3) To eliminate the need for trajectory tracings through anatomical cross sections of the human body in order to establish quantitative values for the probability of incapacitation.

(4) To establish a common basis for the comparison of all fragmenting munitions.

The approach and procedure for establishing the proposed model while accomplishing the above objectives, are presented in the following sections of this paper.

II. APPROACHES AND ASSUMPTIONS

For the purposes of establishing a mathematical model to predict the probability of incapacitation to a human upon impact from a random projectile the following assumptions are made:

(1) The fragmenting pattern of grenades and other exploding munitions have been studied in sufficient detail to provide the probability of a specific fragment striking a human target, P(H).

(2) Each fragment impacting the human target has a distinct probability of penetrating specific distances within the wound recipient's body, P(P).

(3) Each fragment striking the casualty has a distinct probability of encountering a critical organ along its path, P(E). This probability is conditioned by the penetrating ability of the fragment and the location of the critical components of the human anatomy.

(4) Each anatomical component has a distinct probability of being damaged, P(D), to the extent of preventing specific biomechanical motions required by a tactical soldier for the full performance of his mission.

(5) Each specific biomechanical motion has a distinct probability, P(M), of being required during the performance of the soldiers total mission.

(6) The probability of incapacitation from the N'th, P(N), component may be described by the expression:

$$P(N) = P(H) \cdot P(P/H) \cdot P(E/HP) \cdot P(D/HPE) \cdot P(M/HPED)$$

(7) Each of the component probabilities are independent and can be combined mathematically to provide the conditional probability of incapacitation, P(I/HPEDM), to a tactical soldier from a random projectile by the following:

$$P(I/HPEDM) = \frac{\sum_{1}^{N} P(N)}{N}$$

For the purpose of brevity, henceforth the conditional probability of incapacitation of a tactical soldier, P(I/HPEDM), will be shortened to P(I/H). It is to be understood that the second expression includes all of the conditions provided within the first expression.

(8) No synergistic effects occur from multiple wounds. Thus, the probability of incapacitation to a tactical soldier from two or more wounds can be determined by mathematically combining the independent conditional probabilities associated with each wound using the following expression:

$$P(I/H) = 1-(1-P(I/H_1))(1-P(I/H_2))(1-P(I/H_3)).....$$
$$(1-P(I/H_N))$$

As mentioned earlier, the probability of a fragment striking a stationary human target is assumed to have been established by Exterior Ballistics experts. Therefore, the remainder of this paper will concentrate on the procedures to be used to establish the probabilities required for the other portions of the model.

III. PROBABILITY, P(P), THAT PROJECTILE WILL PENETRATE A SPECIFIC DISTANCE, (D), WITHIN THE HUMAN BODY

As assumed by McMillan and Gregg, it seems reasonable to suppose that the probability that a random hit will cause fatal or severe wounding depends upon the fraction of the body's superficial area through which the fragment can wound a vital organ. The identification of the vital regions, and the conditions of striking them necessary to cause fatal or severe wounding may not have been correctly chosen by McMillen and Gregg, nevertheless, it still seems reasonable to suppose that the probability P(I/H) will be a function of the penetrating ability of a fragment. If fragments possess such great penetrating power that they traverse a body completely wherever they hit, then the probability P(I/H) that a random hit will incapacitate is the ratio of a rather large vulnerable area to the total presented area of the body. On the other hand, if the penetrating power of a fragment is so low that it can never reach the critical components of the body, then the probability P(I/H) must approach zero.[5]

It is proposed that an indication of the penetrating potential of fragments be obtained thru the use of facilities currently or soon to be available to wound ballisticians. These include (1) the Ballistic Research Laboratories Computer Man[6] and (2) striking versus residual velocity comparisons for each of the anatomical components of the human body.

A brief diversion is required at this point to acquaint the reader with the BRL Computer Man, in order that he may fully appreciate its potential for aiding in the solution of the current problem.

In brief, the BRL Computer Man is a computer program currently used within the wound ballistics program to determine the extent of incapacitation experienced by wounded tactical soldiers. It consists of coded

versions of human anatomical cross sections extracted from Eycleshymer and Schoemaker, "A Cross Section Anatomy." The combination of the cross sections represents a human male in a specified tactical position. Every major anatomical component illustrated in the original cross section anatomy has been coded within the computer version in approximately the same proportion as found in the published version.

The associated instructions accompanying the coded cross sections within the computer model permit simulated fragment paths to be traced through the individual cross sections at various impact angles. Retardation data is provided for each of the anatomical structures and is used to determine the velocity loss within each coded component as a function of the distance traversed by the fragment through the individual components. Thus the penetrating ability of each fragment becomes a function of (1) striking conditions at impact upon the cross section and (2) the retarding ability of the anatomical components encountered along its path.

Retardation data for several anatomical components has been used within the Wound Ballistics program for several years. However, the data in existence is of inadequate quality for the ultimate solution to our current problem. Fortunately, a program is currently being conducted by Sturdivan and Thompson[6] which should provide the necessary retardation data input to the BRL Computer Man.

Upon completion of the retardation studies, the data will be fed into the Computer Man model. Maximum distances traveled within the anatomical model will be determined as a function of several physical parameters of the impinging projectile (mass, velocity, presented area). Success or failure for each trajectory will be determined by its ability to penetrate at least each of several pre-specified distances within the human anatomy. The probability, P(G), of a particular fragment-velocity combination penetrating at least each of the pre-specified distances within an anatomical subdivision will be the ratio of the total number of successes at that distance to the total number of initiated trajectories not perforating the subdivision. It should be noted that the denominator becomes the number of non-perforating trajectories rather than the total

number of trajectories originated within a subdivision. This becomes necessary because of the varying distances within the human anatomy at which perforations will be achieved by most projectiles because of the variations in tissue structures encountered along each path. In addition, the geometry of the human anatomy is such that the distance available for penetration by any projectile becomes smaller as one moves farther from the center of the body.

If we let P' represent the probability of a projectile perforating the human anatomy, then (1-P') represents the probability that the projectile will stop within the human body. However, it is desired that probability figures be determined as a function of several distances within the non-perforating group. The probability, P(P), associated with each non-perforating projectile penetrating at least each of the prespecified distances will be given by the product of the probability of retention of the fragment within the human anatomy, (1-P'), and the probability of the fragment reaching the required distance, P(G).

Symbolically, this becomes: $P(P) = (1-P') \cdot P(G)$

There appears to be two alternatives for considering the perforating projectiles:

(1) It can be assumed that the penetration pattern of the non-perforating projectiles are representative of that which would have been displayed by the perforating projectiles, if the human anatomy was such that it allowed all projectiles to penetrate as deeply as possible. Based on this assumption, the penetration pattern as a function of distance would be represented by P(P) and the perforating projectile paths would only be used to condition the probability of the body retaining an impinging projectile.

(2) It can be assumed that the perforating fragment paths represent the upper limits of penetration and that because they do perforate, the fragments would be able to reach any desired depth within the human anatomy. Based on this assumption, the penetration pattern as a function of distance would be represented by the sum of P(P) and P'.

As of this writing it has not been decided which of the two alternatives will be used within the proposed model. However, since we are aware of the available options, it is felt that subsequent discussions with knowledgeable personnel will provide insight as to the proper approach.

Probability data of the above type will be obtained for several missile-velocity combinations. It is proposed that regression equations be used to relate the probability of penetration of specific missiles to a physical parameter of the projectile. Thus, generalized equations can be developed which will provide the user with the probability of a particular projectile penetrating a specified distance within a human as a function of its impact point on the anatomy and other relevant physical characteristics of the impinging projectile.

IV. PROBABILITY OF PROJECTILE ENCOUNTERING THE N'TH ANATOMICAL COMPONENT DURING ITS PATH THROUGH THE HUMAN BODY, P(E)

If it is assumed that a projectile will impact the human target randomly, then it must be assumed that the path of the projectile through the body will be random. Although the proposed model is designed to yield primarily, the probability of incapacitation from a single penetrating projectile, it should be realized that an almost infinite combination of organs or tissues can be encountered along the path of a single wound. For this reason one must deal with the probability of encountering each independent anatomical component within a given wound tract, rather than combinations of traumatized organs. The shielding effects afforded some organs by their surroundings, combined with the orientation of the organs to specific impact angles makes the ratio of their size to the total body area an inaccurate measure of the probability of encountering the organ.

In addition, because the geometry of most anatomical components vary as a function of their depth within the anatomy, probabilities of encountering each structure must be developed as a function of the obliquity angle of the impinging projectile and the target depth within the human body. However, for the purposes of this model, final probabilities will be documented as a function of penetration distance only. Several

obliquity angles will be used to determine the probability of encountering each structure as a function of its depth within the human anatomy. The final probability figures will reflect the weighted averages obtained from each of the independent analyses. Not only will this approach reduce the amount of bookkeeping involved within the project, it should also yield probabilities of encountering specific organs more representative of those expected from a truly random penetration.

The BRL Computer Man model will serve as the basis for the data obtained within this phase of the model. Imaginary trajectories will be traced through the coded cross sections and allowed to penetrate specific distances within the human anatomy. For each organ or tissue under consideration, a strike upon the structure within a specified distance will be recorded as a success. The probability of encountering each structure as a function of the specified distance will be the ratio of the total number of recorded successes to the total number of trajectories traced through the anatomy. The above will be documented at unit intervals for each anatomical component. The final probabilistic numbers will reflect the chances of having encountered each component with penetrations up to the indicated depth.

Subsequent additions to the proposed model will provide insight on the effects of missile "bite" on the probability of encountering components within the human anatomy. This physical phenomena considers the fact that a projectile need not actually strike a component in order to damage it. Investigations in these areas are currently underway within the Wound Ballistics program and once these data are available in quantitative form, the proposed model can be modified to include this phenomena.

Subsequent acquisition of modern anatomical cross sections reflecting the internal structure sizes of soldiers from specific military anthropometric percentiles should allow the probabilistic data generated during this subportion of the model to be generalized as a function of a physical measure of the human anatomy. No data exists of this type at the present time, therefore it is impossible to make such generalizations or even speculate as to how the probabilities of organ encounter are expected to vary between population percentiles.

V. PROBABILITY, P(D), THAT DAMAGE TO AN ANATOMICAL COMPONENT WILL PREVENT THE PERFORMANCE OF SPECIFIC BIOMECHANICAL MOTIONS

Before a missile wound can be evaluated in terms of the biomechanics required to perform a given task or role, it is necessary to describe the wound in terms which can be easily related to the physiological function of the body or subsystems within the body. That is, criteria must be developed which relates the probability of producing a given damage level to performance. This must be carried out in three steps. They are:

(1) Determine criteria to be used for damage description of the anatomical components.

(2) Determine the probability that missiles from a given munition will produce a specific damage level to a particular anatomical component.

(3) Determine the probability that this damage level will result in some biomechanical decrement.

At present work is being conducted under area (1). This task is oriented towards wound description in terms which can be related to biomechanical decrement. The criteria used to describe the wound will depend on the particular tissue or tissue type encountered. For instance, injury to muscles may be described in terms of hole size, percent muscle severed, etc; injury to blood vessels may be described in terms of the rate of blood flow from the injured location. Once these criteria have been determined, step (2) can be undertaken.

Presently, there exists a data bank of wound descriptions through animal tissue in terms of various parameters for several penetrating missiles. If, for instance, the criteria chosen is hole size, the wound data will be used to determine the distribution of hole size as a function of the missile parameters (mass, presented area, velocity). The information combined with the distribution of missile parameters for a given munition will provide the distribution of hole size for that munition. Using the probability distribution for hole size for a given munition an expected hole size will be computed, i.e. the

expected damage level, E(D). Likewise, if the criteria for damage was blood loss, a similar E(D) would be computed.

This value of E(D) would then be presented to a medical assessor who would subjectively determine the probability that the indicated damage level would prevent an individual from accomplishing a specific biomechanical motion.

It is recognized that this subjective input into the model prevents the final results from realizing complete objectivity. However, in order to quantify the effects of damage levels to various anatomical components, years and years of experimental laboratory work would be required. As of this writing little experimental work has begun. In view of the fact that the decision made by the medical assessors will be a function of the expected damage to an anatomical component, totally independent of the physical parameters of the impinging projectile and made only once within the lifetime of the model, (except perhaps to improve a previous decision), it is felt that the subjectivity will be minimized as much as possible.

VI. PROBABILITY, P(M), THAT A SPECIFIC BIOMECHANICAL MOTION WILL BE REQUIRED DURING THE PERFORMANCE OF A SOLDIER'S MISSION

Due to the wide variety of specialties among servicemen, it is almost impossible to generalize upon the duties of active combatants. However, basic to the performance of duties associated with every specialty is the ability of the individual to perform controlled movements or biomechanical motions. Regardless of the traumatized anatomical components or region, a soldiers ability to function under battlefield conditions is directly related to his ability to carry out controlled movements. If an individual's wound site is such that it does not hamper or impair his ability to make the necessary biomechanical motions, then for all practical purposes, the individual cannot be regarded as incapacitated. These controlled movements of the body may be resolved into the functioning of the following subsections:

I. Shoulder Girdle

II. Shoulder Joint

III. Elbow and Radio-Ulnar Joint

IV. Wrist and Hand

V. Pelvic Girdle and Hip Joint

VI. Knee Joint

VII. Ankle and Foot

VIII. Spinal Column

IX. Thorax

I. Movement of the Shoulder Girdle:

1. Adduction - movement of the scapula medially toward the spinal column.
2. Abduction - sliding of the scapula laterally and forward along the surface of the ribs.
3. Elevation/depression - the upward and downward motions of the whole scapula without any rotation.
4. Upward rotation - involves an upward turning of the glenoid cavity and the lateral angle in relation to the superior angle and medial border, which turn downward.
5. Downward rotation - reverse of upward rotation.
6. Foreward tilt - occurs when the inferior angle moves backward away from the rib cage.
7. Backward tilt - the inferior angle and the costal surface return to the surface of the rib cage.

II. Movement of the Shoulder Joint:

1. Flexion - foreward elevation of the arm.
2. Extension - return movement.
3. Abduction - sideward elevation of the arm.
4. Adduction - return movement.
5. Inward rotation - turning the humerous around its long axis so that its anterior aspect moves medially.

6. Outward rotation - the opposite with the anterior aspect moving laterally.

III. Movement of Elbow and Radio-Ulnar Joints:

A. Elbow Joint:

1. Flexion.
2. Extension.

B. Radio-Ulnar:

1. Pronation.
2. Supination.

IV. Movements of the Wrist and Hand:

A. Wrist Joint:

1. Abduction.
2. Adduction.
3. Circumduction.
4. Flexion.
5. Extension.

B. Hand:

1. Prehensile movement.
   a. power grip - an object is clamped by the partly flexed fingers and palm with counter pressure applied to the thumb lying more or less in the plane of the palm.
   b. precision grip - object is pinched between the fingers and the opposing thumb.
2. Nonprehensile movement - objects are manipulated by pushing or lifting.

V. Movement of the Pelvis Girdle and Hip-Joint:

A. Pelvis:

1. Forward rotation - increased inclination resulting from lumbo-sacral hyperextention.
2. Backward rotation - opposite movement.
3. Lateral tilt - the lowering or raising of one iliac crest.
4. Rotation - turning about a vertical axis either to the right or left.

B. Hip-Joint:

1. Flexion - forward movement of the femur.
2. Extension - reverse movement.
3. Abduction - movement of one limb away from the other toward the side.
4. Adduction - movement of one limb from the side towards the other.
5. Circumduction - movement of the limb in a circular manner. Combine movements 1 - 4.
6. Rotation - may be outward or inward depending on which way the toes are turned.

VI. Movements of the Knee Joint:

1. Flexion.
2. Extension.
3. Inward rotation.
4. Outward rotation.

VII. Movements of the Ankle and Foot:

A. Ankle and Foot:

1. Dorsiflexion - consists of raising the foot toward the anterior surface of the leg.
2. Plantar flexion - lowering the foot so as to bring its long axis in line with that of the leg.
3. Eversion - the sole is turned laterally or outward.
4. Inversion - sole is turned medially.

B. Toes:

1. Flexion.
2. Extension.

VIII. Movements of the Spinal Column:

A. Cervical Spine:

1. Flexion.
2. Extension.
3. Lateral flexion.
4. Rotation.

B. Thoracic and Lumbar Spines:

1. Flexion.
2. Extension.
3. Lateral flexion.
4. Rotation.

IX. Movements of the Thorax:

1. Elevation.
2. Depression.

As can be seen from the above the possibility of a wide variety of body movements exists for every individual. However, we are concerned only with the essential or necessary movements involved in the tasks of combatants.

It is assumed that every task, regardless of its complexity, can be resolved into a series of controlled movements of the type presented above. It is further assumed that a probabilistic figure which reflects the chances of an individual being required to perform a specific biomechanical motion during the course of his duties can be determined. If an individual is required to assume multiple duties, then the probability of his performing a specific biomechanical motion will be conditioned by the probability of his performing the duty which requires the motion. Thus, the evaluation of a soldier's incapacitation can be related to several duties required within an overall mission.

In order to obtain the desired probabilities, one must be knowledgeable of the specific duties required of today's soldier. In addition, these duties must be analyzed from the kinesiologist point of view in order to reduce them to the series of independent biomechanical motions required for probabilistic analysis.

Once the duties have been reduced to a series of controlled movements, two alternatives for determining the probability, P(M), that a specific biomechanical motion will be required during the performance of a soldier's duties or mission appear to exist. The first assumes that each motion is independent of time and that every motion involved within a specific task requires the same percentage of the soldier's time. This does not eliminate the possibility of repetitions of the same

movement within a given task. Repetitions are accounted for by including each repetition as a separate component of the total sum of movements involved. The probability, P(M), of a particular motion being required within the task will be the ratio of the total repetitions of the movement within the task to the total number of independent movements involved.

The second approach considers the total time required by the combatant to perform his duties as the unit and the probability, P(M), is determined by the functional part of the total allocated to each independent motion.

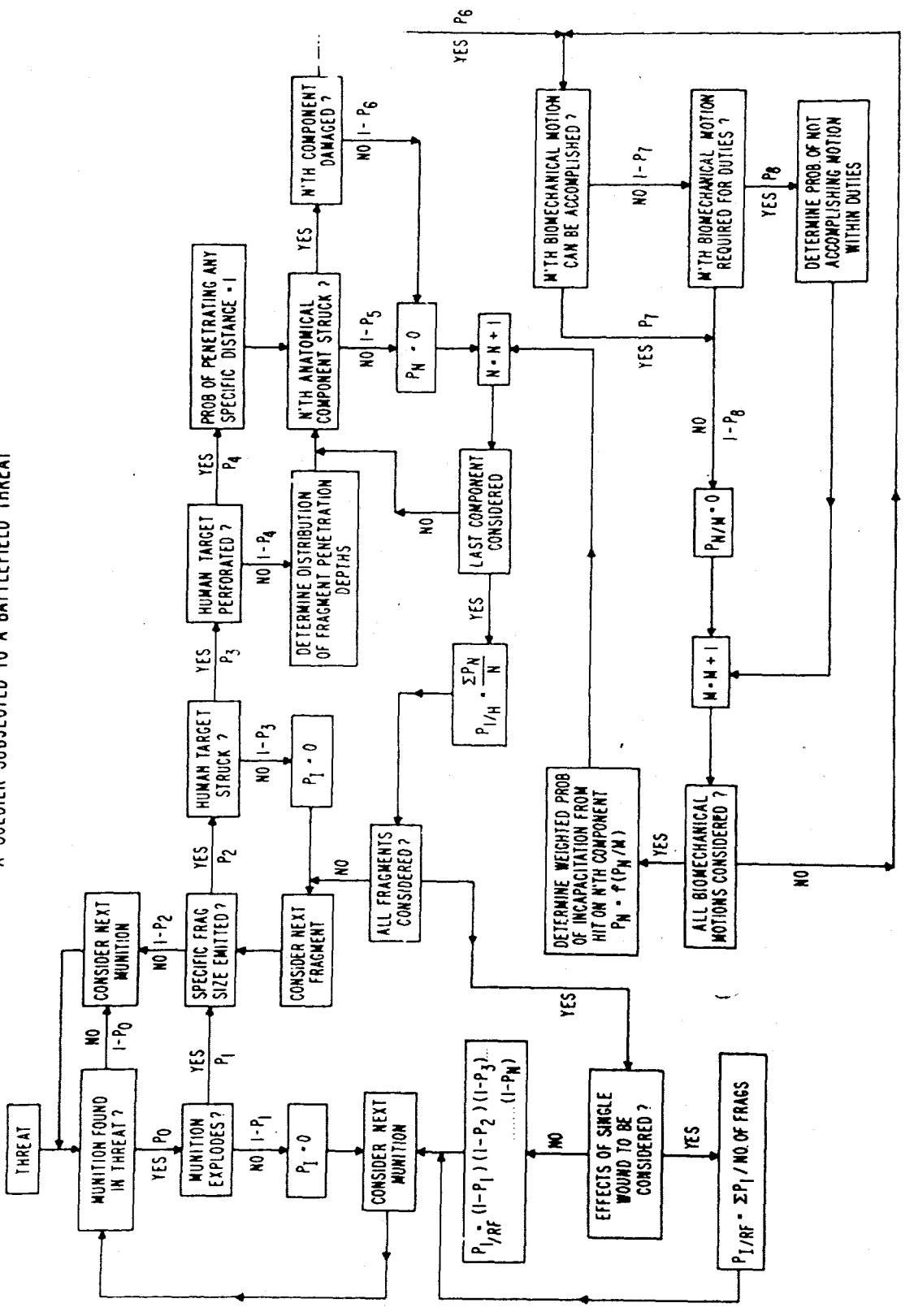
Each of the two approaches has advantages and disadvantages obvious to the authors at this point. However, suffice it to say that the merits of either approach will be discussed in detail with the experts on the subject prior to deciding which approach is most feasible. It appears that either of the alternatives will yield probabilistic data of the type required for input into the model. Therefore, no serious problems appear to exist.

VII. SUMMARY

A mathematical model for assessing the probability of human incapacitation from a penetrating missile wound has been proposed. The model assumes that several probabilistic events must occur, sequentially, in order to cause incapacitation to a tactical soldier. Among these events are (1) the soldier must be hit, (2) the impacting projectile must penetrate deep enough into the human anatomy to strike a critical anatomical component, (3) the anatomical component must be damaged to the extent that its physiological function is impaired, and (4) the dysfunction of the component must be directly related to the soldier's ability to perform specific biomechanical motions required for the performance of the soldier's task.

An approach has been given for obtaining the necessary experimental data needed to relate each of the above events in a probabilistic manner. In addition, the assumptions and anatomical equations used within the model have been detailed.

LOGIC FOR DETERMINING THE PROBABILITY OF INCAPACITATING A SOLDIER SUBJECTED TO A BATTLEFIELD THREAT



Best Available Copy

REFERENCES

1. B. D. Burns and S. Zuckerman, "The Relationship Between Striking Velocity and the Damage Caused by a 3/32 Inch Steel Ball," Great Britain, Ministry of Home Security, R. C. Report No. 232, 1941.

2. R. W. Gurney, "A New Casualty Criterion," Ballistic Research Laboratories Report No. 498, 31 October 1944.

3. T. E. Sterne, "A Provisional Casualty Criterion," Ballistic Research Laboratories Technical Note No. 370, March 1957.

4. F. Allen and J. Sperrazza, "New Casualty Criteria for Wounding By Fragments," Ballistic Research Laboratories Report No. 996, October 1956.

5. T. E. Sterne, "Provisional Criteria for Fatal or Severe Wounding By Fragments," Ballistic Research Laboratories Memorandum Report No. 591, February 1952.

6. R. Kneibert et al, "Wound Tract Analysis by High Speed Digital Computer," Ballistic Research Laboratories Report No. 1623, January 1965.

7. L. Sturdivan, J. Thompson, J. Terner, "A Parametric Study of Bio-Retardation," paper presented at Army Design of Experiments Conference, April 1969.

LIST OF ATTENDEES

Carmelo Aliano, Redstone Arsenal, Alabama
Bernard Alley, Redstone Arsenal, Alabama
Hubert W. Anderson, Redstone Arsenal, Alabama
Edward J. Arber, Picatinny Arsenal, New Jersey
Elwood D. Baas, White Sands Missile Range, New Mexico
Rolf Bargmann, University of Georgia, Athens, Georgia
Garry Barnard, Redstone Arsenal, Alabama
B. W. Barnett, Redstone Arsenal, Alabama
Edwin M. Bartee, Vanderbilt University, Nashville, Tennessee
Hans Baussus von Luetzow, Fort Belvoir, Virginia
Robert Bechhofer, Cornell University, Ithaca, New York
Alan W. Benton, Aberdeen Proving Ground, Maryland
Joe Bigger, Redstone Arsenal, Alabama
Novella S. Billions, Redstone Arsenal, Alabama
Erwin Biser, Fort Monmouth, New Jersey
Mary E. Blome, Ammunition Proc. and Supply Agency, Joliet, Illinois
Joseph I. Bluhm, Army Materials & Mech. Res Lab, Watertown, Mass
E. L. Bombara, NASA, Huntsville, Alabama
D. G. Boyd, Redstone Arsenal, Alabama
Helen M. Boyd, Redstone Arsenal, Alabama
Wallace Branham, Redstone Arsenal, Alabama
Ferrell Bryan, Redstone Arsenal, Alabama
O. P. Bruno, Aberdeen Proving Ground, Maryland
Edwin L. Buchanan, Redstone Arsenal, Alabama
Leo F. Buldhaupt, The Boeing Company, Huntsville, Alabama
Joseph M. Cameron, National Bureau of Standards, Washington, D. C.
D. Ray Campbell, Redstone Arsenal, Alabama
Vera M. Campbell, Redstone Arsenal, Alabama
Ola Carter, Redstone Arsenal, Alabama
Richard W. Clarke, Watervliet Arsenal, New York
A. C. Cohen, Jr., University of Georgia, Athens, Georgia
Charles Colvin, Redstone Arsenal, Alabama
Robert G. Conard, Redstone Arsenal, Alabama
John E. Condon, NASA Headquarters, Washington, D. C.
Gideon A. Culpepper, White Sands Missile Range, New Mexico
C. Alan Cummings, Redstone Arsenal, Alabama
Richard D'Accardi, Fort Monmouth, New Jersey
J. A. Davies, General Electric, Huntsville, Alabama
Carmine DeSanctus, NASA, Huntsville, Alabama
Henry A. Dihm, Jr., Redstone Arsenal, Alabama
James M. Donnini, The Boeing Company, Huntsville, Alabama
Francis G. Dressel, Army Research Office-Durham, North Carolina
Eugene Dutoit, Army Munitions Command, Dover, New Jersey
Alan R. Dyer, Aberdeen Proving Ground, Maryland
Churchill Eisenhart, National Bureau of Standards, Washington, D. C.
Oskar Essenwanger, Redstone Arsenal, Alabama
Wade H. Ewart, Redstone Arsenal, Alabama
Tom G. Feaster, The Boeing Company, Huntsville, Alabama

Preceding page blank

Walter T. Federer, University of Wisconsin MRC, Madison, Wisconsin
Lewis S. Fichter, Picatinny Arsenal, Dover, New Jersey
Thomas O. Finley, Fort Ord, California
Walter D. Foster, Fort Detrick, Maryland
Fred Frishman, Office, Chief of Research and Development, Washington, D.C.
A. S. Galbraith, Army Research Office-Durham, North Carolina
Sidney Gerard, Aberdeen Proving Ground, Maryland
John P. Gibbons, Redstone Arsenal, Alabama
Tom Gill, University of Alabama, Tuscaloosa, Alabama
Dieter Grau, NASA, Huntsville, Alabama
Bruce C. Gray, Fort Detrick, Maryland
Frank E. Grubbs, Aberdeen Proving Ground, Maryland
William V. Gudaitis, Redstone Arsenal, Alabama
John Gurland, MRC, University of Wisconsin, Madison, Wisconsin
Bernard S. Gurman, Fort Monmouth, New Jersey
William B. Harrell, NASA, Huntsville, Alabama
John S. Hagan, Aberdeen Proving Ground, Maryland
Bernard Harris, MRC, University of Wisconsin, Madison, Wisconsin
Boyd Harshbarger, Virginia Polytechnic Institute, Blacksburg, Virginia
Gary Hitchcock, Redstone Arsenal, Alabama
Earl G. Hoard, MASA, Huntsville, Alabama
Chester L. Kopkins, Redstone Arsenal, Alabama
John W. Howerton, Redstone Arsenal, Alabama
David R. Howes, STAG, Bethesda, Maryland
Edmund H. Inselmann, U. S. Army Materiel Command, Washington, D. C.
Jack Irwin, Redstone Arsenal, Alabama
Garhard J. Isaac, Fitzsimons General Hospital, Denver, Colorado
Robert F. James, Redstone Arsenal, Alabama
William R. James, Coastal Engineering Research Center, Washington, D. C.
Toke Jayachandran, Fort Ord, California
Emil H. Jebe, Willow Run Laboratory, Willow Run, Michigan
Andrew H. Jenkins, Redstone Arsenal, Alabama
Jerome R. Johnson, Aberdeen Proving Ground, Maryland
Larry H. Johnson, Redstone Arsenal, Alabama
Wallace E. Jordan, NASA, Huntsville, Alabama
George L. Kinnett, Army Aviation Material Labs, Fort Eustis, Virginia
Raymond V. Knox, Redstone Arsenal, Alabama
Richard Krinsky, University of Alabama, University, Alabama
Richard G. Krutchkoff, Virginia Polytechnic Institute, Blacksburg, Virginia
Harry H. Ku, National Bureau of Standards, Washington, D. C.
William A. Lampe, Redstone Arsenal, Alabama
Raymond E. Lacy, Fort Monmouth, New Jersey
George I. Lavin, Aberdeen Proving Ground, Maryland
Patrick B. Lawler, Jr., Redstone Arsenal, Maryland
Siegfried H. Lehnigk, Redstone Arsenal, Alabama
George A. Lineberry, Redstone Arsenal, Alabama
Joseph Lovinger, Rohm and Haas Company, Huntsville, Alabama
H. L. Lucas, North Carolina State University, Raleigh, North Carolina
Bob B. Lukens, University of Alabama, Huntsville, Alabama

Hat V. Maddix, Redstone Arsenal, Alabama
Clifford J. Maloney, National Institute of Health, Bethesda, Maryland
Nancy R. Mann, North American Rockwell Corporation, Canoga Park, California
Nathan Mantel, National Cancer Institute, Bethesda, Maryland
Stanley M. Martin, National Communicable Disease Center, Atlanta, Georgia
John L. McDaniel, Redstone Arsenal, Alabama
Claudia K. McDonald, Redstone Arsenal, Alabama
William B. McIntosh, Aberdeen Proving Ground, Maryland
Ray L. Miller, Jr., Shaw Air Force Base, South Carolina
James N. Mitchell, Safeguard Logistics Command, Huntsville, Alabama
James R. Moore, Aberdeen Proving Ground, Maryland
Oskar Morgenstern, MATHEMATICA, Princeton, New Jersey
Jacques Naar, Stanford Research Institute, Menlo Park, California
William Nichols, Redstone Arsenal, Alabama
Henry Nocke, SPACO, Inc., Huntsville, Alabama
William H. Nuckols, III, Redstone Arsenal, Alabama
Richard L. Nutt, Redstone Arsenal, Alabama
Robert R. Orr, Redstone Arsenal, Alabama
Bernard Ostle, Florida Technological University, Orlando, Florida
Tsetsy Pangarova, Redstone Arsenal, Alabama
Albert Parent, Thiokol Chemical Corp., Huntsville, Alabama
Gene B. Parrish, Army Research Office-Durham, North Carolina
R. H. Peterson, Aberdeen Proving Ground, Maryland
Charles A. Pollard, Jr., Redstone Arsenal, Alabama
Nixon Powell, Redstone Arsenal, Alabama
Aubrey Presson, Redstone Arsenal, Alabama
Robert C. Raybold, National Bureau of Standards, Washington, D. C.
Vernon H. Rechmeyer, Thiokol Chemical Corp., Huntsville, Alabama
Nancy R. Rich, Redstone Arsenal, Alabama
B. A. Richardson, Canadian Forces Headquarters, Ottawa, Ontario, Canada
Carol D. Rose, Army Tank Automotive Command, Warren, Michigan
Don Royston, Thiokol Chemical Corp., Huntsville, Alabama
Doyce Satterfield, Redstone Arsenal, Alabama
Sam C. Saunders, Boeing Scientific Res. Lab., Seattle, Washington
Raymond B. Schnell, Army Materiel Command, Washington, D. C.
George Schroster, Aberdeen Proving Ground, Maryland
Gordon Satterland, Boeing Company, Huntsville, Alabama
Harold C. Sebring, STANO Systems Office, Washington, D. C.
Seymour M. Selig, Department of the Navy, Washington, D. C.
R. E. Shannon, University of Alabama, Huntsville, Alabama
Chester E. Sharp, Fort Monmouth, New Jersey
Ralph Simanson, Redstone Arsenal, Alabama
Leslie E. Simon, USA Ret, Winter Park, Florida
Orvel E. Smith, NASA, Huntsville, Alabama
Royce W. Soanes, Watervliet Arsenal, New York
Earl B. Stewart, Redstone Arsenal, Alabama
R. G. Stimson, Department of the Army, Washington, D. C.
Glenn A. Stoops, Fort Ord, California
James F. Strickland, Jr., Aberdeen Proving Ground, Maryland

Donald H. Strietzel, Safeguard System Command, Huntsville, Alabama
William C. Taylor, Aberdeen Proving Ground, Maryland
Jerry Thomas, Aberdeen Proving Ground, Maryland
R. M. Thrall, Rice University, Houston, Texas
Myrna Toivanen, Honeywell, Inc., Minneapolis, Minn.
Joseph Weinstein, Fort Monmouth, New Jersey
Burchard E. Wheatley, The Boeing Company, Huntsville, Alabama
Robert D. Williams, Stanford Research Institute, Menlo Park, California
Rosalie C. Woodall, Harry Diamond Laboratories, Washington, D. C.
W. J. Youden, National Bureau of Standards, Washington, D. C.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

ORIGINATING ACTIVITY (Corporate author) U. S. Army Research Office-Durham Box CM, Duke Station Durham, North Carolina 27706	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP N/A

3. REPORT TITLE

PROCEEDINGS OF THE FIFTEENTH CONFERENCE ON THE DESIGN OF EXPERIMENTS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Interim Technical Report

5. AUTHOR(S) (First name, middle initial, last name)

6. REPORT DATE July 1970	7a. TOTAL NO. OF PAGES 882	7b. NO. OF REFS
8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S) ARO-D Report 70-2	
b. PROJECT NO.		
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.		

10. DISTRIBUTION STATEMENT

This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of the U. S. Army Research Office - Durham.

11. SUPPLEMENTARY NOTES None	12. SPONSORING MILITARY ACTIVITY Army Mathematics Steering Committee on behalf of the Office of the Chief of Research and Development

13. ABSTRACT

This is the technical report resulting from the Fifteenth Conference on the Design of Experiments in Army Research, Development and Testing. It contains 35 papers presented at that meeting. These treat various Army statistical and design problems.

14. KEY WORDS:

reliability	manpower planning
performance capability	disconnected block designs
estimation of variance	design of field tests
quality control	infrasonic waves
statistical models	least squares
Multiple triangulation	assay procedures
interlaboratory studies	linear programming
Bayesian confidence bounds	catastrophic threat
real-time simulation	empirical Bayes methods
continuous sampling	man-machine performance
weapons effectiveness	orthogonal latin squares
air defense systems	confidence limits

DD FORM 1 NOV 68 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

unclassified

Security Classification